

MANAGERIAL ECONOMICS

Economic Tools for Today's Decision Makers

SEVENTH EDITION





Paul G. Keat | Philip K.Y. Young | Stephen E. Erfle

Interactive Spreadsheet Modules and Excel Apps

Keat, Young, and Erfle's *Managerial Economics*, Seventh Edition, is accompanied by two different kinds of Microsoft Excel® learning aids. Interactive Spreadsheet Modules are study aids intended to supplement and reinforce concepts in the text. They can be used in class, assigned as homework, or for student exploration. The index of modules on the opposite page contains the number of the module, the topic covered, and the page in the textbook to which the module is relevant.

Excel Apps challenge students to build various economic models from data provided. They also allow readers to turn the static figures and tables in the text into dynamic illustrations, which will strengthen students' understanding of the topics covered and their ability to use Excel. A list of the Excel Apps can be found on the inside back cover of this text.

In the textbook you will see the  icon or the  icon in the margin indicating Interactive Spreadsheet Modules or Excel Apps are available for that topic.

Getting Started with Interactive Spreadsheet Modules

- Step 1)** Students and faculty can download Interactive Spreadsheet Modules at the Keat, Young, and Erfle Companion Website: www.pearsonhighered.com/keat.
- Step 2)** Open the Interactive Spreadsheet Module, and read through the objective and brief description.
- Step 3)** Analyze the solved example in its current form. (Several modules include graphs. To view a graph, press GRAPH at the bottom of the spreadsheet. To return to the text, press name of module at the bottom of the spreadsheet.)
- Step 4)** Enter new data to create alternative outcomes. (Only the cells into which data can be entered are unprotected. All other cells are protected.)
- Step 5)** Analyze the new outcome.

Module No.	Interactive Spreadsheet Module	Page
------------	--------------------------------	------

3A	Supply and demand curves	46
4A	Arc elasticity	70
4B	Point and arc elasticities	80
4C	Elasticity calculation applications	80
5A	Compound growth rate	139
6A	Production functions: quantities and dollars	190
6B	Production functions in dollars (total graph)	190
6C	Production functions in dollars (unit graph)	190
8A	Output and price under perfect competition	258, 320
8B	Output and price under imperfect competition (total graph)	329
8C	Output and price under imperfect competition (unit graph)	329
8BA	Volume-cost-profit (price and unit costs)	347
8BB	Volume-cost-profit (solving for the missing variable)	347
8BC	Volume-cost-profit (revenue and total cost)	348
8BD	Volume-cost-profit (comparing two companies)	350
10A	Revenue maximization model	398
12A	Capital budgeting for mutually exclusive projects	467
12B	Expansion project	470
12C	Replacement project	470
12D	Valuation of common stock with non-constant growth	473
12E	Weighted cost of capital	475
12F	Expected value and standard deviation	479
12G	Expansion project and scenario analysis for chapter "Solution"	495
13A	Transfer pricing	523
13B	Multinational corporate budgeting	527
TVM1	Time value of money	online
TVM2	Annuity calculations	online
TVM3	Bond value calculations	online

MANAGERIAL ECONOMICS

This page intentionally left blank

SEVENTH EDITION

MANAGERIAL ECONOMICS

Economic Tools for Today's
Decision Makers

Paul G. Keat

Thunderbird School of Global Management

Philip K. Y. Young

Nth Degree Systems, Inc. and Duke Corporate Education

Stephen E. Erfle

Dickinson College

PEARSON

Boston Columbus Indianapolis New York San Francisco Upper Saddle River
Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montreal Toronto
Delhi Mexico City São Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo

Editor in Chief: Donna Battista
Executive Acquisitions Editor: Adrienne D'Ambrosio
Editorial Project Manager: Sarah Dumouchelle
Executive Marketing Manager: Lori DeShazo
Managing Editor: Jeff Holcomb
Senior Production Project Manager: Nancy Freihofer
Operations Specialist: Carol Melville
Cover Designer: Bruce Kenselaar
Cover Images: Globe Photo: Thank You/Shutterstock; Bottles Photo: Holbox/Shutterstock
Full-Service Project Management: Tiffany Rupp
Composition: S4Carlisle Publishing
Printer/Binder: Edwards Brothers
Cover Printer: Lehigh-Phoenix Color/Hagerstown
Text Font: ITC New Baskerville Std 10/12

Credits and acknowledgments borrowed from other sources and reproduced, with permission, in this textbook appear on this copyright page and on the appropriate page within text.

p. 55, Figure 3.9: Charts © 2011 Deloitte Global Services Limited (“DGSL”), used with permission. Charts incorporate preexisting information from BP Statistical Review of World Energy, June 2010 edition (based on 2009 year-end data). DGSL has no connection or affiliation with BP or the BP Statistical Review of World Energy publication. Please see the copyright page for the full attribution and disclaimer notices pertaining to these charts. These charts, which are reproduced from a DGSL publication, contain general information only, and none of DGSL, Deloitte Touche Tohmatsu Limited, its member firms, or their related entities (collectively the “Deloitte Network”) is, by means of these charts, rendering professional advice or services. Before making any decision or taking any action that may affect your finances or your business, you should consult a qualified professional adviser. No entity in the Deloitte Network shall be responsible for any loss whatsoever sustained by any person who relies on these charts and no entity in the Deloitte Network is making any representation or warranty of any kind. Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee, and its network of member firms, each of which is a legally separate and independent entity. Please see www.deloitte.com/about for a detailed description of the legal structure of Deloitte Touche Tohmatsu Limited and its member firms. “Deloitte”, “Touche”, “Tohmatsu”, “Deloitte Touche Tohmatsu”, “Deloitte & Touche”, the Deloitte logo, and the Deloitte Touche Tohmatsu logo are trademarks or registered trademarks of the Deloitte Network, which has no connection to the author or publisher of this book and has no responsibility for its contents.

Copyright © 2013, 2009, 2006 by Pearson Education, Inc. All rights reserved. Manufactured in the United States of America. This publication is protected by Copyright, and permission should be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. To obtain permission(s) to use material from this work, please submit a written request to Pearson Education, Inc., Permissions Department, One Lake Street, Upper Saddle River, New Jersey 07458, or you may fax your request to 201-236-3290.

Many of the designations by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and the publisher was aware of a trademark claim, the designations have been printed in initial caps or all caps.

Library of Congress Cataloging-in-Publication Data

Keat, Paul G.

Managerial economics : economic tools for today's decision makers / Paul G. Keat, Philip K.Y. Young, Stephen E. Erfle. – 7th ed.

p. cm.

ISBN 978-0-13-302026-7

1. Managerial economics. I. Young, Philip K. Y. II. Erfle, Stephen E. (Stephen Eugene) III. Title.

HD30.22.K39 2014

338.5024'658–dc23

2012038213

PEARSON

10 9 8 7 6 5 4 3 2 1
ISBN 10: 0-13-302026-6
ISBN 13: 978-0-13-302026-7

To my wife, Sheilah, our children, Diana and Andrew, and our seven grandchildren—P. G. K.

To my grandchildren, Hayden and Emerson—P. K. Y. Y.

To my children, David, Kate, and Vera—S. E. E.

This page intentionally left blank

Brief Contents

Preface *xix*

About the Authors *xxv*

Chapter 1	Introduction	1
Chapter 2	The Firm and Its Goals	18
Chapter 3	Supply and Demand	38
	Appendix 3A The Mathematics of Supply and Demand	65
Chapter 4	Demand Elasticity	68
	Appendix 4A Applications of Supply and Demand	100
Chapter 5	Demand Estimation and Forecasting	114
	Appendix 5A The Demand for White Zinfandel in Los Angeles	167
	Appendix 5B Understanding Consumer Behavior Through Testing	175
Chapter 6	The Theory and Estimation of Production	180
	Appendix 6A Productivity in Services	220
	Appendix 6B The Multiple-Input Case	231
	Appendix 6C Analyzing Production Functions with the Use of Calculus	242
Chapter 7	The Theory and Estimation of Cost	248
	Appendix 7A A Mathematical Restatement of the Short-Run Cost Function	291
	Appendix 7B The Estimation of Cost	295
	Appendix 7C Interview with a Supply Chain Management Executive	306
Chapter 8	Pricing and Output Decisions: Perfect Competition and Monopoly	310
	Appendix 8A The Use of Calculus in Pricing and Output Decisions	343
	Appendix 8B Break-Even Analysis (Volume-Cost-Profit)	345
Chapter 9	Pricing and Output Decisions: Monopolistic Competition and Oligopoly	361
	Appendix 9A A Mathematical Restatement of Monopolistic Competition	387
Chapter 10	Special Pricing Practices	390
Chapter 11	Game Theory and Asymmetric Information	432
Chapter 12	Capital Budgeting and Risk	461
	Appendix 12A The Value of a Corporation	507

Chapter 13	The Multinational Corporation in a Global Setting	509
Chapter 14	Government and Industry: Challenges and Opportunities for Today's Manager	533
Chapter 15	The Global Soft Drink Industry	556
<i>Appendix A</i>	<i>Statistical and Financial Tables</i>	<i>563</i>
<i>Index</i>		<i>579</i>

Contents

Preface *xix*

About the Authors *xxv*

CHAPTER 1 Introduction 1

Introduction: Economics and Managerial Decision Making	2
A Brief Review of Important Economic Terms and Concepts	5
The Case of Global Foods, Inc.: Situations and Solutions	9
Summary of the Situations and Solutions	11
Global Application: The BRIC Countries	14
Summary	15
<i>Important Concepts</i>	16
<i>Questions</i>	16

CHAPTER 2 The Firm and Its Goals 18

Introduction	19
The Firm	19
<i>Coase and the Internet</i>	22
The Economic Goal of the Firm and Optimal Decision Making	22
Goals Other Than Profit	23
<i>Economic Goals</i>	23
<i>Noneconomic Objectives</i>	25
Do Companies Really Try to <i>Maximize Profits</i> ?	26
<i>Profit Maximization, Restated</i>	28
Maximizing the Wealth of Stockholders	29
<i>Market Value Added and Economic Value Added</i>	31
Economic Profits	32
Global Application	34
Summary	35
<i>Important Concepts</i>	35
<i>Questions</i>	36

CHAPTER 3 Supply and Demand 38

Introduction	39
Market Demand	39
Market Supply	42
Market Equilibrium	44
Comparative Statics Analysis	46
<i>Short-Run Market Changes: The “Rationing Function” of Price</i>	46
<i>Long-Run Market Analysis: The “Guiding” or “Allocating Function” of Price</i>	48
<i>Using Supply and Demand in Forecasting</i>	52
Supply, Demand, and Price: The Managerial Challenge	53
Global Application: The BRIC Countries and the Supply and Demand for Oil	53

Summary	56
<i>Important Concepts</i>	56
<i>Questions</i>	57
<i>Problems</i>	58
Appendix 3A The Mathematics of Supply and Demand	65
CHAPTER 4 Demand Elasticity	68
The Economic Concept of Elasticity	69
The Price Elasticity of Demand	69
<i>Measurement of Price Elasticity</i>	70
<i>The Determinants of Elasticity</i>	75
<i>The Effect of Elasticity on Price and Quantity</i>	77
<i>The Elasticity of Derived Demand</i>	78
<i>Elasticity in the Short Run and in the Long Run</i>	79
<i>Demand Elasticity and Revenue</i>	80
<i>The Mathematics of Elasticity and Revenue</i>	83
<i>Empirical Elasticities</i>	84
The Cross-Price Elasticity of Demand	86
<i>Empirical Elasticities</i>	87
Income Elasticity	87
Other Elasticity Measures	90
Elasticity of Supply	90
Global Application: Price Elasticities in Asia	91
Summary	93
<i>Important Concepts</i>	94
<i>Questions</i>	94
<i>Problems</i>	96
Appendix 4A Applications of Supply and Demand	100
CHAPTER 5 Demand Estimation and Forecasting	114
Demand Estimation	115
Introduction	115
<i>Key Chapter Objectives</i>	115
<i>The Critical Importance of Good Data</i>	116
Introduction to Regression Analysis	117
<i>Specifying the Regression Equation and Obtaining the Data</i>	117
<i>Estimating and Interpreting the Regression Coefficients</i>	120
<i>Statistical Evaluation of the Regression Results</i>	122
<i>Review of Key Steps for Analyzing Regression Results</i>	124
<i>Implications of Regression Analysis for Management Decisions</i>	125
Problems in the Use of Regression Analysis	125
<i>The Identification Problem</i>	126
<i>Multicollinearity</i>	127
<i>Autocorrelation</i>	128
Examples of Regression Analysis Across the Disciplines	129
Global Application: Food in Spain, Cigarettes in Taiwan	130
Forecasting	132
Introduction	132
Subjects of Forecasts	132
<i>Demand Estimating and Demand Forecasting</i>	133

Prerequisites of a Good Forecast	133
Forecasting Techniques	133
<i>Expert Opinion</i>	135
<i>Opinion Polls and Market Research</i>	135
<i>Surveys of Spending Plans</i>	136
<i>Economic Indicators</i>	136
<i>Projections</i>	139
<i>Econometric Models</i>	153
Global Application: Forecasting Exchange Rates	156
Summary	159
<i>Important Concepts</i>	159
<i>Questions</i>	161
<i>Problems</i>	162
Appendix 5A The Demand for White Zinfandel in Los Angeles	167
Appendix 5B Understanding Consumer Behavior Through Testing	175
CHAPTER 6 The Theory and Estimation of Production	180
The Production Function	181
A Short-Run Analysis of Total, Average, and Marginal Product	183
<i>The Law of Diminishing Returns</i>	185
<i>The Three Stages of Production in the Short Run</i>	188
<i>Derived Demand and the Optimal Level of Variable Input Usage</i>	190
<i>The Case of Multiple Inputs (Abridged Version)</i>	191
The Long-Run Production Function	193
The Estimation of Production Functions	196
<i>The Various Forms of a Production Function</i>	196
<i>The Cobb-Douglas Production Function</i>	198
<i>Statistical Estimation of Production Functions</i>	200
<i>Aggregate Production Functions</i>	203
The Importance of Production Functions in Managerial Decision Making	205
<i>Careful Planning Can Help a Firm to Use Its Resources in a Rational Manner</i>	205
Going “Beyond the Curves”: Current Production Issues and Challenges for Today’s Managers	207
Call Centers: Applying the Production Function to a Service	209
Global Application: Shifting Trends in Global Outsourcing	210
Summary	211
<i>Important Concepts</i>	212
<i>Questions</i>	213
<i>Problems</i>	214
Appendix 6A Productivity in Services	220
Appendix 6B The Multiple-Input Case	231
Appendix 6C Analyzing Production Functions with the Use of Calculus	242
CHAPTER 7 The Theory and Estimation of Cost	248
The Importance of Cost in Managerial Decisions	249
The Definition and Use of Cost in Economic Analysis	251
<i>Historical Versus Replacement Cost</i>	251

<i>Opportunity Cost Versus Out-of-Pocket Cost</i>	251
<i>Sunk Versus Incremental Cost</i>	252
The Relationship Between Production and Cost	253
The Short-Run Cost Function	255
<i>Increasing Cost Efficiency in the Short Run</i>	257
<i>Alternative Specifications of the Total Cost Function</i>	258
The Long-Run Cost Function	260
<i>The Relationship Between Long-Run Production and Long-Run Cost</i>	260
<i>Economies of Scale</i>	263
<i>The Long-Run Average Cost Curve as the Envelope of Short-Run Average Cost</i>	265
<i>Using Long-Run Average Cost as a Decision-Making Tool: The Importance of Coordinating Production Plans with Market Forecasts</i>	267
The Learning Curve	269
Economies of Scope	272
Economies of Scale: The Short Run Versus the Long Run	273
Supply Chain Management	273
Examples of Ways Companies Have Cut Costs to Remain Competitive	276
Cautionary Note to Managers About the Use of Cost-Cutting as a Strategy	279
Global Applications: Li & Fung Will Do It All for You	280
Summary	282
<i>Important Concepts</i>	282
<i>Questions</i>	283
<i>Problems</i>	285
Appendix 7A A Mathematical Restatement of the Short-Run Cost Function	291
Appendix 7B The Estimation of Cost	295
Appendix 7C Interview with a Supply Chain Management Executive	306

CHAPTER 8 Pricing and Output Decisions: Perfect Competition and Monopoly 310

Introduction	311
Competition and Market Types in Economic Analysis	313
<i>The Meaning of Competition</i>	313
<i>Examples of Market Types</i>	314
<i>Market Types and Competition in Theory and Reality</i>	315
Pricing and Output Decisions in Perfect Competition	316
<i>The Basic Business Decision</i>	316
<i>Key Assumptions of the Perfectly Competitive Market</i>	316
<i>The Total Revenue–Total Cost Approach to Selecting the Optimal Output Level</i>	320
<i>The Marginal Revenue–Marginal Cost Approach to Finding the Optimal Output Level</i>	321
<i>Economic Profit, Normal Profit, Loss, and Shutdown</i>	324
<i>The Competitive Market in the Long Run</i>	327
Pricing and Output Decisions in Monopoly Markets	329
The Implications of Perfect Competition and Monopoly for Managerial Decision Making	333

Global Application: The \$736,000 Bluefin Tuna	335
Summary	338
<i>Important Concepts</i>	338
<i>Questions</i>	339
<i>Problems</i>	340
Appendix 8A The Use of Calculus in Pricing and Output Decisions	343
Appendix 8B Break-Even Analysis (Volume-Cost-Profit)	345
CHAPTER 9 Pricing and Output Decisions: Monopolistic Competition and Oligopoly	361
Introduction	361
Monopolistic Competition	363
Oligopoly	365
<i>Market Concentration</i>	365
Pricing in an Oligopolistic Market: Rivalry and Mutual Interdependence	367
Competing in Imperfectly Competitive Markets	369
<i>Nonprice Competition</i>	369
<i>The Reality of Monopolistic Competition and Oligopoly: “Imperfect” Competition</i>	371
Strategy: The Fundamental Challenge for Firms in Imperfect Competition	372
<i>Industrial Organization</i>	374
<i>The Structure-Conduct-Performance Paradigm</i>	374
<i>The “New” Theory of Industrial Organization</i>	375
<i>Strategy and the Ideas of Michael Porter</i>	376
<i>Concluding Remarks on the Linkages Between Managerial Economics and Strategy</i>	378
Global Application: The World’s Market for Beer	378
Summary	380
<i>Important Concepts</i>	380
<i>Questions</i>	381
<i>Problems</i>	382
Appendix 9A A Mathematical Restatement of Monopolistic Competition	387
CHAPTER 10 Special Pricing Practices	390
Introduction	391
Cartel Arrangements	391
<i>Cases of Price Fixing by Cartels</i>	393
Price Leadership	396
<i>Barometric Price Leadership</i>	396
<i>Dominant Price Leadership</i>	397
Revenue Maximization	398
Price Discrimination	399
<i>Third-Degree Discrimination</i>	401
<i>Examples of Price Discrimination</i>	404
<i>Some Examples of Price Discrimination Practices</i>	406
<i>Pricing in the Hotel Industry: Example of Price Discrimination</i>	407

<i>Tying Arrangements: A Possible Extension of Price Discrimination</i>	409
<i>Social Welfare Implications of Price Discrimination</i>	410
Nonmarginal Pricing	410
<i>Cost-Plus Pricing</i>	411
<i>Incremental Pricing and Costing Analysis</i>	414
Multiproduct Pricing	415
<i>Products Complementary in Demand</i>	416
<i>Products Substitutable in Demand</i>	417
<i>Joint Products with Fixed Proportions</i>	417
<i>Joint Products in Variable Proportions</i>	419
Transfer Pricing	420
<i>No External Markets</i>	420
<i>External Markets</i>	421
Other Pricing Practices	422
Global Application: The Decline of European Cartels	423
<i>The European Carton-Board Cartel</i>	423
<i>The European Vitamin Cartel</i>	423
<i>Some Recent Cases of Price Fixing</i>	423
<i>Price Discrimination by Airlines</i>	424
An Airline Pricing Example	425
Summary	426
<i>Important Concepts</i>	427
<i>Questions</i>	427
<i>Problems</i>	428
CHAPTER 11 Game Theory and Asymmetric Information	432
Introduction	433
Game Theory: Modeling the Strategy of Conflict	434
<i>The Prisoners' Dilemma</i>	437
<i>The Basics of Bargaining</i>	439
<i>A General Framework</i>	442
Asymmetric Information	443
<i>The Lemons Model</i>	443
<i>Using Signals to Avoid the Lemons Problem</i>	446
<i>Job Market Signaling</i>	446
<i>Screening in Insurance Markets</i>	448
<i>Moral Hazard and Principal-Agent Problems</i>	450
<i>Creating Incentive-Compatible Mechanisms</i>	451
Summary	453
<i>Important Concepts</i>	455
<i>Questions</i>	456
<i>Problems</i>	457
CHAPTER 12 Capital Budgeting and Risk	461
Introduction	462
The Capital Budgeting Decision	463
<i>Types of Capital Budgeting Decisions</i>	463
Time Value of Money	464
Methods of Capital Project Evaluation	464
<i>Net Present Value</i>	464
<i>Internal Rate of Return</i>	466
<i>The Profitability Index</i>	467

<i>NPV Versus IRR</i>	467	
<i>Capital Budgeting in Practice</i>	470	
Cash Flows	470	
<i>Types of Cash Flows</i>	471	
Cost of Capital	472	
<i>Debt</i>	473	
<i>Equity</i>	473	
<i>The Weighted Cost of Capital</i>	475	
The Capital Budgeting Model	475	
Capital Rationing	477	
Risk Versus Uncertainty	477	
Sources of Business Risk	478	
The Measures of Risk	478	
<i>Expected Value</i>	479	
<i>The Standard Deviation</i>	480	
<i>Discrete Versus Continuous Distributions and the Normal Curve</i>	481	
<i>The Coefficient of Variation</i>	482	
Capital Budgeting Under Conditions of Risk	483	
Two Other Methods for Incorporating Risk	484	
<i>The Risk-Adjusted Discount Rate</i>	484	
<i>Certainty Equivalents</i>	485	
<i>Present Value Break-Even Analysis</i>	486	
Sensitivity and Scenario Analysis	486	
Simulation	487	
Decision Trees	488	
Real Options in Capital Budgeting	490	
<i>Real Options in Practice</i>	491	
<i>An Abandonment Option</i>	491	
Global Application	493	
<i>Political Risk</i>	493	
<i>Containing International Risk</i>	494	
<i>Capital Budgeting in Practice</i>	494	
Summary	498	
<i>Important Concepts</i>	498	
<i>Questions</i>	499	
<i>Problems</i>	500	
Appendix 12A The Value of a Corporation	507	
CHAPTER 13 The Multinational Corporation in a Global Setting	509	
Introduction	510	
Opportunities for Multinational Corporations in a Global Economy	511	
<i>Expanding the Dressings Category Throughout the World</i>	511	
Doing Business in India	513	
<i>Brief Introduction to Some of India's Key Industries</i>	513	
<i>The Business and Investment Climate</i>	515	
<i>Socioeconomic and Political Factors</i>	515	
<i>McDonald's in India</i>	516	
Risks Faced by a Multinational Corporation	517	
Exchange Rates	518	

Exchange Rate Hedging	518
<i>Offsetting Transactions</i>	519
<i>The Forward Market</i>	519
<i>The Futures Market</i>	519
<i>Currency Options</i>	519
<i>Currency Swaps</i>	520
Foreign Direct Investment	520
Multinational Capital Budgeting	521
<i>Intercompany Fund Flows</i>	521
<i>Inflation Rates</i>	521
<i>Exchange Rates</i>	521
<i>Tax Differences</i>	521
<i>Differences in Cash Flows</i>	522
<i>Cost of Capital</i>	522
<i>The Final Project Valuation</i>	522
The Repositioning of Funds	523
Multinational Transfer Pricing	523
<i>Multinational Transfer Pricing Example</i>	524
<i>Transfer Pricing in Practice</i>	525
Global Application: Tony the Tiger Meets a Bengal Tiger	526
Summary	530
<i>Important Concepts</i>	531
<i>Questions</i>	531
<i>Problems</i>	531

CHAPTER 14 Government and Industry: Challenges and Opportunities for Today's Manager 533

Introduction	534
The Rationale for Government Involvement in a Market Economy	534
<i>Providing a Legal Framework for Competition: The Antitrust Laws</i>	535
<i>The Clayton Act (1914)</i>	535
<i>Dealing with Market Externalities: Another Key Function of Government in the Market Economy</i>	538
Stabilization of the Aggregate Economy: Monetary and Fiscal Policy	541
<i>Monetary Policy</i>	541
<i>Fiscal Policy</i>	541
<i>Lags</i>	542
Subprime Loan Financial Crisis of 2007 to 2009	542
<i>Background</i>	542
<i>Securitization of Mortgaged-Backed Securities</i>	543
<i>Disappearing Liquidity Challenges the Financial System</i>	544
<i>Changing Bank Regulations to Avoid Future Crises</i>	545
<i>Global Financial Deregulation and Growth in International Capital Flows</i>	546
Government Deregulation, Mergers, and Acquisitions	546
<i>Why Firms Merge</i>	548
Government Protection of Intellectual Property (IP)	551
Global Application: The Failed Attempt to Merge by General Electric and Honeywell	553
Summary	554
<i>Important Concepts</i>	554
<i>Questions</i>	555

CHAPTER 15	The Global Soft Drink Industry	556
	Introduction	556
	Factors and Trends Influencing Demand	556
	Factors and Trends Influencing Supply	558
	Energy Drinks	559
<i>Appendix A</i>	<i>Statistical and Financial Tables</i>	<i>563</i>
<i>Index</i>		<i>579</i>

This page intentionally left blank

One day after class, a student in one of our courses commented on the managerial economics text then being used: “This book is very dry. What it needs is a plot!” To a large extent, the idea for this text stemmed from this remark. This is a text that we believe will excite readers about managerial economics as well as inform them about this vital part of management education. Each chapter begins with a Situation, in which managers in a fictional company, Global Foods, Inc., must make certain key decisions about their products in the beverage industry. After the relevant economic concepts or tools of analysis are presented, each chapter ends with a Solution, a suggested way in which these concepts or tools can be used to help managers make the best decision.

We are well aware of the reputation that economics courses have among some business students, that they are “too theoretical and not practical enough for the real world.” In our opinion, nothing could be further from the truth. We know that the instructors in managerial economics will agree with us on this matter. We hope that this text will serve as a solid supplement to their classroom efforts to demonstrate to their students the importance and utility of economic theory for business decision making.


This text is designed for upper-level undergraduate and first-year MBA courses in managerial economics and applied economics. The first two chapters form a general introduction to economics and economic reasoning. A review of the mathematical concepts and tools used in the text has been placed on the Companion Website. In addition to discussing the applications of economic theory to the firm, our text (as is the custom with all texts in managerial economics) includes chapters on various tools of analysis that are helpful to business decision makers but that are not part of the core of traditional microeconomic theory. These are demand, production, and cost estimation using regression analysis, forecasting, capital budgeting, and risk analysis. A discussion of linear programming is also available online, along with a review of the time value of money.

IMPROVEMENTS IN THE SEVENTH EDITION

As in all our previous editions, this edition’s changes are based on our classroom teaching, consulting engagements, and corporate education seminars. In addition, we have received a number of useful suggestions from the reviewers of our sixth edition.

In this seventh edition, we welcome co-author, Stephen Erfle of Dickinson College. Steve has been of particular help in revising and improving the quantitative sections of our text. He has also provided a major addition: the use of Excel in the presentation of many of the numerical and graphical illustrations presented throughout the text.

Here are noteworthy additions, improvements, and enhancements to this edition:

- ▶ We developed Excel Applications (Excel Apps) for a number of the numerical and graphical illustrations used throughout the text. These apps, noted with the icon,  are available to readers on our website and an index delineating what is in each Excel App is provided on the inside rear cover of the text. They allow readers to turn the static figures and tables in the text into dynamic illustrations, and they will also strengthen students’ ability to use Excel, which we believe is a critical skill in today’s job market.

- ▶ We have developed a series of regression Excel Apps that provide readers with a more detailed discussion of many of the topics in regression analysis touched on in Chapter 5. See the Excel Apps listing on the back cover for additional information about coverage about regression topics in various chapters.
- ▶ We have completely rewritten Chapter 11, “Game Theory and Asymmetric Information.” User feedback requested more in-depth coverage of this challenging topic. The chapter now includes increased coverage of game theory and bargaining as well as a more in-depth discussion of adverse selection and moral hazard.
- ▶ Chapter 15, “Managerial Economics in Action,” is an entirely new chapter. Throughout the text, we try to show how the economic concepts and tools of analysis can be applied to the beverage industry by introducing in each chapter a Situation and a Solution for our hypothetical company called Global Foods. In our concluding chapter, we discuss how the basic concepts of supply and demand can be applied to the real global soft drink industry. This chapter was written especially for this edition by a seasoned industry consultant, Farshad Sarmad. Using current industry data and his own experiences, Farshad shows how the factors affecting supply and demand can be applied to various segments of the soft drink industry in countries around the world.
- ▶ Significant developments in information and communications technology (e.g., cloud computing, social media, and Internet commerce) have enabled businesses to store massive amounts of data generated in digital format. We introduce readers to these developments in Chapter 5, “Demand Estimation and Forecasting.” We asked Dr. Mukal Patki, a business analytics specialist at PayPal, to help us with this task. In Appendix 5B, “Understanding Consumer Behavior Through Testing,” he talks about how “big data” has enabled companies to conduct in-depth studies of consumer behavior using a technique called “test and learn.”
- ▶ In Chapter 14, “Government and Industry: Challenges and Opportunities for Today’s Manager,” we have added a discussion of patent laws and the concept of protecting intellectual property (IP), such as trademarks and copyrights. These are significant ways that government affects commerce in a free market economy. This section was written by Riyan Harding, an expert in the commercialization of IP at IBM, a company that is recognized throughout the world for the number of new patents it receives from the U.S. government every year.
- ▶ To give readers a better idea of how managerial economics can be applied in some of the BRIC countries (Brazil, Russia, India, and China), we asked Navin Punjabi, a professor of business in Mumbai, to discuss some of the challenges of doing business in India (Chapter 13). We also asked Lisa Vortsman, a product manager for the dressings category in an actual “Global Foods” company (Lisa requested that her company not be identified) to talk about the challenges of increasing the demand for this category in countries like Russia and Brazil.
- ▶ We have also received the help of outside experts to improve our discussions about other topics of importance and current interest in business. F. John Mathis, Professor of Global Economics and Finance, has written a highly informative summary of the causes and consequences of the 2008 financial crisis (Chapter 14). To provide our readers with a better idea of the actual challenges of initiating and implementing a supply chain management system, we interviewed Steve Martson. Steve, a recently retired executive who has led supply chain systems implementation in companies such as Dell and IBM, talks about some of his experiences working in this very important field (Appendix 7C).

As we have done in all of our previous editions, we have updated our examples wherever appropriate. We have kept some of the examples that we first introduced in our previous editions if we believe they serve as good teaching illustrations, regardless of when they occurred.

FEATURES





The Case of Global Foods

This case, which runs throughout the entire text, helps to integrate and apply the key concepts presented in each chapter with an everyday consumer product: soft drinks and other nonalcoholic beverages. Each chapter's Situation and Solution are based on actual challenges faced by companies in the beverage industry. The stories told in each case are intended to stimulate reader interest by bringing the concepts and tools of analysis to life, which are presented in the graphs and numerical examples.

Global Applications

The Global Applications sections exemplify how the concepts and tools of analysis can be applied in other countries. In this edition, more examples, both in the Global Application sections and in the main body of the chapter, are drawn from world growth markets such as China, India, Brazil, and Russia.

In-Text Icons

References to the Mathematical Appendix are noted by the symbol  and references to the appendix about the Time Value of Money are noted by the symbol . As explained on the inside front cover, there are now two types of Excel icons. References to Excel exercise modules are noted by  and a listing of Excel modules is provided on the inside front cover. References to Excel Apps are noted by  and a list of Excel Apps is provided on the inside rear cover.


Learning Objectives

Each chapter begins with a list of Learning Objectives, which outline the concepts students should be able to take away once they've read the chapter. These Learning Objectives frame the tools that future managers need to know to succeed.

ANCILLARY MATERIALS

Companion Website (www.pearsonhighered.com/keat)

The website contains Internet exercises, activities, and resources related specifically for *Managerial Economics: Economic Tools for Today's Decision Makers*.

A number of other resources are available on the Companion Website such as the Mathematical Appendix, Time Value of Money Appendix, and Excel exercise modules. The modules provide students with templates of the economic models in the text. In this edition, we have introduced a new feature for our Companion Website: Excel Apps . A listing of Excel Apps is provided on the inside back cover. The Excel Apps provide students with instructions to build their own models. In so doing, they gain a deeper understanding of the underlying assumptions of the models themselves.

Business Simulation: New to the seventh edition is an online, computer-based business simulation available on the text's Companion Website. In this simulation, a student becomes the product manager of 'alpha', a consumer product similar to bottled water or soft drinks. The student has the choice of being a low-price competitor (e.g., a private-label bottled water) or a premium-price competitor (e.g., Evian or San Pellegrino). The student competes against two computer-generated companies. One is a low-price competitor; the other is a premium-priced competitor. Students make decisions on price, marketing, process development (to lower costs), and production capacity. By making these decisions and getting the results, students learn in a dynamic and engaging way

about the concepts of price and marketing elasticities and the interdependency of pricing in oligopolistic markets.

Online Study Guide: The Online Study Guide offers students another opportunity to sharpen their problem-solving skills and to assess their understanding of the text material. The Online Study Guide grades each question submitted by the student, provides immediate feedback for correct and incorrect answers, and allows students to e-mail results to up to four e-mail addresses.

Instructor's Resource Center

This password-protected site is accessible from www.pearsonhighered.com/keat and hosts all of the resources listed below. Instructors may click on the Resources link to access files or may contact their sales representative for additional information.

Online Instructor's Manual: This manual, written by the textbook authors, contains all answers to the questions and problems found in the text.

Online Test Item File and TestGen: Written by James Holcomb of the University of Texas, El Paso, and available to instructors in test generator software (TestGen with Quizmaster), this test bank contains multiple-choice questions and a set of analytical questions for use in testing students on the material presented in each chapter of the text. Answers are also provided. TestGen's graphical interface enables instructors to view, edit, and add questions; transfer questions to tests; and print different forms of tests. Instructors also have the option of reformatting tests with fonts and styles, margins, and headers and footers, as in any word-processing document. Search and sort features let the instructor quickly locate questions and arrange them quickly in a preferred order. Quizmaster can work with your school's computer network to automatically grade the exams, store the results electronically, and allow the instructor to view and print a variety of reports.

Online PowerPoint Presentation: This lecture presentation tool, prepared by Gary F. Wilkinson of Indiana Wesleyan University, offers outlines and summaries of important text material, tables and graphs, and additional examples. The package allows instructors to make full-color, professional-looking presentations and custom handouts for students.

ACKNOWLEDGMENTS

We wish to thank our colleagues at Thunderbird School of Global Management and Dickinson College, and former colleagues at IBM and Pace University for their help and encouragement in our work for this and all previous editions. We also wish to thank those who have helped us to improve this seventh edition. As noted earlier, they are: Riyon Harding, Stephen C. Marston, Dr. F. John Mathis, Dr. Navin Punjabi, Dr. Mukal Patki, Farshad Samad, and Lisa Vortsman.

We also thank Dr. Jack Yurkiewicz, professor of management science at Pace University, for writing the material on linear programming that is available on our Companion Website; Professor Gary Wilkinson of Indiana Wesleyan University for preparing the PowerPoint presentation; and Professor James Holcomb of the University of Texas, El Paso, for preparing the Test Item File.

Our appreciation also goes to the reviewers of the seventh edition: Nelson Altamirano, National University; Cassandra DiRienzo, Elon University; Kenneth C. Fah, Ohio Dominican University; Rajeev Goel, Illinois State University; James Holcomb, University of Texas, El Paso; John S. Howe, University of Missouri, Columbia; M. Ebru Kongar, Dickinson College; Matthew Roelofs, Western Washington University; Jennifer VanGilder, Ursinus College; and Gary F. Wilkinson, Indiana Wesleyan University. We also wish to thank Benjamin Greene, Anne Marie Weichert, and Qiaoling

Yuan of Dickinson College, all of whom provided us with comments from a student's perspective on the text, end-of-chapter problems, Excel Apps, and the Online Study Guide.

And we continue to be grateful to all the reviewers of the previous six editions: Michael J. Applegate, Oklahoma State University; Mina Balamoune, University of North Florida; Robert Britt, West Virginia University; Stacey Brook, University of Sioux Falls; Peter Brust, University of Tampa; Charles Callahan, III, State University of New York at Brockport; John Conant, Indiana State University; Richard Cox, University of Arkansas; Brad Ewing, Texas Technical University; Lewis Freiberg, Northeastern Illinois University; Edward H. Heinze, Valparaiso University; George Hoffer, Virginia Commonwealth University; Al Holtmann, University of Miami; Richard A. Jenner, San Francisco State University; Aric Krause, Westminster College; Douglas Lamdin, University of Maryland, Baltimore County; Dale Lehman, Fort Lewis College; Jerry Manahan, Midwestern State University; Cynthia McCarty, Jacksonville State University; Yale L. Meltzer, College of Staten Island; L. W. (Bill) Murray, University of San Francisco; Alex Orlov, Radford University; Jan Palmer, Ohio University–Athens; Leila J. Pratt, The University of Tennessee at Chattanooga; L. B. Pulley, University of Virginia; Mathew Roelofs, Western Washington University; Roy Savoian, Lynchburg College; Frederica Shockley, California State University–Chico; Ken Slaysman, York College of Pennsylvania; William Doyle Smith, University of Texas at El Paso; Robert Stuart, Rutgers University; James Tallant, Cape Fear Community College; Mo-Yin Tam, University of Illinois at Chicago; Yien-I Tu, University of Arkansas; Lawrence White, New York University; Richard Winkelman, Arizona State University; Daryl N. Winn, University of Colorado; Darin Wohlgemuth, Iowa State University; Richard Zuber, University of North Carolina at Charlotte; and Habib Zuberi, Central Michigan University.

In closing, we would like to express our appreciation to the helpful, encouraging, and patient team at Pearson: Donna Battista, Editor in Chief; Adrienne D'Ambrosio, Executive Acquisitions Editor; Nancy Freihofer, Production Project Manager; and Sarah Dumouchelle, Editorial Project Manager, and Tiffany Rupp, Project Editor, S4Carlisle Publishing Services.

This page intentionally left blank

About the Authors

Paul G. Keat has been a member of the Global Business Faculty at Thunderbird School of Global Management for the past twenty-five years. At present he is an Associate Professor Emeritus. Prior to his coming to Thunderbird, he was associated for many years with the International Business Machines Corporation in professional and managerial capacities.

His education includes a B.B.A. in accounting from the Baruch School of the City University of New York, an M.A. from Washington University, and an M.A. and Ph.D. in economics from the University of Chicago.

Dr. Keat began his IBM career in the department of economic research and then moved into the long-range planning area. Later, as a member of the finance function, he spent several years at IBM's European headquarters in Paris, as manager in the financial planning area and then as the financial manager for the company's European software business. After his return to the United States, Dr. Keat served as manager in the pricing area of one of the company's manufacturing groups. Before leaving IBM in 1987, he was associated with the company's International Finance, Planning and Administration School (IFPA), where he taught managerial economics, lectured on finance in a number of company-related courses, and managed academic courses. He also taught at IBM's IFPA School at La Hulpe, Belgium.

Dr. Keat has taught at several U.S. universities, including Washington University, the City University of New York (CUNY), and Iona College. He was an adjunct professor of finance at the Lubin Graduate School of Business at Pace University, and he also taught in Pace University's Executive MBA program.

Philip K. Y. Young (www.philipkyoung.com) is the founder and president of Nth Degree Systems, Inc., a consulting firm that provides customized education and training programs to major corporations around the world. He recently co-founded Learning Burst Academy (www.learningbursts.com), a company that produces courses in business education in an innovative, digital format. He is also a member of the global faculty network of Duke Corporate Education. He has thirty years of teaching experience as a professor of economics in MBA programs and over twenty-five years of experience developing and teaching customized education and training programs.

Most of Dr. Young's teaching career was spent in the Lubin School of Business at Pace University in New York, followed by several years as clinical professor of management at Thunderbird School of Global Management. His list of clients includes a number of multinational corporations in industries such as information technology, telecommunications, fast-moving packaged consumer goods, consulting services, advertising and public relations, pharmaceuticals, semiconductor manufacturing and design, diversified manufacturing, and financial services. He teaches for these companies in the United States, Latin America, western and central Europe, Asia, and the Middle East.

Dr. Young received a B.A. from the University of Hawaii, a master's degree in international relations from Columbia University, and a Ph.D. in economics from New York University.

Stephen E. Erfle began his career as a managerial economist during a 1994–1995 sabbatical at Seagram Classics Wine Company (SCWC). During those fourteen months, he maintained offices at Sterling Vineyards and at Mumm Cuvée Napa, where, respectively, the finance and marketing departments of SCWC resided. Trained as a microeconomic theorist, he began to use his economist’s toolkit to analyze concrete business questions, such as, Should Mumm raise the price of Brut Prestige a dollar a bottle? When does it make sense to have another tasting room associate on the floor in Sterling’s tasting room?

On returning to Dickinson College, Dr. Erfle decided to refocus his teaching in a more applied direction. He helped found the International Business and Management department and major during the late 1990s. One of the core courses in that major is his course, Managerial Economics, which uses Excel as a teaching platform. This course is modeled after what he did during his SCWC sabbatical. In the past fifteen years, he has taught more than a thousand undergraduates how to build economic models in order to do comparative statics analysis and how to do regression modeling in Excel.

Dr. Erfle received a B.S. in mathematics and B.A. in economics from the University of California, Davis, and a master’s and Ph.D. in economics from Harvard University. He has also taught in the Economics Department at Dickinson College and in the School of Social Sciences at the University of California, Irvine. He is also involved in wine education; he has taught wine-tasting classes and conducted wine tastings since his graduate school days as the resident economics and wine tutor for Harvard’s Leverett House.

Introduction

Learning Objectives

Upon completion of this chapter, readers should be able to:

- Define managerial economics and discuss briefly its relationship to micro-economics and other related fields of study such as finance, marketing, and statistics.
- Cite the important types of decisions that managers must make concerning the allocation of a company's scarce resources.
- Cite and compare the three basic economic questions from the standpoint of both a country and a company.

The Situation

The last of the color slides was barely off the screen when Bob Burns, the CEO of Global Foods, Inc., turned to his board of directors to raise the question that he had been waiting all week to ask. "Well, ladies and gentlemen, are you with me in this new venture? Is it a 'go'? Shall we get into the soft drink business?"

"It's not that easy, Bob. We need some time to think it over. You're asking us to endorse a very major *decision*, one that will have a long-term impact on the direction of the company."

"I appreciate your wish to deliberate further, Dr. Breakstone," Bob responded, "but I would like to reach a decision today. As the president of a major university, you have been especially valuable in advising this company in matters relating to social and governmental policies. But we must diversify our business very soon in order to maintain the steady growth in profits that we have achieved in recent years. As my presentation showed, the manufacturing and marketing of our own brand of soft drink is one of the best ways to do this. It represents a significant diversification, yet it is very closely related to our core business: food.

"The *economics* of the soft drink market tell us that we would be foolish to pass up the kind of *investment return* that the market offers to those newcomers willing to take the *risk*. The food business is generally a mature one. On the other hand, our *forecast* indicates that there is still a lot of room for growth in the soft drink market. To be sure, there is a tremendous amount of *competition* from the 'red team' and the 'blue team.' But we already have expertise in the food business, and it should carry over into the beverage market."

"That's just it, Bob," interjected another board member. "Are we prepared to take this risk? You yourself acknowledged that the *market power* wielded by the two dominant companies in this business is not to be taken lightly. Others have tried to take market share from them and have failed miserably. Moreover, the projections that you have shown for a growing soft drink market are based on the *assumption* that the growth rate will remain the same as it has been in the past ten years or so. As we all know, the soft drink market has been growing, but it has also been very fickle. Only recently, Americans were on a health kick, and fruit juices and bottled waters along with health foods were in fashion. Now it seems that soft drinks are back in style again.

(continued)

(continued)

Who knows what people will want in the future? Maybe we'll all go back to drinking five cups of coffee a day. And what about all the money that we're going to have to spend up front to *differentiate* our product? As you well know, in the processed-food business, establishing brand recognition—not to mention brand loyalty—can be extremely difficult and costly.”

“Well, ladies and gentlemen, all your concerns are certainly legitimate ones, and believe me, I have given much thought to these drawbacks. This is one of the biggest decisions that I will have made since becoming CEO. My staff has spent hundreds of hours analyzing all available data to arrive at a judgment. Our findings indicate a strong probability of earning an above-average return on an investment in the soft drink business, a return commensurate with the kind of risk we know exists in that market. But if we could make all our decisions with 100 percent certainty simply by feeding numbers into a computer, we'd all be out of a job. To be sure, details on production, cost, pricing, distribution, advertising, financing, and organizational structure remain to be ironed out. However, if we wait until all these details are worked out, we may be missing a window of opportunity that might not appear again in this market for a long time. I say that we should go ahead with this project as soon as possible. And unanimity among the board members will give me greater confidence in this endeavor.”

INTRODUCTION: ECONOMICS AND MANAGERIAL DECISION MAKING

Managerial Economics is one of the most important and useful courses in your curriculum of studies. It will provide you with a foundation for studying other courses in finance, marketing, operations research, and managerial accounting. It will also provide you with a theoretical framework for tying together courses in the entire curriculum so you can have a cross-functional view of your studies.

Economics is “the study of the behavior of human beings in producing, distributing and consuming material goods and services in a world of scarce resources.”¹ *Management* is the discipline of organizing and allocating a firm’s scarce resources to achieve its desired objectives.² These two definitions clearly point to the relationship between economics and managerial decision making. In fact, we can combine these two terms and define **managerial economics** as the use of economic analysis to make business decisions involving the best use of an organization’s scarce resources.

Joel Dean, author of the first managerial economics textbook, defines managerial economics as “the use of economic analysis in the formulation of business policies.” He also notes a “big gap between the problems of logic that intrigue economic theorists and the problems of policy that plague practical management [which] needs to be bridged in order to give executives access to the practical contributions that economic thinking can make to top-management policies.”³

William Baumol, a highly respected economist and industry consultant, stated that an economist can use his or her ability to build theoretical models and apply them to any business problem, no matter how complex, break it down into essential components, and describe the relationship among the components, thereby facilitating a systematic search for an optimal solution. In his extensive experience as a consultant to both industry and government, he found that every problem that he worked on was

¹Campbell McConnell, *Economics*, New York: McGraw-Hill, 1993, p. 1.

²For books supporting this definition, see Peter Drucker, *Management*, New York: Harper & Row, 1973.

³Joel Dean, *Managerial Economics*, Englewood Cliffs, NJ: Prentice-Hall, 1951, p. vii.

helped in some way by “the method of reasoning involved in the derivation of some economic theorem.”⁴

William H. Meckling, the former dean of the Graduate School of Management at the University of Rochester, expressed a similar sentiment in an interview conducted by *The Wall Street Journal*. In his view, “economics is a discipline that can help students solve the sort of problems they meet within the firm.” Recalling his experience as the director of naval warfare analysis at the Center for Naval Analysis and as an economic analyst at the Rand Corporation, one of the nation’s most prominent think tanks, Meckling stated that these institutions are “dominated by physical scientist types, really brilliant people.” However, he went on to say that “the economists knew how to structure the problems . . . the rest of the people knew a lot about technical things but they had never thought about how you structure big issues.”⁵

As it has evolved in undergraduate and graduate programs over the past half century, managerial economics is essentially a course in applied microeconomics that includes selected quantitative techniques common to other disciplines such as linear programming (management science), regression analysis (statistics, econometrics, and management science), capital budgeting (finance), and cost analysis (managerial and cost accounting). From our perspective as economists, we see that many disciplines in business studies have drawn from the core of microeconomics for concepts and theoretical support. For example, the economic analysis of demand and price elasticity can be found in most marketing texts. The division of markets into four types—perfect competition, pure monopoly, monopolistic competition, and oligopoly—is generally the basis for the analysis of the competitive environment presented in books on corporate strategy and marketing strategy.⁶

There are a number of other examples to be found. The economic concept of opportunity cost serves as the foundation for the analysis of relevant cost in managerial accounting and for the use of the “hurdle rate”⁷ in finance. As shown in Chapter 2, opportunity cost also plays an important part in understanding how firms create “economic value” for their shareholders. Finally, in recent years, certain authors have linked their managerial economics texts thematically with strategy and human resources.⁸ Figure 1.1 illustrates our view that managerial economics is closely linked with many other disciplines in a business curriculum.

Our approach in this text is to show linkages of economics with other business functions, while maintaining a focus on the heart of managerial economics—the microeconomic theory of the behavior of consumers and firms in competitive markets. When clearly understood and exemplified in actual business examples, this theory provides managers with a basic framework for making key business decisions about the allocation of their firm’s scarce resources. In making these decisions, managers must essentially deal with the following questions listed in abridged form:

1. What are the economic conditions in a particular market in which we are or could be competing? In particular:
 - a. Market structure?
 - b. Supply and demand conditions?

⁴William Baumol, “What Can Economic Theory Contribute to Managerial Economics?” *American Economic Review*, 51, 2 (May 1961), p. 114.

⁵“Economics Has Much to Teach the Businessman,” *Wall Street Journal*, May 3, 1983.

⁶Professor Michael Porter, whose books on strategy have greatly influenced this field of study, is himself a Ph.D. in economics.

⁷Essentially, this is a company’s cost of funds expressed as a percentage (e.g., 15 percent). Any project funded by the company should have a rate of return that is greater than this level.

⁸See for example, David Besanko et al., *Economics of Strategy*, New York: John Wiley & Sons, 2009, and James A. Brickley et al., *Managerial Economics and Organizational Architecture*, New York, McGraw-Hill, 2003.

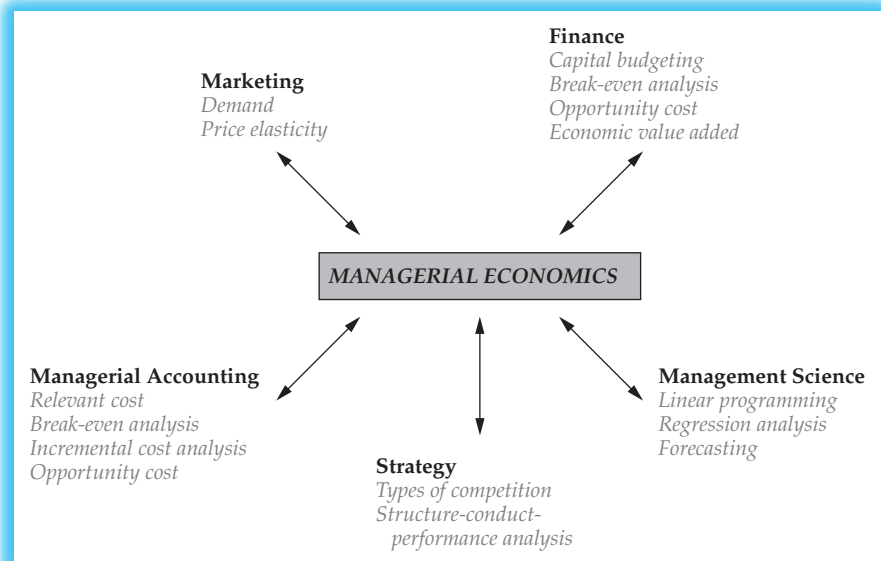


Figure 1.1 Managerial Economics and Other Business Disciplines

- c. Technology?
 - d. Government regulations?
 - e. International dimensions?
 - f. Future conditions?
 - g. Macroeconomic factors?
2. Should our firm be in this business?
 3. If so, what price and output levels should we set in order to maximize our economic profit or minimize our losses in the short run?
 4. How can we organize and invest in our resources (land, labor, capital, managerial skills) in such a way that we maintain a competitive advantage over other firms in this market?
 - a. Cost leader?
 - b. Product differentiation?
 - c. Focus on market niche?
 - d. Outsourcing, alliances, mergers, acquisitions?
 - e. International dimension—regional or country focus or expansion?
 5. What are the risks involved?

Perhaps the most fundamental management question is question 2, which concerns whether a firm should be in the business in which it is operating. This is the very question addressed by Bob Burns and the rest of the board of directors of Global Foods in this chapter's "The Situation" vignette.

Note that question 5 has to do with a firm's risk. Uncertainty pervades our everyday lives, especially when we are considering what may happen in the future, and uncertainty, or risk, is always present in the operations of a business. Of course, some things are fairly certain. A company that buys steel can get a price quote and be certain what it will pay for a ton. A company with temporary excess cash to invest for a short period of time can ascertain the interest rate it will earn. An investor purchasing a U.S. Treasury bill is virtually certain that he or she will be repaid in full at maturity.

However, when it comes to future impacts, few things are certain. We can define risk or uncertainty (we explain the difference between these two terms in Chapter 12) as a chance or possibility that actual future outcomes will differ from those expected today. Actually we are usually only concerned with unfavorable results. Thus we can say that risk is the possibility that the outcomes of an action will turn out to be worse than expected. Typical of the types of risk that businesses face include:

- Changes in demand and supply conditions
- Technological changes and the effect of competition
- Changes in interest rates and inflation rates
- Exchange rate changes for companies engaged in international trade
- Political risk for companies with foreign operations

You may not literally see the term *risk* in many of the chapters of this book. In some of the chapters we implicitly assume that we know the level of demand, the product price, production cost, and the economic profit resulting from operations. However, we really know that risk is present in most situations. In Chapter 12, we show how businesses attempt to quantify risk and how decisions are made under these conditions, but this is not all. Chapter 5 deals with estimating the effects of changes in the variables that determine the demand for products. It also looks at predicting future results based on past experience, assuming we have sufficient historical data. It also talks about the challenges of estimating what may happen in the future if historical data do not exist.

A BRIEF REVIEW OF IMPORTANT ECONOMIC TERMS AND CONCEPTS

For purposes of study and teaching, economics is divided into two broad categories: microeconomics and macroeconomics. The former concerns the study of individual consumers and producers in specific markets, and the latter deals with the aggregate economy. Topics in microeconomics include supply and demand in individual markets, the pricing of specific outputs and inputs (also called factors of production, or resources), production and cost structures for individual goods and services, and the distribution of income and output in the population. Topics in macroeconomics include analysis of the gross domestic product (also referred to as “national income analysis”), unemployment, inflation, fiscal and monetary policy, and the trade and financial relationships among nations.

Microeconomics is the category that is most used in managerial economics. However, certain aspects of macroeconomics must also be included because decisions by managers of firms are influenced by their views of the current and future conditions of the macroeconomy. For example, we can well imagine that the management of a company producing capital equipment (e.g., computers, machine tools, trucks, or robotic instruments) would indeed be remiss if they did not factor into their sales forecast some consideration of the macroeconomic outlook. For these and other companies whose businesses are particularly sensitive to the business cycle, a recession would have an unfavorable effect on their sales, whereas a robust period of economic expansion would be beneficial. However, for the most part, managerial economics is based on the variables, models, and concepts that embody microeconomic theory.

As defined in the previous section, economics is the study of how choices are made regarding the use of scarce resources in the production, consumption, and distribution of goods and services. The key term is *scarce resources*. **Scarcity** can be defined as a condition in which resources are not available to satisfy all the needs and wants of a specified group of people. Although scarcity refers to the supply of a **resource**, it makes sense only in relation to the demand for the resource. For example,

there is only one Mona Lisa. Therefore, we can safely say that the supply of this particular work of art by da Vinci is limited. Nevertheless, if for some strange reason no one wanted this magnificent work of art, then in purely economic terms it would not be considered scarce. Let us take another example: broken glass on the streets of New York City. Here we have a case of a “resource” that is not scarce not only because there is a lot of broken glass to be found, but also because nobody wants it! Now suppose there were a new art movement inspired by the use of materials retrieved from the streets of urban areas, with broken glass from the streets of New York being particularly desirable. The once-plentiful resource would fast become a “scarce” commodity.

The relative nature of scarcity is represented in Figure 1.2. As seen in Figure 1.2, the supply of resources is used to meet the demand for these resources by the population. Because the population’s needs and wants exceed the ability of the resources to satisfy all the demands, scarcity exists.

In an introductory economics course, the concept of scarcity is usually discussed in relation to an entire country and its people. For example, you will probably recall from your first course in economics the classic example of “guns” (representing a country’s devotion of resources to national defense) versus “butter” (representing the use of resources for peacetime goods and services). To be sure, scarcity is a condition with which individual consumers and producers must also deal. This text is primarily concerned with the way in which managers of the producing organizations contend with scarcity. However, before discussing this particular aspect of the problem, let us review the condition of scarcity from the perspective of an entire country.

The intent of the “guns versus butter” example is to illustrate that scarcity forces a country to choose the amounts of resources that it wants to allocate between defense and peacetime goods and services. In so doing, its people must reckon with the **opportunity cost** of their decision. This type of cost can be defined as the amount or subjective value that must be sacrificed in choosing one activity over the next best alternative. In the “guns versus butter” example, one activity involves the production of war goods and services, and the other pertains to the production of peacetime goods and services. Because of the scarcity of resources, the more that the country allocates to guns, the less it will have to produce butter, and vice versa. The opportunity cost of additional units of guns are the units of butter that the country must forgo in the resource allocation process. The opposite would apply as resources are allocated more for the production of butter than for guns.

In the presence of a limited supply relative to demand, countries must decide how to allocate their scarce resources. This decision is central to the study of economics. In fact, economics has been defined as “the science which studies human behavior as a relationship between ends and scarce means which have alternative

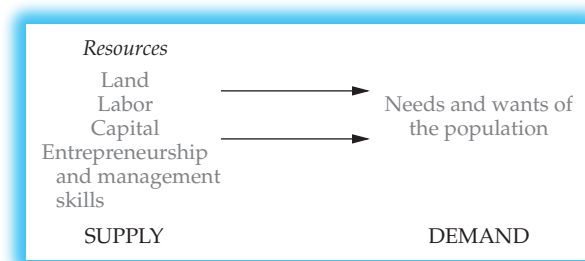


Figure 1.2 Supply, Demand, and Scarcity

uses.”⁹ Essentially, the allocation decision can be viewed as comprising three separate choices:

1. *What* goods and services should be produced and in what quantities?
2. *How* should these goods and services be produced?
3. *For whom* should these goods and services be produced?

The first question incorporates the “guns versus butter” decision. Should a country with scarce resources produce guns? Should it produce butter? If so, how much butter and how many guns? The same applies to the countless other goods and services or product groups that a country is capable of producing.

The second question involves the allocation of a country’s resources in the production of a particular good or service. Suppose a country decides to produce a certain amount of butter. What amounts of land, labor, capital, and entrepreneurial efforts should it devote to this end? Should it use more workers than machinery (a labor-intensive process), or vice versa (a capital-intensive process)?

The meaning of the third question should be readily apparent. It is a decision that must be made about the distribution of a country’s output of goods and services among the members of its population.

All countries must deal with these three basic questions because all have scarce resources. Scarcity is a more serious problem in some countries than in others, but all have needs and wants that cannot be completely met by their existing resources. Precisely how these countries go about making allocation decisions is the question to which we now turn.

There are essentially three ways a country can answer the questions of what, how, and for whom. These ways, referred to as *processes*, are as follows:

1. **Market process:** The use of supply, demand, and material incentives to answer the questions of what, how, and for whom.
2. **Command process:** The use of the government or some central authority to answer the three basic questions. (This process is sometimes referred to as the *political process*.)
3. **Traditional process:** The use of customs and traditions to answer the three basic questions.

Countries generally employ a combination of these three processes to allocate their scarce resources. The market process is predominant in the United States, although the command process plays an important role. Hence, the United States is said to have a mixed economy. Based on the levels of spending by the federal, state, and local governments, we can state that approximately one-fifth of the goods and services produced in the United States are influenced by the command process. The command process does not necessarily mean that a government literally orders the production of certain amounts of guns, butter, or other goods or services; rather, a government may use the material incentives of the market process to allocate resources in certain ways, a process often referred to as *indirect command*. For example, the government offers defense contractors the opportunity to earn a profit by producing military goods and services. In addition, the government can control the allocation of resources in a more direct way through various laws governing the actions of both consumers and producers. For example, the government controls manufacturing and distribution through such agencies as the U.S. Food and Drug Administration. It attempts to control consumer use of certain foods and drugs through various laws and regulations. A simple but important example of this pertains to the tobacco industry. Over the past several decades, the U.S. government has made determined efforts to convince people to stop smoking. These efforts range from warnings on cigarette packages to the banning of smoking on airline flights and in restaurants.

⁹Lionel Robbins, *An Essay on the Nature and Significance of Economic Science*, 2nd ed., London: Macmillan, 1935, p. 16.

In addition to using rules and regulations and its fiscal power, the government can also influence the allocation of scarce resources through subsidies, tariffs, and quotas. Further discussion of these aspects of the command process in the U.S. economy is found in other sections of this chapter and throughout the text. In fact, Chapter 14 is devoted to a discussion of the role of government in the market economy.

The traditional process is also at work in the U.S. economy, but this process can be better understood by considering its impact on different countries throughout the world, particularly those whose economies are still developing. Examples of the traditional process are found in the eating habits, and in the patterns of work and social interaction, in such countries. Two examples of how the traditional process influences the allocation of scarce resources are religious restrictions on certain foods, such as beef and pork, and hiring practices based primarily on familial relationships. A branch of anthropology called economic anthropology is particularly concerned with the impact of customs and traditions on the economic questions of what, how, and for whom. In the business curriculum, students will find this subject of particular importance in courses on international business.

Because of the predominance of the market process in the U.S. economy, our discussion of the allocation of scarce resources is based on the assumption that managers operate primarily through the mechanisms of supply, demand, and material incentive (i.e., the profit motive). Their decisions about what goods to produce, how they should be produced, and for whom they should be produced are essentially market oriented. That is, firms choose to produce certain goods and services because, given the demand for these products and the cost of using scarce resources, they can earn sufficient profit to justify their particular use of these resources. Moreover, they combine their scarce resources to produce maximum output in the least costly way. Finally, they supply these goods and services to those segments of the population expected to provide the most material reward for their efforts.

Table 1.1 compares the three basic questions from the standpoint of a country and from the standpoint of a company, where they form the basis of the **economic decisions for the firm**. From the firm's point of view, question 1 is the product decision. At some particular time, a firm may decide to provide new or different goods or services or to stop providing a particular good or service. For example, consider Apple's decision to get into the music business by offering its iPod and iTunes music download service. Another good example is the various "non-computer-service" businesses that have gotten into the market for providing cloud-computing services. For example, telecom companies such as Verizon, AT&T, Deutsche Telecom, and British Telecom (BT) all provide cloud-computing services. Even companies such as Amazon and Dell have gotten into this market.

Question 2 is a basic part of a manager's responsibility. It involves personnel practices such as hiring and firing, as well as questions concerning the purchase of items ranging from raw materials to capital equipment. For example, the decision to automate certain clerical activities using a network of personal computers results

Table 1.1 The Three Basic Economic Questions

From the Standpoint of a Country	From the Standpoint of a Company
1. What goods and services should be produced?	1. The product decision
2. How should these goods and services be produced?	2. The hiring, staffing, procurement, and capital budgeting decisions
3. For whom should these goods and services be produced?	3. The market segmentation decision

in a more capital-intensive mode of production. The resolution to use more supplementary, part-time workers in place of full-time workers is another example of a management decision concerning how goods and services should be produced. A third example involves the selection of materials in the production of a certain item (e.g., the combination of steel, aluminum, and plastic used in an automobile).

The firm's decision concerning question 3 is not completely analogous to that of a country. Actually, a firm's decision regarding *market segmentation* (a term used in the marketing field) is closely related to question 1 for a country. In deciding on what segment of the market to focus, the firm is not literally deciding who gets the good or service. For example, suppose a firm decides to target a certain demographic segment by selling only a "high-end" or premium version of a product. However, the way in which a company markets the product (which includes its pricing and distribution policies) makes certain segments of the market more likely to purchase the product.

Perhaps one of the best ways to link the economic problem of making choices under conditions of scarcity with the tasks of a manager is to consider the view put forth by Professor Robert Anthony that a manager is essentially a person who is responsible for the allocation of a firm's scarce resources.¹⁰

It is interesting to note that "managers" or "management skills" was not delineated as a separate factor of production by early economic theorists. The four traditional categories of resources are land, labor, capital, and entrepreneurship. The last category can be treated as broad enough to include management, but the two classifications do involve different characteristics or skills.

The term *entrepreneurship* is generally associated with the ownership of the means of production. In addition, it implies willingness to take certain risks in the pursuit of goals (e.g., starting a new business, producing a new product, or providing a different kind of service). Management, in contrast, involves the ability to organize and administer various tasks in pursuit of certain objectives. An important part of a manager's job is to monitor and guide people in an organization. In the words of Peter Drucker, who has been called "the founding father of the science of management,"

It is "management" that determines what is needed and what has to be achieved [in an organization]. . . . Management is work. Indeed, it is the specific work of a modern society, the work that distinguishes our society from all earlier ones. . . . As work, management has its own skills, its own tools, its own techniques.¹¹

THE CASE OF GLOBAL FOODS, INC.: SITUATIONS AND SOLUTIONS

Prior sections of this chapter cited various reasons why an understanding of economics is important to managerial decision making. An effective way of demonstrating this importance is to cite real-world examples gleaned from the popular press and distilled from the findings of research studies on the use of economics in managerial decision making. All other texts in managerial economics do this, and this book is no exception. In addition, we want to show how economic terms and concepts can be applied

¹⁰Actually, Anthony divided the planning and control process in a firm into three activities: strategic planning (i.e., setting the firm's overall objectives), management control (i.e., making sure scarce resources are obtained and used effectively and efficiently in the firm's accomplishment of its objectives), and operational control (i.e., making sure specific tasks are carried out effectively and efficiently). These ideas were first put forth in R. N. Anthony, *Planning and Control Systems: A Framework for Analysis*, Boston: Harvard Business School, Division of Research, 1965.

¹¹Drucker, *Management*, p. xi.

to managerial decision making through the use of a series of hypothetical situations such as the one presented at the beginning of this chapter. In fact, each chapter begins with a *situation* requiring some sort of decision or action relating directly to the subject matter of the chapter. For example, in this chapter, a decision must be made about whether to enter the soft drink market. This is a fundamental business decision involving the allocation of a firm's scarce resources, a major theme of this chapter.

At the end of each chapter, a *solution* for the situation is presented based on the knowledge gained from reading the chapter. We use the term *solution* rather loosely because it may not involve a specific answer, as one might expect in the solution to a mathematical problem. In our view, the ambiguity of a solution is very much in keeping with conditions in the real world. Often in an actual business problem, there is no unique formula that one can use to compute the answer. Either the formula does not exist or is not entirely applicable to the problem, or the problem itself is not amenable to a straightforward quantitative solution technique. Even when a specific numerical solution is achieved—as is the case in Chapter 12 on capital budgeting—there may be other considerations of a qualitative nature that temper the acceptability of the solution. Therefore, the solutions offered at the end of the chapters are only suggested outcomes of the situations. (You may want to consider alternative ways for the managers depicted in the situations to deal with their tasks or problems.)

The situations used throughout the book are based on one industry and one firm in that industry. As you have already learned, we use the soft drink industry. Moreover, we follow the trials and tribulations of the managers of Global Foods, Inc., and, in certain cases, the managers of firms that do business with Global. This helps tie together the disparate aspects of economic analysis. Also, we believe that a focus on one firm in one industry creates added interest in the events depicted in the situations, further motivating mastery of each chapter's material.

A number of industries were initially considered. We chose the soft drink industry based on the following criteria:

1. The industry should be one that all readers can relate to as consumers.
2. The goods or services sold in the industry should be essentially nontechnical, and the means of production should be relatively easy for the layperson to understand.
3. The competitive environment should be intense.
4. Information about the industry should be readily available (e.g., from trade journals and research monographs), and news about current activities in this industry should be frequently reported in the popular media.

The soft drink industry closely meets all these criteria. Just about everyone consumes this product, and the product itself is rather simple: carbonated water, sweeteners, and various flavorings. The packaging is also uncomplicated. Soft drinks are sold today in 12-ounce aluminum cans, in 1- and 2-liter plastic bottles, and in glass bottles (particularly in foreign countries). The making of the syrup and the bottling of the beverage involve various manufacturing processes that are relatively easy to understand. The two most important trade publications in the soft drink industry are *Beverage World* and *Beverage Digest*. We found them to be excellent sources of background information on the industry. Moreover, the major soft drink companies are constantly reported on in major periodicals. Recent articles from these sources are cited throughout the text.

The situations used in each chapter, along with the characters portrayed at Global, are entirely fictitious. However, the features of each situation closely resemble actual business problems or circumstances that managers must often address. The verity of the issues involved in each situation is based on the authors' experiences in private industry, as well as on extensive interviews with managers from various companies

in the soft drink business. The following section gives a summary of the situations and solutions presented in the chapters. The main decisions to be made by the characters portrayed in the situations are included under the heading “Key Question.”

SUMMARY OF THE SITUATIONS AND SOLUTIONS

1 INTRODUCTION

Situation: Bob Burns, CEO of Global Foods, Inc., asks the board of directors to approve a decision to enter the soft drink business.

Key Question: “Should we enter the soft drink business?”

Solution: The board approves the decision, and Global Foods enters the soft drink business.

2 THE FIRM AND ITS GOALS

Situation: In an effort to increase the company’s revenues, Bob Burns considers entry into the market for energy drinks, a segment of the beverage market that has grown rapidly in recent years.

Key Question: “How can we improve the value of our company when the Wall Street analysts are judging us primarily on our ability to grow our revenue and profit?”

Solution: Bob decides to take Global Foods into the energy drink business.

3 SUPPLY AND DEMAND

Situation: Bob Burns, CEO, and Nicole Goodman, vice president of marketing, consider developing and launching a product line of gourmet tea.

Key Question: “What are the current and future supply and demand conditions in the consumer market for premium and specialty tea?”

Solution: Anecdotal information and a survey of articles in trade journals and business periodicals indicate that tea could be the next “big thing” in the U.S. beverage industry. However, Bob is not completely convinced and asks Nicole to do further quantitative analysis of the supply and demand for tea.

4 DEMAND ELASTICITY

Situation: Henry Caulfield, the proprietor of a Gas ’n Go convenience store, must evaluate the desirability of various pricing schedules for soft drinks set by the major beverage companies.

Key Question: “For what price should I sell this new soft drink?”

Solution: He decides that the relative inelasticity of the products in question makes it difficult to increase sales by lowering the price.

5 DEMAND ESTIMATION AND FORECASTING

Situation: Frank Robinson, newly appointed head of Global’s forecasting department, is asked to estimate the next year’s sales of Citronade, the company’s lemon-lime soda.

Key Question: “What will next year’s sales for Citronade be?”

Solution: Frank uses a trend analysis, adjusted seasonally as well as cyclically, to forecast the coming year’s sales.

6 THE THEORY AND ESTIMATION OF PRODUCTION

Situation: Christopher Lim, production manager, is concerned about the best way to bottle the water that the company now intends to sell. To differentiate Global Foods’

product in a highly competitive market, the marketing people want to use glass bottles. In Chris's view, this may help in the marketing of the product, but may well increase production costs significantly.

Key Question: "Should we package the water in glass bottles?"

Solution: Chris recommends that plastic bottles should be used to package both the carbonated soft drink and the bottled water products.

7 THE THEORY AND ESTIMATION OF COST

Situation: Shayna Soda Company, an independent bottler of Global Foods soft drinks, is looking for new ways to increase the profitability of its soda production. Adam Michaels, the plant manager, receives a marketing flyer from a new potential supplier in the next state, Lawrence Aluminum Products.

Key Question: "Should the company stay with its current main supplier in the next town, Kayla Containers, Inc., or switch to this new supplier?"

Solution: Shayna Soda Co. decides to stay with Kayla Containers, Inc., after considering all the other factors involved in the switch. Raw material costs were lower with Lawrence Aluminum, but other associated costs resulted in a net increase.

8 PRICING AND OUTPUT DECISIONS: PERFECT COMPETITION AND MONOPOLY

Situation: Frank Robinson is appointed product manager of the new energy drink product. One of his first tasks is to recommend a price for the product.

Key Question: "What price should we charge for our new product?"

Solution: After analyzing the demand elasticity and short-run cost structure of the product, Frank recommends a price based on the $MR = MC$ rule.

9 APPENDIX B BREAK-EVEN ANALYSIS (VOLUME–COST–PROFIT)

Situation: Suzanne Prescott, a senior analyst for the energy drink division, is asked to prepare a profit plan for the coming year.

Key Question: "What is the profit outlook for the coming year for our energy drink product, Go-nuts?"

Solution: She uses break-even analysis to forecast the coming year's profit for this product. She also uses sensitivity analysis to provide best-case and worst-case scenarios.

10 PRICING AND OUTPUT DECISIONS: MONOPOLISTIC COMPETITION AND OLIGOPOLY

Situation: The management committee of Global Foods asks Frank to reconsider his pricing recommendation because his analysis did not take into account certain competitive and market issues.

Key Question: "What is the best price for a product, given its demand, cost, and competition?"

Solution: After much debate, the management committee, with Frank's additional help, decides to set the price of its product slightly lower than the premium-priced competitors but slightly higher than the "value brands."

11 SPECIAL PRICING PRACTICES

Situation: Rebecca James must decide what price bid she should submit to a large airport caterer that wants to award a contract to a single supplier.

Key Question: “How should the bid price be set to give Global Foods a good shot at obtaining the large caterer’s contract?”

Solution: Because demand elasticities differ in different markets, the price offered in this price-sensitive market will have to be sufficiently low to give Global a good chance of winning the contract.

12 GAME THEORY AND ASYMMETRIC INFORMATION

Situation: Henry Caulfield’s daughter, Erica, believes she can help her father better understand and anticipate the pricing reactions of his closest competitors by applying to his business some of the principles of game theory that she learned in her college business courses.

Key Question: “How can managers deal with dynamic business conditions in which their decisions often trigger reactions by their competitors?”

Solution: Although game theory can help Henry understand the underlying dynamics of his pricing tactics against his competitors, in practical terms it falls short of providing him with a definite solution about what to do.

13 CAPITAL BUDGETING AND RISK

Situation: George Kline, the manager of Global’s capital planning department, is considering a proposal for the expansion of company activities into a new geographical region.

Key Question: “Should we expand into a new geographical area?”

Solution: Using capital budgeting techniques involving the calculation of net present value and internal rate of return, George recommends that the firm accept the project. He also performs a scenario analysis to present an optimistic and a pessimistic set of results.

14 MANAGERIAL ECONOMICS IN A GLOBAL ENVIRONMENT

Situation: Global Foods is interested in building a plant in Central Europe, where costs are lower than in Western Europe, but where productivity and quality are not as good as any country in which it operates around the world. George Kline, manager of Global’s capital planning department, is asked to conduct a capital budgeting analysis of an existing plant in the Czech Republic.

Key Question: “Should Global make an investment in a plant in the Czech Republic?”

Solution: George calculates the NPV and IRR for both the subsidiary and for the parent company. Although the results for the subsidiary are favorable, the NPV for the parent is negative. He therefore recommends that the company should not make the investment at this time.

15 THE ROLE OF GOVERNMENT IN THE MARKET ECONOMY

Situation: Because of growing concern about the move by NYC to limit the size of servings of carbonated soft drinks as well as the growing trend among consumers for healthier foods and beverages, Bob considers shifting the company into healthier foods and beverages.

Key Question: How do we respond proactively to the changing market and legal environment?

Solution: The senior leadership team advises Bob not to shift its focus too quickly or too intensively on the healthy food and beverage industry. This advice is based on the recent experiences of another major global food and beverage company that tried to expand their healthier products at the expense of its core carbonated soft drink products.

16 MANAGERIAL ECONOMICS IN ACTION: THE GLOBAL SOFT DRINKS MARKET

All the hypothetical “situations” and “solutions” in the previous chapters are based on facts and events in the global beverage industry. In this final chapter, we forego the mixture of fact and fiction in these narratives and simply discuss the actual market conditions in today’s global beverage industry. We then show how these market conditions can be explained in terms of the key factors affecting supply and demand discussed earlier in Chapters 3 and 4.

GLOBAL APPLICATION: THE BRIC COUNTRIES

In 2001, Jim O’Neill, chair of Goldman Sachs, came up with the acronym BRIC (Brazil, Russia, India, and China) for a paper he wrote called “The World Needs Better Economic BRICs.”¹² As a group the BRICs make up over 40 percent of the world’s population and about one-third of the land mass. The sheer size of their populations as well as their rapid economic growth over the past decade (with the possible exception of Russia) make them obvious attractions for companies seeking new markets in which to grow. Over the past decade, the gross domestic product (GDP) growth rates of the BRIC countries have been easily over twice the rate of the United States and the other developed countries around the world, with China and India enjoying rates that have been over three times faster. Estimated data for 2011 on the GDP by country reveals the BRIC countries to be among the top ten in GDP adjusted for purchasing power.

Rank	Country	GDP (Purchasing Power Parity)	Date of Information
1	European Union	\$ 15,390,000,000,000	2011 est.
2	United States	\$ 15,040,000,000,000	2011 est.
3	China	\$ 11,290,000,000,000	2011 est.
4	India	\$ 4,463,000,000,000	2011 est.
5	Japan	\$ 4,389,000,000,000	2011 est.
6	Germany	\$ 3,085,000,000,000	2011 est.
7	Russia	\$ 2,380,000,000,000	2011 est.
8	Brazil	\$ 2,282,000,000,000	2011 est.
9	United Kingdom	\$ 2,250,000,000,000	2011 est.
10	France	\$ 2,214,000,000,000	2011 est.

Source: <https://www.cia.gov/library/publications/the-world-factbook/rankorder/2001rank.html>.

Food and beverage companies have been very active in growing their businesses in the BRIC countries.¹³ For example, Coca-Cola considers its expansion over the past decade into the BRIC markets to be a major reason for the revival of its profitability and stock price in recent years.¹⁴ We will be providing more examples involving the BRIC countries in future chapters of this textbook. In fact, marketing in the BRIC countries will be a major theme in the concluding chapter.

¹²[http://en.wikipedia.org/wiki/Jim_O%27Neill_\(economist\)](http://en.wikipedia.org/wiki/Jim_O%27Neill_(economist))

¹³A useful summary of recent efforts by food and beverage companies to market health and wellness products to the BRIC countries can be found by going to: http://www.ghy.com/wennstrom_AF3_2011.pdf

¹⁴<http://www.compassintlconsulting.com/2012/01/25/qa-coca-cola-ceo-muhtar-kent/>

The Solution

After about an hour of heated debate, Bob had a suggestion to make to the board. “Look, we’ve been discussing this to such an extent that perhaps the key arguments I made in my presentation have gotten lost or confused. Let me summarize the seven key reasons why we want to enter the soft drink business, and then let’s vote on this matter.

1. *Outlook for the industry:* Prospects for growth in the industry continue to be strong. Therefore, we can expect the demand for our products to be a part of this positive industry trend.
2. *Market size and structure:* Although the industry is dominated by Coca-Cola and Pepsi-Cola, we believe there is still room for the entry of niche marketers. A number of regional and specialty companies have emerged over the past few years, particularly those offering sparkling fruit juices. We believe we can be as successful, if not more successful, than these new entrants.
3. *Manufacturing, packaging, and distribution:* Our experience in the manufacturing of food products will give us a significant head start when we enter the soft drink business. Moreover, we do not plan to build bottling facilities from scratch. Instead, we look to purchase and consolidate existing bottling plants currently owned and operated by independent firms or by multiple franchise operations. We are also encouraged by the number of new cost-reducing technologies that have been introduced and the fact that the cost of the artificial sweetener aspartame should decline now that Monsanto’s patent has expired.
4. *Transportation and distribution:* We already have a well-managed fleet of vehicles that deliver our food products. We also have important influence and contacts in the retail food business. These will be essential in establishing a presence on the shelves of supermarkets and convenience stores throughout the country.
5. *Pricing, advertising, and promotion:* As the ‘new kid on the block,’ we recognize that we will have to enter the market as a price follower. However, in time, as our products are developed and marketed, we should be able to establish some independence either to raise or lower our prices in comparison with the rest of the industry. As far as advertising and promotion are concerned, we have an excellent advertising agency that has served us well with our current product line. However, we shall be flexible enough to consider other agencies if the need arises. Moreover, our experience with various promotional programs (e.g., discount coupons through direct mail and magazine inserts) should be transferable to the soft drink industry. Most important, all advertising and promotional efforts should be greatly aided by the fact that our company name—Global Foods—enjoys a high degree of consumer recognition (along with the specific brand names of our products).
6. *New products:* As you have seen in the detailed report, we have exciting plans for several new products as well as a full line of naturally and artificially sweetened carbonated drinks. Through an independent market research company, we have tested consumer preferences for our new offerings. The results have been most satisfactory.
7. *Financial considerations:* As stated at the very beginning of this presentation, our main goal is to create value for our shareholders. We must continue to grow in a profitable manner if we are to continue satisfying the financial expectations of our shareholders. We compete in a mature industry that enables us to generate a considerable amount of cash from our current line of products with well-entrenched brand names. We believe that we should use this cash to expand into the soft drink business. As all our financial projections and analyses indicate, this effort should yield a rate of return that is more than enough to compensate us for the investment and its associated risk.”

After giving this executive summary, Bob Burns asked the board for a final decision. “All right, let’s vote. All those in favor of entering the soft drink market? Opposed? Great, it’s unanimous. Ladies and gentlemen, we’re going into the soft drink business.”

SUMMARY

Managerial economics is a discipline that combines microeconomic theory with management practice. Microeconomics is the study of how choices are made to allocate scarce resources with competing uses. An important function of a manager is to decide how to allocate a firm’s scarce resources. Examples of such decisions are the selection of a firm’s products or services, the hiring of personnel, the assigning of personnel to particular functions or tasks, the purchase of materials and equipment, and the pricing of products and services. This text shows how the application of economic theory and concepts helps managers make allocation decisions that are in the best economic interests of their firms.

Throughout the text, numerous examples are cited to illustrate how economic theory and concepts can be applied to management decision making. References are also made to business cases and economic events that have been reported in the popular press. However, a unique feature of this book is a unifying case study of a food and beverage company, Global Foods, Inc. Each chapter begins with a situation in which the managers of this firm have to make key economic decisions. The solutions that end the chapters suggest ways that economic analysis can assist in the decision-making process.

IMPORTANT CONCEPTS

Command process. The use of central planning and the directives of government authorities to answer the questions of *what, how, and for whom*. (p. 7)

Economic decisions for the firm. “What goods and services should be produced?”—the product decision. “How should these goods and services be produced?”—the hiring, staffing, and capital-budgeting decision. “For whom should these goods and services be produced?”—the market segmentation decision. (p. 8)

Economics. The study of how choices are made under conditions of scarcity. The basic economic problem can be defined as: “What goods and services should be produced and in what quantities?” “How should these goods and services be produced?” “For whom should these goods and services be produced?” (p. 2)

Managerial economics. The use of economic analysis to make business decisions involving the best use of a firm’s scarce resources. (p. 2)

Market process. The use of supply, demand, and material incentives to answer the questions of *what, how, and for whom*. (p. 7)

Opportunity cost. The amount or subjective value forgone in choosing one activity over the next best alternative. This cost must be considered whenever decisions are made under conditions of scarcity. (p. 6)

Resources. Also referred to as *factors of production* or *inputs*, economic analysis usually includes four basic types: land, labor, capital, and entrepreneurship. This chapter also includes managerial skills and entrepreneurship. (p. 5)

Scarcity. A condition that exists when resources are limited relative to the demand for their use. In the market process, the extent of this condition is reflected in the price of resources or the goods and services they produce. (p. 5)

Traditional process. The use of customs and traditions to answer the questions of *what, how, and for whom*. (p. 7)

QUESTIONS

1. Define *scarcity* and *opportunity cost*. What role do these two concepts play in the making of management decisions?
2. Elaborate on the basic economic questions of *what, how, and for whom*. Provide specific examples of these questions with respect to the use of a *country’s* scarce resources.
3. Following are examples of typical economic decisions made by the managers of a firm. Determine whether each is an example of *what, how, or for whom*.
 - a. Should the company make its own spare parts or buy them from an outside vendor?
 - b. Should the company continue to service the equipment that it sells or ask customers to use independent repair companies?
 - c. Should a company expand its business to international markets or concentrate on the domestic market?
 - d. Should the company replace its own communications network with a “virtual private network” that is owned and operated by another company?
 - e. Should the company buy or lease the fleet of trucks that it uses to transport its products to market?
4. Define the market process, the command process, and the traditional process. How does each process deal with the basic questions of *what, how, and for whom*?

5. Discuss the importance of the command process and the traditional process in the making of management decisions. Illustrate specific ways in which managers must take these two processes into account.
6. Explain the differences between management skills and entrepreneurship. Discuss how each factor contributes to the economic success of a business.
7. Compare and contrast microeconomics with macroeconomics. Although managerial economics is based primarily on microeconomics, explain why it is also important for managers to understand macroeconomics.
8. What do you think is the key to success in the soft drink industry? What chance do you think Global Foods has in succeeding in its new venture into the soft drink market? Explain. (Answer these questions on the basis of the information provided in the chapter and any other knowledge you might have about the food and beverage business.)
9. (Optional) Have you been personally involved in the making of a decision for a business concerning *what*, *how*, or *for whom*? If so, explain your rationale for making such decisions. Were these decisions guided by the market process, the command process, or the traditional process? Explain.

The Firm and Its Goals

Learning Objectives

Upon completion of this chapter, readers should be able to:

- Understand the reasons for the existence of firms and the meaning of transaction costs.
- Explain the economic goals of the firm and optimal decision making.
- Describe the meaning of the “principal-agent” problem.
- Distinguish between “profit maximization” and the “maximization of the wealth of shareholders.”
- Demonstrate the usefulness of Market Value Added[®] and Economic Value Added.[®]

The Situation

Bob Burns looked over the last few numbers provided to him in the consultant’s report on the bottled water industry, closed the binder and turned to Nicole Goodman, Global Foods’ vice president of marketing. “Looking back, our decision to get into the soft drink industry was a good one, but we need to get into new products and market segments in order to support our growth strategy. Our growth in bottled water has slowed because competition continues to be tough and because consumers are starting to resist paying for high priced alternatives to tap water.”

“I just happened to be reading a report from the people from *Food Science and Safety* about the growing popularity of energy drinks. They say it is a fast growing product in a key market segment: 18- to 34-year-olds. Why don’t we go into this market?”¹

“I hate to say this, Bob, but it may be a case of ‘too little, too late,’” said Nicole. “Established brands such as Red Bull have been around for a long time and dominate this market space. Red Bull is particularly popular in Asia, which as you know is a rapidly growing region for our soft drink products.”

“I’m not prepared to give up so easily, Nicole. The fact that energy drinks are popular in growth markets in Asia is all the more reason for us to consider getting into this business. And I really like the fact that this product is so popular among young people. We have good distribution channels, bottling know-how, and marketing savvy. This business is close to our core competency and it is still growing. We have told analysts that we expect double-digit growth of our top line as well as our bottom line. At our upcoming shareholder meeting, our shareholders and the analysts will be expecting to hear our plans for further business expansion. In the past, we did it with the help of soft drinks and bottled water. Now it is time to grow our entire beverage division with the help of energy drinks.”

¹M. A. Heckman, et al., “Energy Drinks: An Assessment of the Market Size, Consumer Demographics, Ingredients Profile, Functionality, and Regulation in the United States,” *Food Science and Food Safety*, Volume 9, 2010.

INTRODUCTION

Chapter 1 explains that managerial economics deals primarily with the problem of deciding how best to allocate a firm's scarce resources among competing uses. The best or optimal decision is the one that enables the firm to meet its desired objectives most closely. This chapter elaborates on the process of making decisions under conditions of scarcity by discussing the goals of a firm and the economic significance of the optimal decision. An online appendix explains the role of marginal analysis in economic decision making. This appendix also presents a review of the mathematics used in this text to illustrate key economic concepts and methods of analysis.

The major portion of this chapter is devoted to a discussion of the goals of the firm. However, to carry on this discussion sensibly, we must first define and explain *the firm*.

THE FIRM

The traditional (neoclassical) theory of economics defined the **firm** as a collection of resources that is transformed into products demanded by consumers. The costs at which the firm produces are governed by the available technology, and the amount it produces and the prices at which it sells are influenced by the structure of the markets in which it operates. The difference between the revenue it receives and the costs it incurs is *profit*. It is the aim of the firm to maximize its profit.

The preceding theory assumes the existence of the firm. But this leaves the reason for its existence unanswered. Why does a firm perform certain functions internally and others through the market? It appears that the size of the firm is not determined strictly by technological considerations. Then why are some firms small and others large?

Answers to the preceding questions began to appear in 1937 when Ronald Coase postulated that a company compares costs of organizing an activity internally with the cost of using the market system for its transactions.²

If there were no costs of dealing with the outside market, a firm would be organized so all its transactions would be with the outside. However, it is incorrect to assume the marketplace does not involve any costs. In dealing through the market, the firm incurs **transaction costs**.

Transaction costs are incurred when a company enters into a contract with other entities. These costs include the original investigation to find the outside firm, followed by the cost of negotiating a contract, and later, enforcing the contract and coordinating transactions. Transaction costs are influenced by uncertainty, frequency of recurrence, and asset specificity.³

Uncertainty, the inability to know the future perfectly, increases transaction costs because it is not possible to include all contingencies in a contract, particularly a long-term contract. Frequent transactions also tend to make it necessary for explicit contracts to exist.

But probably the most important of these characteristics is asset specificity. If a buyer contracts for a specialized product with just one seller, and furthermore, if the product necessitates the use of some specialized machinery, the two parties become tied to one another. In this case, future changes in market conditions (or in production technology) may lead to **opportunistic behavior**, where one of the parties seeks to take advantage of the other. In such cases, transaction costs will be very high.

²The seminal work in this area was by Ronald H. Coase in "The Nature of the Firm," *Economica*, 4 (1937), pp. 386–405, reprinted in R. H. Coase, *The Firm, the Market and the Law*, Chicago: University of Chicago Press, 1988, pp. 33–55. Coase was awarded the Nobel Prize in Economics in 1991.

³Much of this discussion is based on Oliver E. Williamson, "Transaction-Cost Economics: The Governance of Contractual Relations," *Journal of Law and Economics*, 22 (October 1979), pp. 233–61.

When transaction costs are high, a company may choose to provide the service or product itself. However, carrying out operations internally creates its own costs. A major cost is that, in hiring workers to do the work within the firm, the firm incurs monitoring and supervision costs to ensure the work is done efficiently. Quite possibly, employees who work for a fixed wage or salary may have less incentive to work efficiently than an outside contractor.

Employers will try to decrease monitoring costs by using incentives to increase employees' output. Among such incentives are bonuses, benefits, and perquisites ("perks"). Another popular incentive is to provide workers with the possibility of stock ownership, using stock options and employee stock plans. Stock ownership is also used to attract new employees. Such employees will, of course, benefit when the company is profitable and its stock increases in value. However, incentives come with a price tag.

The trade-off between external transaction costs and the cost of internal operations can be shown on the simple graph in Figure 2.1. When a company operates at the vertical axis, all its operations are conducted with the outside. As we move to the right on this graph, the firm substitutes internal for external operations. The cost of external transactions decreases, while the cost of internal operations increases. The total cost is the vertical summation of the two costs, and it decreases at first as the company finds that internalizing some operations is efficient. However, as more of the operations are internalized, some efficiency is lost, and the total cost begins to rise again. The company will choose to allocate its resources between external transactions and internal operations so the total cost is at a minimum, which in this case will occur about midway between the two extremes.

If transaction costs for a specific product or service are higher than the costs of carrying on the activity internally, then a company benefits from performing this particular task in-house. An independent firm may not find it profitable to produce a product if only a few customers demand it. However, as markets expand, the demand for a product or service, which may have been limited in the past, now expands. This will permit new firms to specialize in activities that previously had to be performed by the firm that needed this task to be performed. Thus new companies and industries come into existence. This is true not only in the case of products, but also for services that at one time were performed by the firm itself and are now produced by

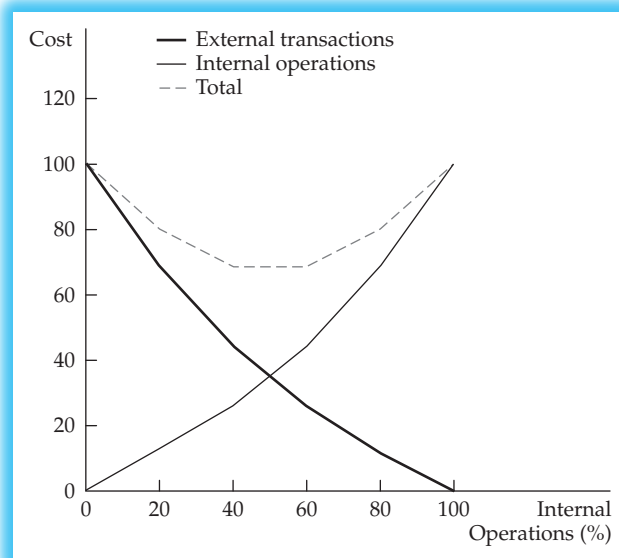


Figure 2.1
Trade-off Between
Transaction Costs and
Internal Operating Costs

independent firms—for example, cleaning services, security services, and cafeterias are often now run by specialized firms. Another example would be the college bookstore that is operated by one of the large companies in the book industry. This idea is actually rather old. It really started with Adam Smith, who stated that “the division of labor is limited by the extent of the market.” George Stigler discussed this point in a 1951 article, and concluded that as industries expand, companies that previously had produced everything internally will tend to experience “vertical disintegration.”⁴ What has happened, of course, is that transaction costs have decreased and that the possibility of “opportunistic behavior” has also diminished.

Although the outsourcing of peripheral, noncore activities has been around for a long time, the outsourcing of a business’s core activities is a somewhat more recent activity. In the past we have seen the outsourcing of private-label or house-brand merchandise. However, now the outsourcing of highly technical products and services is widespread. But outsourcing cannot only occur within a given country. When production of goods and services is transferred to another country, we call it offshoring. Transferring production of manufactured goods to foreign countries has been going on for a long time. Now, because of great advances in information technology and communication, offshoring of technical services, including call centers, accounting, human resources and legal services, and even engineering and research, have become quite common. The question of the benefits of offshoring has been a very controversial topic among economists and it has also become a political issue.

By transferring various services from developed economies to developing countries where costs are lower, developed economies become more competitive, which tends to increase their gross domestic product (GDP). Further, developing countries gain employment and production and this tends to increase their GDP.

However, there are some disadvantages for the developed economies: for example, a potential loss of jobs, and also the question of the quality of products and services as well as the decreased amount of control that the originating countries incur. In addition, the offshoring of technical work and knowledge may erode the developed countries’ competitive advantage.

While we read about offshoring in the press almost every day, there are some small signs of certain operations returning to the United States as well as to other developed countries, a phenomenon called reshoring. There are several reasons for this trend.

Wages in developing countries have been rising. This is particularly true of China, where wages are reported to have been rising by 15 percent per year, even though they are still quite low compared to wages in developed countries. In addition, the decrease in the value of the dollar has increased the cost of importing. Increases in energy costs have made it more expensive to ship products from faraway countries. Communications and quality control have often been instrumental in increasing costs. The lead time for orders from foreign countries is longer. Due to great improvements in technology in U.S. manufacturing, there has been a significant increase in productivity, thus making U.S. production more competitive.

However, reshoring has been quite limited so far and probably has had very little effect on employment in U.S. manufacturing. There is another facet that must be considered. While costs have been rising in some of the countries to which products and services have been offshored, there are still many countries where production has remained quite inexpensive. So, while some of the countries may be losing their

⁴George J. Stigler, “The Division of Labor Is Limited by the Extent of the Market,” *Journal of Political Economy*, 59 (June 1951), pp. 185–93.

advantage, production of goods and services will shift to the still-low-cost countries rather than back to the high-wage economies.⁵

Coase and the Internet

When Ronald Coase wrote his article in 1937, he and the rest of the world knew nothing about the advent of the Internet in the last years of the twentieth century. However, his contribution to economic theory turns out to have great relevance to today's business transactions over the Internet. If you should surf the Internet and type in the following keywords, "Transaction costs Coase Internet" you will get a large number of hits attesting to today's popularity of Coase's ideas.

The basic idea of the trade-off between the costs of internal operations and external transactions remains as valid as ever. But the revolutionary event that occurred in recent years is that the Internet has caused transaction costs to decrease drastically, making it easier and more efficient for companies to curtail their own operations and farm out much of the work they would have been performing to outside companies that specialize in specific operations.

As mentioned previously, transaction costs include the cost of search and investigation, contract negotiations, and coordination or enforcement. How has the Internet affected these transactions and their costs?

Search and investigation have been made significantly easier. Potential suppliers can be quickly and easily identified. Information on their reliability and credit standing is readily available, and so are evaluations of their financial conditions. "Online clearing houses . . . permit a purchaser to contract for price, quality and delivery dates with few clicks of the mouse," making contracting a much easier task. As far as coordinating costs are concerned, it is now much more simple to follow a shipment as it progresses toward its location, and to take action in real time, if this should become necessary.⁶

Although Coase has stated that he is not paying much attention to e-commerce, he believes that understanding transaction costs in the new economy "enables you to have more specialization and greater production, because you are more efficient. You'll get more small firms as a result, but large firms will also get larger, because they can concentrate on core activities and contract out what they can't do well."⁷

We end this discussion with the conclusion that a firm will trade off costs incurred in conducting transactions with the outside market with the costs of internalizing such transactions in order to minimize the combination of the two. This is consistent with the overall economic goals of the firm, the subject to which we turn next.

THE ECONOMIC GOAL OF THE FIRM AND OPTIMAL DECISION MAKING

Every business has a goal. Most students would assert that the primary goal of a business is to earn a certain amount of profit (i.e., to "make money"), and, in fact, the economic theory of the firm—the foundation on which much of managerial economics

⁵Much of the material on reshoring has been obtained from the following sources: Stan Abrams, "Thoughts on the Reshoring Mini-Wave," *Business Insider*, July 6, 2011; Bob Adelman, "ReShoring: American Manufacturing Jobs Come Home," *The New American*, October 18, 2011; Peter Marsh, "China labour costs push jobs back to US," *Financial Times*, October 6, 2011; Tom Leunig, "Stop thinking of 'reshoring' jobs from China. It just won't happen," *Financial Times (USA)*, October 31, 2011.

⁶This section, including the direct quotation, is based on Don Tapscott, David Ticoll, and Alex Lowy, "Internet Nirvana," *eCompany Now*, December 2000, pp. 98ff.

⁷Bob Tedeschi, "Coase's Ideas Flourish in the Internet Economy," *New York Times*, October 2, 2000.

rests—assumes the principal objective of a firm is to maximize its profits (or minimize its losses).⁸ Thus, throughout this text, unless otherwise stated, we assume this same objective, known among economists as the **profit maximization hypothesis**.

To be sure, there are other goals that a firm can pursue, relating to market share, revenue growth, profit margin, return on investment, technology, customer satisfaction, and shareholder value (i.e., maximizing the price of its stock). It is crucial to be precisely aware of a firm's goals. Different goals can lead to very different managerial decisions given the same limited amount of resources. For example, if the main goal of the firm is to maximize market share rather than profit, the firm might decide to reduce its prices. If the main goal is to provide the most technologically advanced products, the firm might well decide to allocate more resources to research and development. The added research and development expenses would most likely reduce the amount of profit the firm earns in the short run but may result in increased profits over time as the company increases its technological lead over its competitors. If the main goal of the firm is to carry a complete line of products and services, it may choose to sell certain products even though they might not be earning a profit.

Given the goal (or goals) that the firm is pursuing, we can say that the **optimal decision** in managerial economics is one that brings the firm closest to this goal. For example, as you see in Chapter 8, to maximize its profit (or minimize its loss), a firm should price its product at a level where the revenue earned on the last unit of a product sold (called *marginal revenue*) is equal to the additional cost of making this last unit (called *marginal cost*). In other words, the optimal price equates the firm's marginal revenue with its marginal cost.

One additional concept should be presented in our discussion of a firm's goals. In economics, a distinction is made between the “short-run” time period and the “long-run” time period. As explained in greater detail in later sections of this text (see Chapters 3, 6, and 7), these time periods actually have nothing directly to do with calendar time. During the short run, we assume a firm can vary the amount of certain resources (e.g., labor hours) but must operate with a fixed amount of at least one of its resources (e.g., factory space). Theoretically, in the long run, a firm is able to vary the quantities of all resources being used. In this text, we look at both short-run and long-run decisions made by the firm. We assume a company's goal is to maximize profits both in the short and long run. However, it must be understood that a business will, at times, sacrifice profitability in the short run with the anticipation of maximized long-run profits.

GOALS OTHER THAN PROFIT

Economic Goals

The concept of profit maximization has been attacked as incomplete by many writers. They point out that companies may have other economic objectives, such as those mentioned previously.

For the time being, we omit discussion of the objective of “value” or “shareholder wealth” maximization and consider some of the other alternatives concerning a company's activity during a single period of time (such as a year). It is readily

⁸As we will see in Chapter 8, a firm may lose money in the short run and still be better off than it would be if it were to shut down operations, as long as its losses are less than its fixed costs. However, if it is going to lose money, from an economic standpoint it is optimal to minimize its losses.

admitted that profit maximization is a rather vague term. How does a company know that its profits in a given period are the largest they can be? Or, more correctly (from an ex ante, or planning, viewpoint), how does a company know that the actions it is taking in this time frame will result, if all goes as expected, in the greatest possible profit?

Let us look at the objectives set out by a company's CEO (or a committee representing the company's top management). It is not unusual for the CEO or his or her representatives, having decided on the achievable results for the next fiscal period, to distribute objectives to the various operating heads at the beginning of the planning cycle. Now imagine this memorandum from the firm's CEO to the general manager of one of the company's operating units:

Dear Alex,

We have had a pretty good year in 2012, and we believe that 2013 should be even better.

I am therefore issuing the following objective for your unit in 2013. Take any and all actions that will ensure your profit is maximized.

Corporate management is confident that you will not disappoint us. We know that the objective we have given you is challenging. We also are convinced that it is achievable.

ERIC, CEO

This memorandum is obviously an extreme simplification, but what is Alex to do with his marching orders to maximize profit? What resources does he have to do this? How can his performance be measured at the end of the year? What is his maximum profit?

Now let us look at another "objective" memorandum:

Dear Alex,

We have had a pretty good year in 2012, and we expect that 2013 should be even better. We are assigning specific objectives to each of our operating units in such a way that the total result will be a financial posture consistent with our economic and industry forecasts, our available resources, and good increases in productivity. With this in mind, we want you to build your 2013 plan to correspond to the following objectives for your unit:

1. Your revenue should increase by 10 percent from 2012.
2. The profit margin of your unit should increase from 8 percent to 9 percent, and your return on assets should be 10 percent.
3. Your division will receive \$10 million of company funds for expansion projects whose minimum internal rate of return should be 12 percent.
4. The head count of your unit can increase by no more than 2 percent.

Corporate management is confident that you will not disappoint us. We know that the objective we have given you is challenging. We also are convinced that it is achievable.

ERIC, CEO

Assuming that this memorandum makes more sense (which it certainly should, for otherwise our point has been lost), does this mean that the company's objective is not really profit maximization, but rather a growth rate, a profit margin, or a return on its assets? This is what many writers on this subject say.

Such a conclusion is, however, misleading. Any of these measures in itself is incomplete, and each should be seen as a realistic target consistent with the ultimate objective of maximizing the firm's overall profits. Management, in this example, advised by its expert staff regarding the company's economic environment, competition, technological advances, and market potential, has come to the conclusion that maximum profits can be achieved by the combination of growth and profit measures included in its memorandum.

Thus, the specific objectives assigned to an operating unit are really proxies for the overall objective of profit maximization. The achievement of these proxies is also measurable at the end of the fiscal period; the division executive's performance and contribution toward the company's profits can be evaluated, and rewards in terms of bonuses or incentive plans can then be determined.

Noneconomic Objectives

In this complex world, companies may have objectives that are not strictly economic or at least do not appear to be governed by economic thinking. Indeed, some large companies have published statements of principles that, if accepted at face value, would indicate that making profits is the last thing for which they strive. Profits may be mentioned as only one of several objectives, and they may actually be listed last. Furthermore, the statements do not mention any maximum but rather concentrate on such measures as "adequate" or "reasonable" return to stockholders. Such modesty is certainly more palatable to the public. What, then, are some of the guiding principles such companies publish?

1. Provide a good place for our employees to work.
2. Provide good products/services to our customers.
3. Act as a good citizen in our society.

These actions are costly, and at first glance may seem to interfere with profit maximization. However, consider the following: Satisfied employees not only tend to be more productive, but will remain with the company longer, thus decreasing expensive labor turnover. Without satisfied customers, a company will not remain in business. Supporting good causes, such as charitable and other nonprofit organizations, will create goodwill and ultimately potential sales. Therefore, it would be worthwhile for a company to spend resources on such **noneconomic objectives** consistent with increases in revenues and profit. If this is the case, then attaining these objectives is not incompatible with profit maximization, and indeed, these objectives could be classified as economic.

We could enlarge this discussion of so-called noneconomic objectives, but the point has been made. Today's markets and institutions constrain companies in many ways that did not exist in the past. Therefore, companies must concern themselves with creating employee and customer satisfaction and maintaining social responsibility to a much higher degree than in the past. But these considerations do not contradict the profit maximization principle. If companies were maximizers in the past, under less restrictive conditions, they are still maximizers today but have to operate within the requirements imposed by current standards and the costs that accompany them.

DO COMPANIES REALLY TRY TO *MAXIMIZE PROFITS*?

There still remains for us to discuss a major criticism, that has been leveled at the view of profit maximization as a company's primary objective.

The argument is that today's corporations do not maximize. Instead, their aim is to "satisfice." To understand this argument, we have to consider two parts of this idea:

1. The position and power of stockholders in today's corporation
2. The position and power of professional management in today's corporation

Years ago the owner or owners of a business also managed it. Businesses were predominantly quite small and lent themselves to being operated as individual proprietorships, partnerships, or small, closely held corporations. Modern businesses, particularly medium-size or large corporations, of course, cannot be managed by the owners, who are the shareholders and number in the thousands or even hundreds of thousands. Many stockholders own only minute pieces of a corporation. Furthermore, stockholders tend to diversify their holdings; thus, they may hold small interests in many different corporations. The argument asserts that most stockholders are not well informed on how well a corporation can do and will be satisfied with an adequate dividend and some reasonable growth. Because they own different stocks, poor performance on one of their holdings may be offset by some of their other assets. The stockholder is more concerned with the portfolio of stocks than with any individual stock. Shareholders may not be capable of knowing whether corporate management is doing its best for them, and they actually may not be very concerned as long as they receive what they consider a satisfactory return on their investment—hence "**satisficing**."

Second, in a modern corporation professional managers—the chair of the board, the president, a group of vice presidents, and other high-level managers—direct the operations of a company. Although they are overseen by a board of directors (which often includes a large number of insiders), they are responsible for major decision making. It is claimed by a number of writers that managers (who commonly hold a relatively small number of shares) have their own objectives, which do not include maximization of shareholder earnings. Indeed, it is often said that managers tend to be more conservative—that is, risk averse—than stockholders would be because their jobs will most likely be safer if they turn in a competent and steady, if unspectacular, performance. They could probably benefit stockholders in the long run by taking some well-calculated risks. However, they may be too cautious to do so, and thus they miss opportunities. They fear that they may not survive the reverses that could result from risk taking. If stockholders need only be satisfied, this may be the appropriate way for management to go.

Management's interests may actually be contrary to those of stockholders. For instance, management may be more interested in revenue growth than profits. Why? It has been claimed that management remuneration tends to be a function of revenue size rather than profits. Several studies have been made on this subject, but the evidence is considerably less than overwhelming. Also, company management may be more interested in maximizing its own income, may indulge in various perquisites, and in general may not act in the best interest of the widely dispersed, somewhat disinterested and lethargic stockholder population. The divergence in the objectives between owners and management has been the subject of much discussion in economic

literature and is known as the “principal-agent” problem or simply as the “agency problem.”⁹

The two sides in this relationship tend to complement one another. The owners of the corporation—the stockholders—are not interested in maximization, or even if they are, they are not well informed and have too little power. The corporation’s management, whose selfish motives lead them to act in their own favor when stockholder and management goals differ, will manage in a way that serves their interest, while keeping the stockholder satisfied with adequate return and moderate growth.

Like all ideas presented by intelligent people, this one probably contains a certain amount of truth. Each of the points seems eminently reasonable and, for all we know, could be valid over limited periods of time. However, let us look at some of the realities of life and some recent events in the business world that tend to contradict this argument.

You, the reader of this book, may be among that group of far-flung stockholders owning a hundred shares in a company with millions of shares outstanding. However, particularly in the case of large corporations, much of the outstanding stock is held by institutions in professionally managed accounts. Among these are banks that manage large pension funds, insurance companies with their extensive portfolios, and mutual funds. These organizations employ expert analysts (who are only human, and therefore, at least occasionally make mistakes) who study companies and pass judgment on the quality of their management and their promise for the future. Of course, they deal mostly with stock prices, but after all, stock prices are a reflection of a company’s profitability.¹⁰ These analysts make recommendations to their management on which stocks to buy and which to sell. Companies that underperform would be weeded out of these institutional portfolios, with a consequent drop in their stock prices.

Now, what happens when certain stocks tend to underperform in the market? They become targets for takeovers by others. We really do not have to belabor this point because anyone reading the business sections of daily newspapers or other business publications is very much aware of recent events in the takeover and buyout arenas. In addition to the accumulation of stock and subsequent tender offers by outside financiers, we have also witnessed the existence of proxy fights by dissident large stockholders. Thus, it appears that management in today’s corporation is not insulated from outside pressures. Management is constrained to act in agreement with stockholders, who look for increases in stock values and returns and who act to “punish” the managements of those companies that appear to underperform.

Another argument leads to a similar conclusion. Competitive pressures also act to stimulate management to performance. If a company’s results lag behind those of competitors, those lethargic stockholders who do not challenge the company directly will tend to sell its shares and turn to those companies providing better returns and better prospects of returns. The price of the company’s stock will suffer relative to prices of the others; such a scenario will not go unnoticed in financial markets.

⁹A formal theory dealing with the potential conflicts between shareholders and management was developed by Michael C. Jensen and William H. Meckling in their article “Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure” (*Journal of Financial Economics*, October 1976, pp. 350–60). These conflicts arise whenever managers own less than 100 percent of the stock, which is, of course, the predominant situation in today’s large corporation. To ensure managers act on behalf of the stockholder, the latter will have to incur “agency costs,” which are expenditures to monitor managers’ actions, to structure the organization in such a way as to limit management’s action, and so forth.

¹⁰The connection between profits and stock prices is examined in the next section, when we expand the maximization principle to include the wealth of stockholders.

Company management will come under the gun to improve performance, and ultimately management may be replaced because of pressure by outside board members, a successful proxy fight, or even a takeover. A very vocal and sometimes effective advocate of shareholder rights has been the California Public Employees' Retirement System (Calpers). With about \$236 billion in assets as of March 31, 2012, it has demonstrated that changes in corporate governance can be accomplished. Calpers reviews the performance of companies in its investment portfolio and selects companies that have performed poorly and for which it may seek a change in corporate governance.¹¹ Until 2010, Calpers published a "Focus List" naming companies that performed poorly. Now, instead of publishing such a list, it requests meetings with representatives of these companies to suggest specified reforms.

The Sarbanes-Oxley Act was passed in 2002 in response to a number of corporate scandals. The act sets new, stricter standards on the behavior of public corporations and accounting firms. Since then, shareholders have become much more active in proposing changes in corporate policies in proxies that are to be considered at annual stockholder meetings. Several of these reforms, such as changing the way by which directors are elected and stockholders casting nonbinding votes on executive pay, have received much greater consideration and in some cases actually were adopted.¹²

In addition, there is the managerial labor market. Managers who have performed well for their stockholders will most likely be in greater demand and will be better compensated than managers with mediocre records.

Management has another, more direct, motivation to act in concert with the objectives of stockholders. Parts—frequently large parts—of an executive's remuneration are tied to performance in terms of operating profits for the corporation or for units supervised by the particular executive.¹³ Furthermore, an executive's compensation package is usually enhanced by the issuance of stock options. Because the value of stock options depends on the price of the company's stock, which in turn is a function of the company's profit performance, self-serving company managers may find that their objectives (less than miraculously) coincide with those of the stockholders.

Profit Maximization, Restated

It is readily agreed that the existence of the profit maximization objective can never be proven conclusively. We must note, however, that lack of financial success by a company is not necessarily a contravention of the principle. The best of plans may go awry, and management's judgment certainly is not error-proof. Under certain circumstances, the aim for loss minimization may replace the goal of profit maximization, but this too supports our basic premise. As difficult as it is to point to acts of profit maximization by management, none of the alternative constructions lends itself as well as a yardstick by which to measure business activity. As long as a corporation strives to do better—that is, prefers higher profits to lower profits and lower costs to higher costs, and acts consistently in those directions—the assumption of profit maximization serves as a better basis for judging a company's decisions than any of the other purported objectives. Incidentally, this "striving to do better" can include a multitude of decisions, including those that lead to a revenue increase greater than a cost increase, a revenue decrease smaller than a cost decrease, or a constant revenue with decreased costs. All these decisions involve an increase in profits.

¹¹Much information about Calpers can be obtained by visiting www.calpers.org or www.calpers.ca.gov.

¹²Jena McGregor, "Activist Investors Get More Respect," *BusinessWeek*, June 11, 2007, pp. 34–35.

¹³The fact that these performance incentives may be tied to near-term profits can create a problem because the executive's horizon may be shortened.

However, maximizing profits in the very short term (e.g., 1 year) can always be accomplished by management. If, for instance, revenue in the coming year is expected to decline, a company can keep up its profits by cutting expenses. If management seeks to do this without an immediate further reaction on revenue, it can eliminate some development projects. The effect of a lack of new products will not be felt right away, but the shortsightedness of this management decision will come home to roost in a few years. This is the decision area in which the objective of period profit maximization can be attacked more logically. Profit maximization for one period is an incomplete measure from the viewpoint of a business organization that is expected to operate into the infinite—or at least the foreseeable—future.

MAXIMIZING THE WEALTH OF STOCKHOLDERS

Because period profit maximization is an extremely useful way to look at day-to-day decision making in the firm, we use it as our model throughout most of this book. However, there is another view of maximization that is usually adopted in finance textbooks and that takes into consideration a stream of earnings over time. This concept includes not only the evaluation of a stream of cash flows; it also considers the all-important idea of the time value of money.¹⁴ Because it is an obvious fact that a dollar earned in the future is worth less than a dollar earned today, the future streams must be discounted to the present. Both the shape of these streams through time and the interest rate at which they are discounted affect the value of the stockholders' wealth today. The discount rate in particular is affected by risk, so risk becomes another component of the valuation of the business. Financial theorists differentiate various types of risk, with the two major types commonly identified as business risk and financial risk.

Business risk involves variation in returns due to the ups and downs of the economy, the industry, and the firm. This is the kind of risk that attends all business organizations, although to varying degrees. Some businesses are relatively stable from period to period, whereas others incur extreme fluctuations in their financial returns. For instance, public utilities (i.e., suppliers of electricity and gas, as well as the operating telephone companies) tend to have more stable earnings over time than do industrial companies, particularly those in industries that are highly cyclical (e.g., steel, automobiles, and capital goods), or companies in high-tech fields.

Financial risk concerns the variation in returns that is induced by leverage. *Leverage* signifies the proportion of a company financed by debt. Given a certain degree of leverage, the earnings accruing to stockholders will fluctuate with total profits (before the deduction of interest and taxes). The higher the leverage, the greater the potential fluctuations in stockholder earnings. Thus, financial risk moves directly with a company's leverage.

How do we obtain a measure of stockholders' wealth? We do so by discounting to the present the cash streams that stockholders expect to receive out into the future. Because we know today's price of a company's stock, we can—given the expected dividends to be received by the stockholders—determine the discount rate the investment community applies to the particular stock. This discount rate includes the pure time value of money and the premiums for the two categories of risk. The dividend stream is used to represent the receipts of stockholders because that is all they really receive from the company. Of course, a stockholder also looks for a capital gain, but selling the stock at some point involves someone else buying it; thus, this



¹⁴Time value of money and discounting of cash flows is discussed in greater detail in the appendix found on the Companion Website.

payment represents only a trade, an exchange of funds. However, dividends represent the returns on the stock generated by the corporation. In equation form, we have the following:

$$P = \frac{D_1}{(1+k)} + \frac{D_2}{(1+k)^2} + \frac{D_3}{(1+k)^3} + \dots + \frac{D_n}{(1+k)^n}$$

where P = present price of stock

D = dividends received per year (in year 1, year 2, . . . year n)

k = discount rate applied by financial community, often referred to as cost of equity capital of a company

If it is assumed that the corporation will have an infinitely long life and dividends will remain the same year after year, then the price of each share of stock can be calculated as a perpetuity with the following formula:

$$P = D/k$$

Investors, however, will usually expect dividends to rise. In the case where dividends grow at a constant rate each year, the formula for share price becomes

$$P = D_1/(k - g)$$

where D_1 = dividend to be paid during coming year

g = annual constant growth rate of dividend expressed as a percentage¹⁵

Multiplying P by the number of shares outstanding gives the total value of the company's common equity.¹⁶

A simple example will help clarify the previous equation. Assume that a company expects to pay a dividend of \$4 in the coming year, and expects dividends to grow at 5 percent each year. The rate at which stockholders discount their cash flows (which is really the rate of return stockholders require to earn from this stock) is 12 percent. There are 1 million shares outstanding. We would expect the price of each share to be

$$P = 4/(0.12 - 0.05) = 4/.07 = \$57.14$$

The value of the company's stock would be \$57.14 million. This is the expected market value given the variables that we have assumed. However, this may not be the maximum value the company could achieve. The variables in the equation may have to change. Because k is a function of the company's level of risk (both business and financial), the company may be able to decrease k by lowering the riskiness of its operations or by changing its leverage. It can affect g and D by retaining more or less of its earnings. By retaining a larger portion of its earnings and devoting a smaller portion of its earnings to dividends, the company may be able to increase its growth rate, g .

Thus, under this construction, maximizing the wealth of the shareholder means that a company tries to manage its business in such a way that the dividends over time paid from its earnings and the risk incurred to bring about the stream of dividends always create the highest price and thereby the maximum value for the company's stock.

This **wealth maximization** hypothesis tends to weaken even further the management versus stockholder argument. Corporate executives, for whom stock options

¹⁵The derivation of these formulas is discussed in greater detail in Chapter 12.

¹⁶The value of a company's equity can also be obtained by calculating the present value of the expected stream of "free cash flows." However, when free cash flow is correctly constructed, it is essentially equal to dividends paid. This subject is discussed at greater length in Appendix 12A when the calculation of the value of a corporation is presented.

represent a significant portion of remuneration, now have an even greater incentive to aim at results that conform to the objectives of the stockholders.

This is a rather complex if quite obvious development of the maximization principle. As stated previously, we work primarily with the profit maximization hypothesis because it is quite sufficient for most of our purposes. We return to the wealth maximization rule in Chapter 12 when we discuss a company's investment and replacement decisions involving expenditures for which the resulting payoffs flow into the corporation over a considerable period of time. In that chapter, we also briefly discuss how the market tends to determine the rate of return it requires from a company (and thus sets the discount rate k , the company's cost of capital). In Chapter 12, we also examine the question of risk and uncertainty and attempt to find ways to deal with it.

Market Value Added and Economic Value Added

Various publications have measured the wealth of stockholders by taking the price per share quoted in the stock market pages and multiplying it by the number of shares outstanding. The product is, of course, the current value of the shares, and thus reflects the value of the company accorded to it by the market. However, such a measure does not show the wealth that has been created by the company. After all, suppose the stockholders had contributed more capital than the stock was worth currently. Then, actually, the company would have "destroyed" some of the stockholders' wealth. What is really important is how much the stockholders' investment is worth today relative to what they have contributed to the corporation in originally buying the stock and then having earnings retained by the corporation for reinvestment.

A relatively new measure has become popular with the financial community as well as with many corporations. It is called **Market Value Added (MVA[®])** and has been developed by the consulting firm of Stern Stewart.¹⁷ (The firm is now called EVA Dimensions, LLC.) MVA represents the difference between the market value of the company and the capital that investors have paid into the company.

The market value of the company includes the value of both equity and debt. The capital includes the book value of debt and equity on the company's balance sheet plus a number of adjustments that increase the basic number. Among these adjustments is the inclusion of research and development (R&D) expense (which accountants treat as expense). Prior years' R&D is cumulated and amortized over a number of years. Another item that is included is the amortization of goodwill. Thus, in the end, the contributed capital of the corporation will turn out to be larger than merely the book value of equity and debt. Although the market value of a corporation will always be positive, the MVA may be positive or negative, depending on whether the market value of the company is greater than the capital that investors contributed. Where a corporation's market value is less than the contributed capital, investors' wealth has actually been "destroyed."

A recent ranking of 3,000 corporations based on 2011 stock prices showed Apple at the top, with an MVA of \$423 billion, followed by Exxon (at \$222 billion), Microsoft (at \$184 billion), and IBM (at \$162 billion). At the other end of the scale were the American International Group, Citigroup, and Bank of America, which actually showed negative MVAs.¹⁸

Basically, MVA is a forward-looking measure. If market value reflects the financial markets' appraisal of a company's future cash streams, then MVA represents the

¹⁷This concept was originally introduced in 1990. See G. Bennett Stewart III, "Announcing the Stern Stewart Performance 1,000: A New Way of Viewing Corporate America," *Journal of Applied Corporate Finance*, Summer 1990, pp. 38–59.

¹⁸"Ranking of Companies by MVA," EVA Dimensions LLC, 2011. Such a measure favors large companies and penalizes smaller companies. To show the relative market value added, one could divide the MVA by the company's annual revenue.

financial markets' assessment of the company's future net cash flows (i.e., after subtracting the investments the company must make to achieve those cash streams).

Another measurement developed by Stern Stewart is **Economic Value Added (EVA[®])**. EVA is calculated as follows:

$$\text{EVA} = (\text{Return on Total Capital} - \text{Cost of Capital}) \times \text{Total Capital}$$

Actually, the calculation of return on capital (profit divided by capital) is nothing new. However, EVA subtracts an estimated cost of capital from return. If the resulting number is positive, then the company has earned more than its investors require, and thus will add to investors' wealth. In contrast, if cost is greater than return, then value is being destroyed. While EVA can be calculated for past periods to see how a company has been performing, it can also be used in evaluating future plans.

To avoid distortions created by accounting conventions, Stern Stewart makes numerous adjustments to the return and capital numbers. Actually, EVA could be said to be very much like "economic profits," which are mentioned briefly in the next section of this chapter and are discussed thoroughly in Chapter 9. However, when these numbers are calculated they are generally based on past results and do not necessarily say anything about a company's future profitability. Still, "Stern Stewart says that there is a close correlation between EVA and MVA—if managers improve EVA, the company's MVA is highly likely to improve too."¹⁹

Over the last few years, many companies have begun emphasizing the EVA measure over more traditional measures such as earnings per share and return on equity, as have money managers such as Oppenheimer, Calpers, and others.²⁰

EVA Dimensions has also added a new measure, called EVA Momentum. It basically measures the growth rate of EVA by dividing the change in EVA for a given period by the company's sales at the beginning of the period.

ECONOMIC PROFITS

Throughout this chapter, we use the term *profit* and assume it has some kind of meaning. But we have not defined it. We only said that profit—and its maximization—is uppermost in the company owner's and manager's minds. In a way, profit is easy to define. Every company that closes its books annually and whose accountants construct a statement of earnings (whether this company is public so everybody can see the published statement and its "bottom line," or whether it is private) knows its profits. The accountants report the level of profits, and they also affirm that everything in the financial statements has been done in conformance with generally accepted accounting principles (GAAP).

Unfortunately, things are not quite that simple. Profits as reported on an earnings statement are not necessarily definitive. Accountants have certain amounts of freedom in recording items leading to the "bottom line."²¹ A few examples will suffice:

1. There are different ways of recording depreciation. In the past, the straight-line method, the sum-of-the-years'-digits method, the declining balance method, and probably others

¹⁹"A Star to Sail By?" *The Economist*, August 2, 1997, p. 54.

²⁰S. Tully, "America's Greatest Wealth Creators," *Fortune*, November 9, 1998, p. 195. The following articles discuss how companies use the EVA concept: "Stern Stewart EVATM Roundtable," *Journal of Applied Corporate Finance*, Summer 1994, pp. 46–70; S. Milunovich and A. Tsuei, "EVA[®] in the Computer Industry," *Journal of Applied Corporate Finance*, Spring 1996, pp. 104–15; A. Jackson, "The How and Why of EVA[®] at CS First Boston," *Journal of Applied Corporate Finance*, Spring 1996, pp. 98–103.

²¹Some writers in this field have said that accountants take too many liberties. Professor Abraham Briloff has written a number of books and articles on this subject.

have been used. Under present tax law, the Modified Accelerated Cost Recovery System (MACRS) is most frequently employed.

2. There are various ways of recording inventories, the famous FIFO (first-in, first-out) and LIFO (last-in, first-out) being just two alternatives.
3. Amortization of such items as goodwill and patents can be recorded differently.

This is just a small sample of the better-known alternative treatments by accountants, and any of these are in conformance with GAAP. Moreover, the tax return that a company completes and sends to the IRS may be quite different from the published statement of a public company.

As if the question of what accounting profits really are were not enough, the economist compounds this problem even further. Everybody agrees that profit equals revenue minus costs (and expenses). But economists do not agree with accountants on the concept of costs. An accountant reports costs on a historical basis. The economist, however, is concerned with the costs that a business considers in making decisions, that is, future costs. We concern ourselves with this concept more thoroughly later in this book, but we must touch on the subject now, albeit briefly. Basically, economists deal with something they call *opportunity costs* or *alternative costs*. This means that the cost of a resource is what a business must pay for it to attract it into its employ or, put differently, what a business must pay to keep this resource from finding employment elsewhere. To get down to specific examples, we can mention the following:

1. **Historical costs versus replacement costs:** To an economist, the replacement cost of a piece of machinery (and, therefore, the level of periodic depreciation on the replacement cost) is important, whereas an accountant measures cost—and depreciation—on a historical basis.
2. **Implicit costs and normal profits:**
 - a. The owners' time and interest on the capital they contribute are usually counted as profit in a partnership or a single proprietorship. However, the owners could work for someone else instead and invest their funds elsewhere. So these two items are really costs to the business and not profit.
 - b. The preceding item is not relevant in the case of a corporation because even top executives are salaried employees, and interest on corporate debt is deducted as an expense before profits are calculated. However, the payments made to the owners/stockholders—dividends—are not part of cost; they are recorded as a distribution of profits. But surely part of the shareholders' return is similar to the interest on debt because stockholders could have invested their funds elsewhere and required a certain return in order to leave the investment with the corporation. Thus, on this account, corporate profits as recorded by accountants tend to be overstated.

It appears, therefore, that an economist includes costs that would be excluded by an accountant. Indeed, the economist refers to the second category of costs—which are essential to obtain and keep the owners' resources in the business—as **normal profits**, which represent the return that these resources demand to remain committed to a particular firm.

Thus, **economic costs** include not only the historical costs and explicit costs recorded by the accountants, but also the replacement costs and implicit costs (normal profits) that must be earned on the owners' resources. In the rest of this book, profits are considered to be **economic profits**, which are defined as total revenue minus all the economic costs we describe in this section.

GLOBAL APPLICATION

The model of a firm's goals discussed in this chapter applies predominantly to firms operating in the United States and possibly the United Kingdom. However, one must ask whether profit maximization or shareholder wealth maximization is also valid for other countries. It is often said that for many reasons (e.g., political, cultural, legal, and institutional), firms in other countries pursue goals that include the interests of other groups, such as labor, community, government, and so on, in addition to interests of stockholders. In some countries, for instance, labor unions are represented on the board of directors. Thus, it may be necessary to consider such interests in our discussions. However, even if such considerations are important, it is possible for us to treat them as constraints on the actions of a firm. Even if profit or shareholder wealth maximization is not the only objective, as long as firms attempt to take actions that will improve their earnings—within specific constraints—our maximization model can still be used.²² It is important to recognize, however, that multinational firms (e.g., a U.S. parent corporation operating in many different countries through subsidiaries or branches) will encounter restrictions and complications, which they must consider in doing business abroad. We list these and explain them briefly:²³

1. Foreign currencies and their exchange rates must be considered. Thus, revenues, costs, and other cash flows that are denominated in other currencies must be translated into domestic currencies, and their potential changes must be analyzed for their impact on the business. Under certain circumstances, a profitable activity abroad can become unprofitable from the viewpoint of the domestic parent corporation.
2. Legal differences must be taken into account. Dissimilarities in tax laws can have important consequences on results of transactions between the domestic parent corporation and its foreign subsidiary. Differences in legal systems make the tasks of executives considerably more complex.
3. Most Americans have in the past mastered only their own language, and thus, are often at a disadvantage when dealing with their multilingual counterparts in other countries.
4. The differences in cultural environments influence the defining of business goals and attitudes toward risk. Thus, such differences can greatly affect the way business is conducted.
5. The role of government in defining the rules under which companies operate varies from country to country. Although in some countries market competition prevails, in others the political process dictates the behavior of firms in much greater detail.
6. Corporations operating in different countries may be restricted from transferring corporate resources out of the country and may even face the danger of expropriation. This is political risk, which must be included in any economic analysis of a company's prospects.²⁴

The points just discussed, as well as others, must always be considered by companies doing business abroad. Although some of the differences may have adverse effects on

²²It is interesting to note that, in an interview, Heinrich von Pierer, then chief executive officer of the German electronics firm Siemens, stressed "German corporate values as concern with quality, reliability and long-term thinking." He also made the following statement: ". . . if people think that German businesses have an aversion to shareholder value, we are not a very good representative of such companies. . . . To improve profitability and market capitalisation is the main goal." Peter Marsh, "A Conglomerate with an Air of Confidence," *Financial Times*, January 21, 2002.

²³We discuss some of these subjects at greater length in Chapter 13.

²⁴The preceding points can be found in Eugene F. Brigham and Phillip R. Daves, *Intermediate Financial Management*, Thomson Southwestern, 2004, pp. 211–212.

a company, participation in a global market is a necessity for most large (and even small) firms today. Profitability, and even survival, can depend on a company entering global markets and competing worldwide.

The Solution

It was a lively stockholder meeting. Bob Burns was somewhat worried about what the reactions from the audience would be regarding his decision to go into the energy drink market. But he knew that he would have to be convincing because this move was an essential part of the company's growth strategy. After covering the results for the most recent fiscal year, Bob continued:

"In recent years, we've experienced some tough going largely due to the recession brought on by the 2008 financial crisis. Throughout this period, your management has maintained as its primary objective to continue to increase the value of your investment in the company. We are well aware that the price of our stock has not been increasing at the rate it did prior to the recession. However, throughout this period we have remained committed to a long-run increase in the price of our stock. To accomplish this goal, we need to return to the double-digit annual increase in revenue as well as profit.

"As part of this growth strategy, we have decided to enter the growing market for energy drinks. Over the past decade, this has been a rapidly growing segment of the beverage business, particularly among 18- to 34-year-olds. This is a segment that spends a lot for beverages. Energy drinks are popular in the growth markets in Asia. It also has great potential in other growth countries such as Brazil.

"To maintain and increase profits in recent years, we have been focusing mainly on cutting costs through productivity increases and operational excellence. But there is a limit to such endeavors. In the long run, we must find new products and growing markets to increase our top and bottom lines and thus discharge our responsibility to you, our stockholders, to increasing the value of your stock. Entering the market for energy drinks is one of the ways that we believe will bring this about."

SUMMARY

In this text, we generally assume a firm's short-run or long-run objective is the maximization of its profit or the minimization of its loss. Although a firm can select from a number of other goals, both in the short run and in the long run, the assumption of profit maximization provides us with a clear-cut model for explaining how firms can use economic concepts and tools of analysis to make optimal decisions. In presenting these concepts and tools of analysis, a certain amount of mathematics will be employed. Thus, before proceeding to the next chapter, we believe that a brief review of the mathematics used in this text will be helpful. This review is contained in the online appendix.

IMPORTANT CONCEPTS

Business risk. The variability of returns (or profits) due to fluctuations in general economic conditions or conditions specifically affecting the firm. (p. 29)

Economic cost. All cost incurred to attract resources into a company's employ. Such cost includes explicit cost usually recognized on accounting records as well as opportunity cost. (p. 33)

Economic profit. Total revenue minus total economic cost. An amount of profit earned in a particular endeavor above the amount of profit that the

firm could be earning in its next best alternative activity. Also referred to as *abnormal profit* or *above-normal profit*. (p. 33)

Economic Value Added (EVA). The difference between a company's return on total capital and its cost of capital. (p. 32)

Financial risk. The variability of returns (or profits) induced by leverage (the proportion of a company financed by debt). The higher the leverage, the greater the potential fluctuation in stockholder earnings for a given change in total profits. (p. 29)

- Firm.** An organization that transforms resources into products demanded by consumers. The firm chooses to organize resources internally or to obtain them through the market. (p. 19)
- Market Value Added (MVA).** The difference between the market value (equity plus debt) of a company and the amount of capital investors have paid into the company. (p. 31)
- Noneconomic objectives.** A company's objectives that do not appear to be governed by economic thinking but rather define how a business should act. "Acting as a good corporate citizen" is an example of a noneconomic objective. (p. 25)
- Normal profit.** An amount of profit earned in a particular endeavor that is just equal to the profit that could be earned in a firm's next best alternative activity. When a firm earns normal profit, its revenue is just enough to cover both its accounting cost and its opportunity cost. It can also be considered as the return to capital and management necessary to keep resources engaged in a particular activity. (p. 33)

- Opportunistic behavior.** One party to a contract seeks to take advantage of the other. (p. 19)
- Optimal decision.** The decision that enables the firm to meet its desired objective most closely. (p. 23)
- Profit maximization hypothesis.** One of the central themes in economics, the claim that a company will strive to attain the highest economic profit in each period. (p. 23)
- Satisficing.** A concept in economics based on the principle that owners of a firm (especially stockholders in a large corporation) may be content with adequate return and growth since they really cannot judge when profits are maximized. (p. 26)
- Transaction costs.** Cost incurred by a firm in dealing with another firm, including the cost of investigation, negotiation, and enforcement of contracts. (p. 19)
- Wealth maximization.** A company's management of its business in such a manner that the cash flows over time to the company, discounted at an appropriate discount rate, will cause the value of the company's stock to be at a maximum. (p. 30)

QUESTIONS

1. The following is a quote from a *New York Times* article: "If a company makes product donations to the school—computers for instance—then the image of a company goes up as graduate students use the company's products." Does such action square with a company's objective of profit maximization? Discuss.
2. Is the maximization of profit margin (profit as a percentage of total sales) a valid financial objective of a corporation? Discuss.
3. "The growth of consumer information organizations, legal requirements, and warranty requirements has caused significant increases in the cost of customer satisfaction. Thus it is no longer useful to talk about profit maximization as a company objective." Comment on this quote.
4. Discuss the difference between profit maximization and shareholder wealth maximization. Which of these is a more comprehensive statement of a company's economic objectives?
5. Explain the term *satisfice* as it relates to the operations of a large corporation.
6. Discuss the meaning of the term *principal-agent problem*. Why does this problem exist?
7. Why may corporate managers not specifically aim at profit (or wealth) maximization for their companies?
8. What are some of the forces that cause managers to act in the interest of shareholders?
9. Do you believe that profit (or shareholder wealth) maximization still represents the best overall economic objective for today's corporation?
10. Because of inflation, a company must replace one of its (fully depreciated) machines at twice the nominal price paid for a similar machine 8 years ago. Based on present accounting rules, will the company have covered the entire cost of the new machine through depreciation charges? Explain by contrasting accounting and economic costs.
11. How do implicit costs lead to a difference between accounting and economic profits?

12. You have a choice of opening your own business or being employed by someone else in a similar type of business. What are some of the considerations in terms of opportunity costs that you would have to include in arriving at your decision?
13. Various depreciation methods can be used to arrive at an accounting profit number. From the viewpoint of the economist, how should annual depreciation be determined?
14. Do you believe that the profit maximization model can be applied to the activities of a multinational corporation? Explain.
15. What are transaction costs? How does *opportunistic behavior* tend to increase transaction costs?
16. The outsourcing of important parts of a company's production has been growing in recent years. How would you explain these changes? How has the Internet contributed to these changes?
17. What are some reasons for companies internalizing transaction costs?
18. A company has 2 million shares outstanding. It paid a dividend of \$2 during the past year, and expects that dividends will grow at 6 percent annually in the future. Stockholders require a rate of return of 13 percent. What would you expect the price of each share to be today, and what is the value of the company's common stock?
19. Discuss the difference between the calculation of shareholder wealth and the concept of *Market Value Added*. Which of the two would appear to be more meaningful from the viewpoint of a shareholder?

Supply and Demand

Learning Objectives

Upon completion of this chapter, readers should be able to:

- Define supply, demand, and equilibrium price.
- List and provide specific examples of the nonprice determinants of supply and demand.
- Distinguish between the short-run rationing function and the long-run guiding function of price.
- Illustrate how the concepts of supply and demand can be used to analyze market conditions in which management decisions about price and allocations of resources must be made.
- Use supply and demand diagrams to show how the determinants of supply and demand interact to determine market price in the short run and in the long run.

The Situation

While sipping a cup of green tea with honey and ginseng, CEO Bob Burns began to reminisce. It seemed like yesterday, but it was actually more than 10 years ago when he had convinced the board of directors of Global Foods, Inc., to go into the soft drink business. Here he was a decade later, sampling a product that his VP of Marketing, Nicole Goodman, was telling him would be an even stronger “growth engine” for the company than bottled water. She had pointed out to him that in 2003 American consumers spent \$5 billion on tea. Although this amount was far less than the \$20 billion spent annually for coffee, it was five times as much as people spent on tea a decade ago.¹ “It’s obvious,” she told him, “Global Foods must get into the tea business.” As he poured his second cup of tea, Bob had to admit that even he had begun to prefer tea over coffee. He decided to call Nicole into his office to discuss the matter further.

“Okay, Nicole,” Bob began, “You’ve always had a good instinct for what’s new in the market. But before we leap into this, I want a report on exactly why you believe tea will be the real spark to our company’s growth in the coming 5 years. After all, in our business it’s all about ‘share of stomach.’ If people are drinking more tea, then they might be drinking fewer soft drinks and bottled water, so we’d be cannibalizing our own products. I’d feel much better if you could help me understand why this wouldn’t be the case. Furthermore, what are the key determinants of the demand for tea? Could this be just a fad? Already people are starting to tire of their low-carb diets. Nicole, I want you to provide me with a report on the elements that are driving the demand for tea. How responsive will people be to price changes, to changes in the price of competing products such as bottled water and carbonated soft drinks? Is tea a ‘luxury’ good, or is it a necessity? The answers to these questions will help us better understand how to price and position our brand in the marketplace.”

¹Much of the situation and solution is based on Elizabeth Olsen, “A Tea Party, and All Boomers Are Invited,” *New York Times*, October 17, 2004.

INTRODUCTION

In this chapter, we introduce the basic elements of supply and demand. Although for some of you, this chapter serves as a review of material covered in an economics principles course, it has been included because it is essential for every reader to have a thorough grounding in supply and demand before proceeding to the particulars of managerial economics. There may be situations—such as those described in this chapter’s “The Situation” vignette—in which you may be required to conduct or evaluate a study with a considerable use of supply-and-demand analysis. Regardless of how directly this chapter’s material may apply to your work, most of the material covered in this book relates in some way to supply or demand. Indeed, supply and demand can be considered the conceptual framework within which the specifics of managerial economics are discussed.

MARKET DEMAND

The **demand** for a good or service is defined as

Quantities of a good or service that people are ready to buy at various prices within some given time period, other factors besides price held constant.

Note that in this definition “ready” implies that consumers are prepared to buy a good or service both because they are willing (i.e., they have a preference for it) and they are able (i.e., they have the income to support this preference).

Demand can first be illustrated with an example in which we imagine that you, the reader, are part of a simple market experiment. Suppose you were asked to respond to the following survey question: “In a 1-week period, how many slices of pizza would you be prepared to buy at the following prices: \$2.00, \$1.50, \$1.00, \$.50, and \$.05?” Every reader would obviously have his or her own pattern of response. Let us assume a sample of three readers respond in the following way:

Price (per Slice)	Q_{D1}	Q_{D2}	Q_{D3}	Q_{DM}
\$2.00	0	2	3	5
1.50	1	2	5	8
1.00	2	2	8	12
0.50	3	3	10	16
0.05	4	4	12	20

As you can see, the combined responses of the three individuals make up the total **market demand** (Q_{DM}) for pizza, the sum of all individual demands.

Market demand is illustrated with a simple numerical function, as shown in Table 3.1. This table shows a hypothetical demand for pizza. As the price of a slice of pizza falls from \$7.00 to zero, the amount that consumers in this market are willing to buy increases from zero to 700 slices. This inverse relationship between price and the **quantity demanded** of pizza is called the **law of demand**. There may be instances in which consumers behave in an “irrational” manner by buying more as the price rises and less as the price falls because they associate price with quality. But in the economic analysis of demand, it is assumed buyers do not associate price with quality and will therefore follow the law of demand.

The law of demand can be observed in the curve shown in Figure 3.1, derived from the schedule of numbers in Table 3.1. Notice that the curve in this figure slopes

Table 3.1 Market Demand for Pizza

Price (per Slice)	Q_D
\$7.00	0
6.00	100
5.00	200
4.00	300
3.00	400
2.00	500
1.00	600
0	700

downward and to the right, indicating that the quantity of pizza demanded increases as the price falls, and vice versa.

A change in the demand for pizza or any other product is indicated by a change in the entire schedule of quantities demanded at a list of prices, or a shift in the demand curve either to the left or to the right. We see these changes in Table 3.2 and Figure 3.2.

To summarize, we can say the following:

Changes in price result in **changes in the quantity demanded** (i.e., movements *along* the demand curve).

Changes in the nonprice determinants result in **changes in demand** (i.e., *shifts* in the demand curve).

This difference can be seen in Table 3.2 in the following manner. At the price of \$5 the quantity demanded in the first list of responses (Q_{D1}) is 200. If the price drops to \$4, then the *quantity demanded* increases to 300. However, if the *demand* increases to Q_{D2} , then at the price of \$5 the quantity increases to 300 and in fact increases by 100 units at each price being offered.

Figure 3.1 Market Demand Curve for Pizza

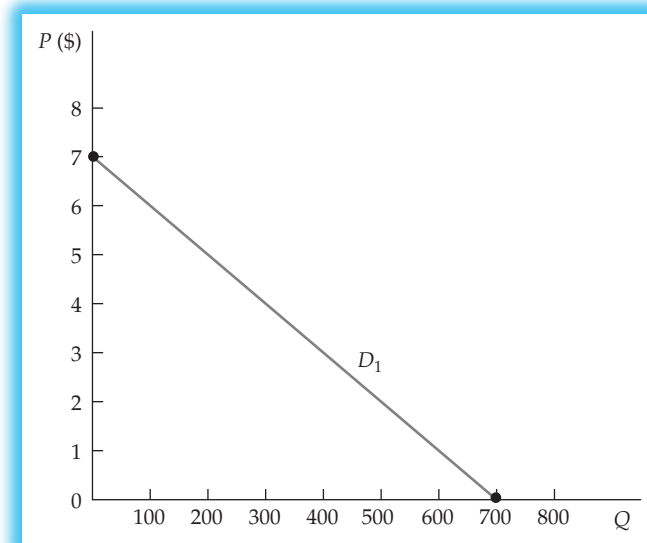
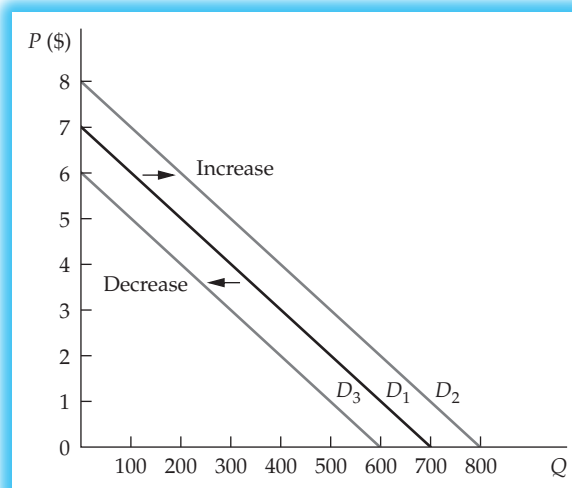


Table 3.2 Different Levels of Market Demand for Pizza

Price (per Slice)	Q_{D1}	Q_{D2}	Q_{D3}
\$7.00	0	100	0
6.00	100	200	0
5.00	200	300	100
4.00	300	400	200
3.00	400	500	300
2.00	500	600	400
1.00	600	700	500
0	700	800	600

Factors that can cause demand to change are called **nonprice determinants of demand**. Following is a list of these determinants and a brief elaboration of their impact on demand:

1. **Tastes and preferences.** Why do people buy things? Marketing professors, corporate market researchers, and advertising executives spend their careers trying to answer this question. Economists use a general-purpose category in their list of nonprice determinants called *tastes and preferences* to account for the personal likes and dislikes of consumers for various goods and services. These tastes and preferences may themselves be affected by other factors. Advertising, promotions, and even government reports can have profound effects on demand via their impacts on people's tastes and preferences for a particular good or service.
2. **Income.** As people's incomes rise, it is reasonable to expect their demand for a product to increase, and vice versa. In the next chapter, the possibility of demand moving in the *opposite direction* to changes in income is discussed.
3. **Prices of related products.** A good or service can be related to another by being a substitute or by being a complement. If the price of a substitute product changes, we expect the

Figure 3.2 Shifts in the Market Demand for Pizza

demand for the good under consideration to change in the *same direction* as the change in the substitute's price. Consider, for example, what would happen to the demand for software if the price of computer hardware falls, or to the demand for music downloads (from legitimate sites!) if the price of MP3 players falls. It is reasonable to expect that the demand for the two items would *rise* as a result of a *fall* in the price of their respective *complementary* products.

4. **Future expectations.** If enough buyers expect the price of a good or service to rise (fall) in the future, it may cause the current demand to increase (decrease). In markets for various financial instruments (e.g., stocks, bonds, negotiable certificates of deposit, U.S. Treasury bills), as well as for agricultural commodities and precious metals, expectations of future price changes among both buyers and sellers play an important part in determining the market demand. In most of these types of markets, speculation among buyers and sellers is an important factor to consider. Buyers and sellers act on a current price of a product not for its immediate consumption but because of the possibility of gaining from some future transaction. (Recall the old adage “buy low and sell high.”)

This factor can also affect the demand for consumer and commercial products. For example, the demand for DVD recorders, digital cameras, home entertainment systems, laptop computers, and personal digital assistants was probably not as high as sellers expected when these products were first introduced because buyers were waiting for their prices to come down at a later time.

5. **Number of buyers.** The impact of the number of buyers on demand should be apparent; as far as sellers are concerned, the more the merrier. What is interesting, nonetheless, is how changing demographics and tastes and preferences within demographic groups can affect the pool of potential buyers for a particular good or service. In other words, sheer numbers (i.e., population) may not be as important as differences within the population.

We discuss further how changes in these factors change demand and market price. But first we introduce the concept of supply. By combining supply with demand, we can conduct a complete analysis of the market, both in the short run and in the long run.

MARKET SUPPLY

The **supply** of a good or service is defined as

Quantities of a good or service that people are ready to sell at various prices within some given time period, other factors besides price held constant.

Notice that the only difference between this definition and that of demand is that in this case the word *sell* is used instead of *buy*. Just as in the case of demand, supply is based on an assumed length of time within which price and the other factors can affect the **quantity supplied**.

Recall that the law of demand states that the quantity demanded is related inversely to price, other factors held constant. In contrast, the law of supply states that quantity supplied is related *directly* to price, other factors held constant. Thus, any schedule of numbers representing a relationship between price and quantity supplied would show a *decrease* in the quantity supplied as price falls.

Table 3.3 shows a hypothetical supply schedule. Also shown are two additional supply schedules, one indicating a greater supply and the other showing a reduced supply. These schedules are shown as supply curves in Figure 3.3. The supply curve has a positive slope, reflecting the direct relationship between price and quantity supplied.

Table 3.3 Market Supply for Pizza

P	Q_{S1}	Q_{S2}	Q_{S3}
\$7	600	700	500
6	500	600	400
5	400	500	300
4	300	400	200
3	200	300	100
2	100	200	0
1	0	100	0
0	0	0	0

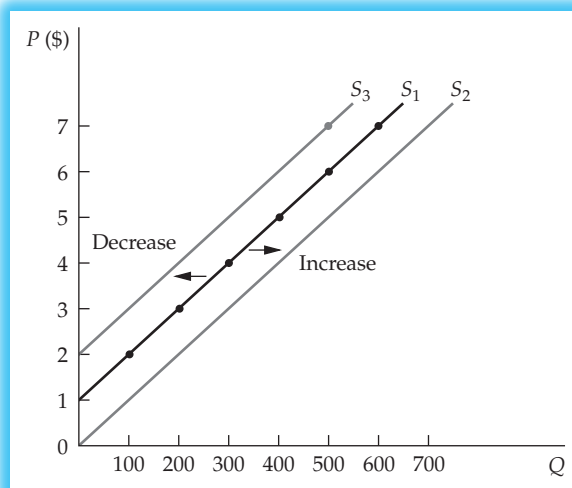
In analyzing the supply side of the market, it is important to make the distinction between *quantity supplied* and *supply*. The distinction between these two terms is the same as that used for the demand side of the market:

Changes in *price* result in **changes in the quantity supplied** (i.e., movements along the supply curve).

Changes in *nonprice determinants* result in **changes in the supply** (i.e., shifts of the supply curve).

Just as there are nonprice determinants of demand, there are **nonprice determinants of supply**. A change in any one or a combination of these factors will change market supply (i.e., cause the supply line to shift to the right or the left). We briefly discuss each factor to understand why this is expected to occur:

1. **Costs and technology.** The two factors of costs and technology can be treated as one because they are so closely related. *Costs* refer to the usual costs of production, such as labor costs, costs of materials, rent, interest payments, depreciation charges, and general and administrative expenses—in other words, all items usually found in a firm's income statement. *Technology* refers to technological innovations or improvements introduced to reduce the unit cost of production (e.g., automation, robotics, and computer hardware

**Figure 3.3** Supply Curves for Pizza

and software utilization). Technological changes that result in entirely new products for final consumption are not considered part of this category. These new products would have to be considered in an entirely different market analysis. In any event, unit cost reductions, whether from technological innovations or simply management decisions, will result in an increase in market supply. Increases in the unit cost of production will have the opposite effect.

2. **Prices of other goods or services offered by the seller.** From the consumer's standpoint, any good or service has other goods or services related to it either as substitutes or as complements. From the producer's standpoint, there can also be substitutes or complements for a particular good or service offered in the market. For example, suppose the sellers of pizza notice that the price of hot dogs increases substantially. In the extreme case, they may drop their line of pizza and substitute hot dogs. Or they may at least reduce the amount of resources (e.g., labor and store space) devoted to the selling of pizza in favor of hot dogs. In either case, the market supply of pizza would decrease. If the sellers were already selling two (or more) products, the change in market conditions would prompt them to reallocate their resources toward the more profitable products. (Given this possibility, it may be more appropriate to say that the sellers consider pizza and hot dogs as "competing" products rather than as "substitute" products.)
3. **Future expectations.** This factor has a similar impact on sellers as on buyers; the only difference is the direction of the change. For example, if sellers anticipate a rise in price, they may choose to hold back the current supply to take advantage of the higher future price, thus decreasing market supply. As we discuss in the "Market Demand" section, an expected rise in price will increase the current demand for a product.
4. **Number of sellers.** Clearly, the number of sellers has a direct impact on supply. The more sellers, the greater the market supply.
5. **Weather conditions.** Bad weather (e.g., floods, droughts, unusual seasonal temperatures) will reduce the supply of an agricultural commodity. Good weather will have the opposite effect.

With this discussion of supply, we are now able to combine supply with demand into a complete analysis of the market.

MARKET EQUILIBRIUM

Now that we have reviewed the definitions and mechanics of demand and supply, we are ready to examine their interaction in the market. Market demand and supply are compared in Table 3.4 and Figure 3.4.

You can see in both the table and the graph that at the price of \$4, the market is cleared in the sense that the quantity demanded (300) is equal to the quantity

Table 3.4 Supply and Demand for Pizza

P	Q_D	Q_S
\$7	0	600
6	100	500
5	200	400
→4	300	300
3	400	200
2	500	100
1	600	0
0	700	0

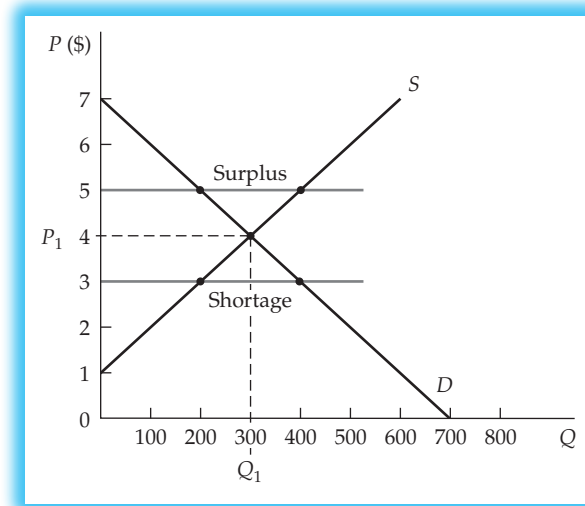


Figure 3.4 Supply and Demand Curves for Pizza, Indicating Market Equilibrium

supplied (300). Thus, \$4 is called the **equilibrium price**, and 300 is referred to as the **equilibrium quantity**. Another way to view this market situation is to imagine what would happen if the price were not at the equilibrium level. For example, suppose the price were at a higher level, say, \$5. At this price, as you can see in Table 3.4, the quantity supplied would exceed the quantity demanded, a condition called a **surplus**. At a lower price, say, \$3, the situation is reversed: the quantity demanded exceeds the quantity supplied. This situation is called a **shortage**. Both the surplus and the shortage conditions are indicated in Figure 3.4.

In the event of a surplus or a shortage, various competitive pressures cause the price to change (decrease in the case of a surplus, and increase in the event of a shortage). The price thus serves to clear the market of the imbalance. The clearing process continues until equilibrium (i.e., quantity demanded equals quantity supplied) is achieved. In the case of a surplus, sellers wanting to rid themselves of the extra items offer the product at a lower price to induce people to buy more. At the same time, as the price falls, suppliers are discouraged from offering as much as before. In the case of a market shortage, as the price rises toward the equilibrium level, the market is cleared because the quantity demanded decreases while the quantity supplied increases. In the event of a shortage, sellers try to take advantage of the situation by raising their prices, and people are thus discouraged from buying as much as before. Also, sellers are induced to offer a greater number of items in the market. Both actions serve to clear the market of a shortage.

To summarize the material in this section, remember the following definitions:

Equilibrium price: The price that equates the quantity demanded with the quantity supplied (i.e., the price that clears the market of a surplus or shortage).

Equilibrium quantity: The amount that people are willing to buy and sellers are willing to offer at the equilibrium price level.

Shortage: A market situation in which the quantity demanded exceeds the quantity supplied, *at a price below the equilibrium level*.

Surplus: A market situation in which the quantity supplied exceeds the quantity demanded, *at a price above the equilibrium level*.

COMPARATIVE STATICS ANALYSIS

The model of market demand, supply, and equilibrium price and quantity developed in the preceding sections can now be used to analyze the market. The particular method of analysis we use is called *comparative statics analysis*. This is a commonly used method in economic analysis and is used throughout the text. In general, this method of analysis proceeds as follows:

1. State the assumptions needed to construct the model.
2. Begin by assuming the model is in equilibrium.
3. Introduce a change in the model. In so doing, a condition of disequilibrium is created.
4. Find the new point at which equilibrium is restored.
5. Compare the new equilibrium point with the original one.

In effect, comparative statics analysis is a form of sensitivity analysis, or what business people often refer to as *what-if* analysis. For example, if we were doing a what-if analysis of a company's cash flow, we would start with a given pro forma income statement adjusted to provide the cash flow for a given period of time. We would then conduct sensitivity analysis by supposing certain factors changed, such as revenue, cost, or the rate of depreciation. We would then inspect how changes in these factors would change the cash flow of the firm over time. In the same manner, economists conduct a what-if analysis of their models.

The term *statics* alludes to the theoretically stable point of equilibrium, and *comparative* refers to the comparison of the various points of equilibrium. The ensuing sections explain exactly how comparative statics analysis is used in the analysis of the market.



Module 3A

Short-Run Market Changes: The “Rationing Function” of Price

Let us continue with our analysis of pizza. Following the steps involved in comparative statics analysis, we start by assuming all factors except the price of pizza are held constant, and the various patterns of response to price among buyers and sellers are represented by the supply and demand lines in Figure 3.4. We make a fresh start by redrawing this graph in Figure 3.5. It would also be useful to recall all nonprice determinants that could affect the demand or supply for a product. They are listed for you in Table 3.5.

As noted in step 2 in the previous section, we begin this analysis in the condition of equilibrium. This is denoted in Figure 3.5 as the point where the supply line intersects with the D_1 demand line (i.e., the price level where quantity supplied is equal to quantity demanded).

Based on step 3, we introduce a change in one or more of the assumptions made when the model was constructed. Any one or more of the factors shown in Table 3.5 can cause this change. Let us assume a new government study shows pizza to be the most nutritious of all fast foods and that consumers substantially increase their demand for pizza as a result of this study. In Figure 3.5, this increase is represented by a shift in the demand curve from D_1 to D_2 . As you can see, this shift results in a new, higher equilibrium price of \$4.50. Notice also that the new equilibrium quantity is greater than the original equilibrium quantity.

The comparison of the new equilibrium point with the original one (step 5 in comparative statics analysis) leads us to conclude that, as a result of a change in tastes and preferences, the price of pizza rises, and so does the quantity bought and sold.

This analysis can be repeated using other possible changes in market conditions (e.g., the price of cheese rises, the price of soft drinks falls). Each time, the same

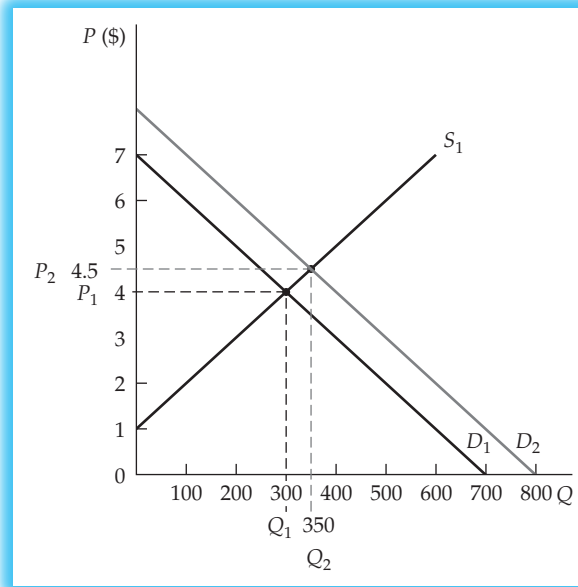


Figure 3.5 Increase in Demand for Pizza and Resulting Impact on Market Equilibrium

procedure should be followed. If we consider only one possible change at a time, the effects on equilibrium price and quantity can be illustrated as in Figure 3.6. Instead of using specific numbers, we have designated the prices and quantities with the symbols P and Q along with appropriate subscripts. We can summarize the effects shown in the graphs as follows:

- An increase in demand causes equilibrium price and quantity to rise. (See Figure 3.6*a*.)
- A decrease in demand causes equilibrium price and quantity to fall. (See Figure 3.6*b*.)
- An increase in supply causes equilibrium price to fall and quantity to rise. (See Figure 3.6*c*.)
- A decrease in supply causes equilibrium price to rise and quantity to fall. (See Figure 3.6*d*.)

In Figure 3.6, we observe that the shift in demand or supply has in effect created either a shortage or a surplus at the original price P_1 . Thus, the equilibrium price has to rise or fall to clear the market. When the market price changes to eliminate the

Table 3.5 Nonprice Determinants of Demand and Supply

Demand	Supply
1. Tastes and preferences	1. Costs and technology
2. Income	2. Prices of other products offered
3. Prices of related products	3. Future expectations among sellers
4. Future expectations among buyers	4. Number of sellers
5. Number of buyers	5. Weather conditions (particularly for agricultural products)

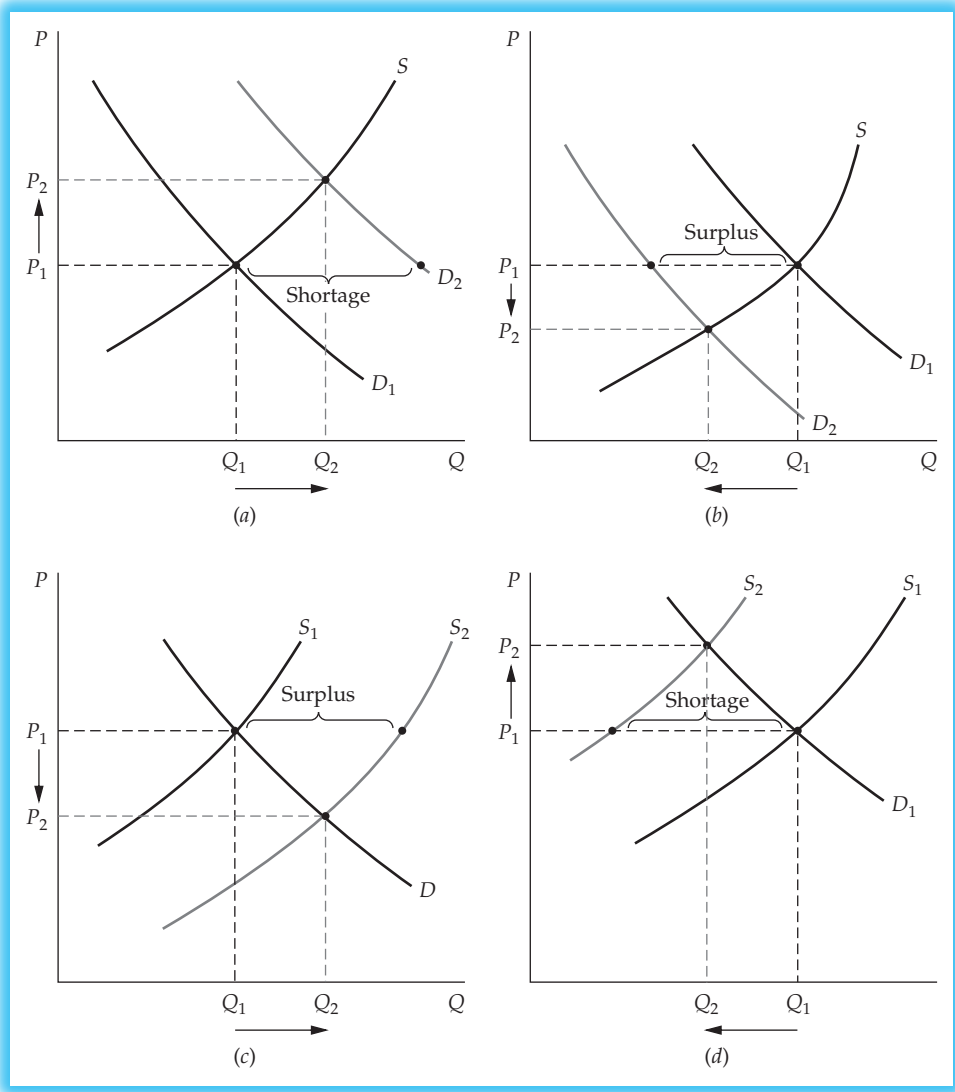


Figure 3.6 Changes in Supply and Demand and Their Short-Run Impact on Market Equilibrium (the Rationing Function of Price)

imbalance between quantities supplied and demanded, it is serving what economists call the **rationing function of price**. The term *rationing* is often associated with shortages, but we define it to also include a surplus situation.

Long-Run Market Analysis: The “Guiding” or “Allocating Function” of Price

The comparative statics analysis presented earlier required only that you consider the response of equilibrium price and quantity to a given change in supply or demand. This response was dubbed the “rationing function” of price. Let us consider what might happen as a result of this change in market price. To illustrate this, we examine the market for hot dogs, a presumed substitute for pizza. The two markets are represented by the supply and demand diagrams in Figure 3.7.

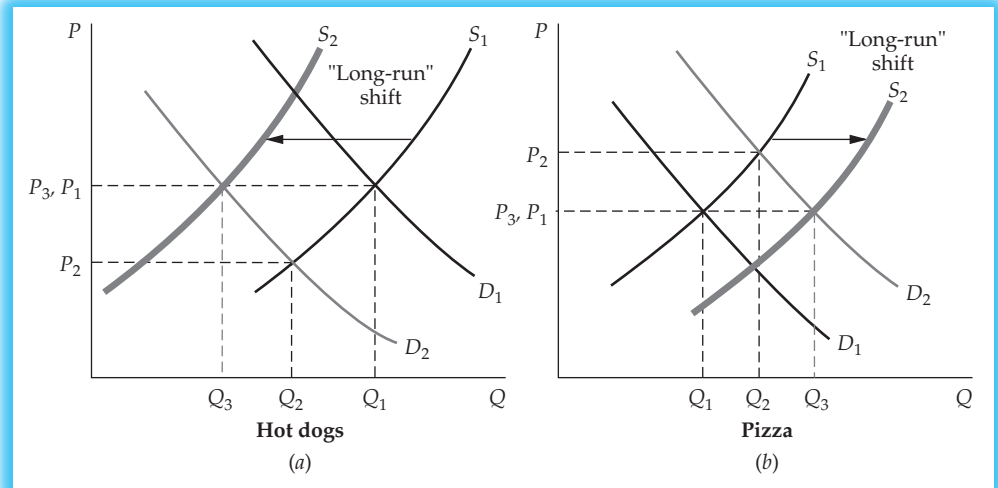


Figure 3.7 Short-Run and Long-Run Changes in Supply (in Response to an Initial Change in Demand)

Now let us assume that at the same time people's tastes and preferences change in favor of pizza, their tastes and preferences become more adverse to hot dogs (e.g., for health reasons). The changes in the demand for the two products are shown in Figure 3.7 by a downward shift in the demand for hot dogs and an upward shift in the demand for pizza (D_1 to D_2). This would cause a shortage in the pizza market and a surplus in the hot dog market. As we know, the rationing function of price will immediately start to correct these market imbalances. As the price of hot dogs falls, the surplus is eliminated; as the price of pizza rises, the shortage is eliminated. (For the purpose of the analysis, it really does not matter where the price of pizza stands in relation to the price of hot dogs. To simplify matters, we have assumed the two prices were about equal before the change in tastes and preferences occurred. The point is that after price performs its rationing function, the equilibrium price of pizza will be higher than the equilibrium price of hot dogs in relative terms.)

Now suppose the prices have indeed changed, and the two markets are once again in equilibrium. What do you suppose will happen next? As you might imagine, the depressed price of hot dogs will cause the sellers to begin allocating less of their resources to this market. Some may even go out of the business of making or selling hot dogs. In contrast, the higher price of pizza will induce the allocation of more resources into this market. New pizza stands and restaurants may be opened. Food companies may build new plants to produce frozen pizza for distribution through supermarkets. The effect of these follow-on adjustments to the initial change in equilibrium prices can be seen in the figure as a rightward shift in the supply of pizza and a leftward shift in the supply of hot dogs.

After this "long-run" adjustment is made, equilibrium price and quantity may return to the levels at which they were before the initial changes in demand took place (i.e., P_3 in each market may be close to or equal to P_1). But the main point is that Q_3 is considerably less than Q_1 in the hot dog market and considerably more than Q_1 in the pizza market. These differences represent the shifting of resources out of the hot dog market and into the pizza market. Several centuries ago, Adam Smith referred to this shifting of resources into and out of markets in response to price changes as

the “invisible hand.”² Another way to express these shifts in supply is that they represent a response to “price signals” sent to the owners of the factors of production. In any event, when resources have been shifted out of the market for hot dogs and into the market for pizza, price is fulfilling its **guiding or allocating function**. Defined in a more formal manner, the guiding or allocating function of price is the movement of resources into or out of markets in response to a change in the equilibrium price of a good or service.

The preceding example illustrates a basic distinction made in economic analysis between the “short run” and the “long run.” This distinction has nothing to do directly with a specific calendar time. Instead, it refers to the amount of time it takes for sellers and buyers to react to changes in the market equilibrium price. The following descriptions of the short run and the long run helps readers distinguish the two time periods:

1. Short run

- a. Period of time in which sellers already in the market respond to a change in equilibrium price by adjusting the amount of certain resources, which economists call *variable inputs*. Examples of such inputs are labor hours and raw materials. A short-run adjustment by sellers can be envisioned as a movement along a particular supply curve.
- b. Period of time in which buyers already in the market respond to changes in equilibrium price by adjusting the quantity demanded for a particular good or service. A short-run adjustment by buyers can be envisioned as a movement along a particular demand curve.

2. Long run

- a. Period of time in which new sellers may enter a market or the original sellers may exit from a market. This period is long enough for existing sellers to either increase or decrease their *fixed factors* of production. Examples of fixed factors include property, plant, and equipment. A long-run adjustment by sellers can be seen graphically as a shift in a given supply curve.
- b. Period of time in which buyers may react to a change in equilibrium price by changing their tastes and preferences or buying patterns. (The *Wall Street Journal* and other sources of business news may refer to this as a “structural change” in demand.) A long-run adjustment by buyers can be seen graphically as a shift in a given demand curve.

Another good way of distinguishing the short run from the long run is to note that the rationing function of price is a short-run phenomenon, whereas the guiding function is a long-run phenomenon.

Let us summarize the short-run “rationing function” and the long-run “guiding function” of price in terms of our example involving pizza and hot dogs:

1. Changing tastes and preferences cause the demand for pizza to increase and the demand for hot dogs to decrease.
2. The changing demand for the two products causes a shortage in the pizza market and a surplus in the hot dog market.
3. In response to the surplus and shortage in the two markets, price serves as a *rationing agent* by decreasing in the hot dog market and increasing in the pizza market. That is, the short-run response by suppliers of the two products is to change their variable inputs (i.e., movement downward along the supply line in the market for hot dogs, and movement upward along the supply line in the market for pizza).
4. In the *long run*, price fulfills its *guiding* function by causing sellers and potential sellers to respond by increasing capacity or entering the market for pizza and by decreasing

²For Smith, the “visible” hand was that of the government, which might try to dictate the allocation of resources among different markets by the command process rather than by the market process.

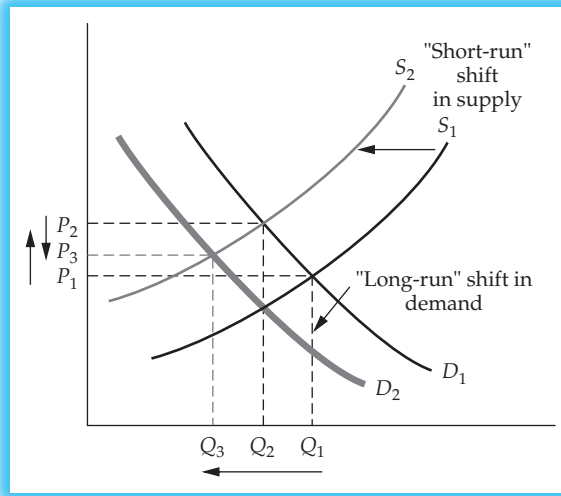


Figure 3.8 Short-Run and Long-Run Changes in Demand (in Response to an Initial Change in Supply)

capacity or leaving the market for hot dogs (i.e., rightward shift in the supply line for pizza and leftward shift in the supply line for hot dogs).

5. As a result of the shifts in supply, new equilibrium levels of price and quantity are established. The new quantities bought and sold represent shifts in resources out of one market and into the other.

The distinction between short- and long-run changes in the market can also be made in cases that begin with changes in supply rather than in demand. A classic example is the case of the Organization of Petroleum Exporting Countries (OPEC) and the world oil market. In the early 1970s, OPEC conspired to raise the price of oil by limiting production to an amount that would support a price above the current level. The supply-and-demand diagram shown in Figure 3.8 illustrates this action. Table 3.6 provides a brief description of this illustration. As we can see, limiting the production of oil can be envisioned as a leftward shift in the supply line to the level where it intersects the demand curve for oil at some designated point above the current market price (i.e., P_2 rather than P_1). The short-run response by consumers to the increase in oil prices was to reduce their consumption of oil. However, in the terms specific to our analysis, this reduction can be seen as a decrease in the *quantity demanded* for oil. In other words, the decrease in the supply of oil (i.e., the shift of the supply line to the left) prompted a *movement back along the demand curve* for oil.

Over time, U.S. consumers began to change their pattern of oil consumption as a result of the high price. They formed car pools, bought more fuel-efficient cars, lowered their thermostats in their homes, and even tried to follow the new 55-mph speed limit established on highways throughout the country. Industrial users of oil responded by substituting more fuel-efficient machinery as soon as it became cost efficient to make these changes. The effect of this *long-run* change in the pattern of oil usage caused the demand for oil to gradually fall. Graphically, this is presented by a leftward shift of the demand curve for oil from D_1 to D_2 . Notice that as a result of this long-run shift in demand, the equilibrium price and quantity fell. As seen in Figure 3.8, the long-run quantity that is bought and sold (i.e., Q_3) was now even less than it was before this decrease in demand took place. This indicates a further shift of resources out of this market.

Table 3.6 Short-Run and Long-Run Changes in the Market

INITIAL CHANGE	FOLLOW-ON CHANGE
(Short-Run Time Period)	(Long-Run Time Period)
Increase in demand causes <i>price to rise</i>	<i>Supply increases</i> as new sellers enter the market and original sellers increase production capacity
Decrease in demand causes <i>price to fall</i>	<i>Supply decreases</i> as less profitable firms or those experiencing losses exit the market or decrease production capacity
Increase in supply causes <i>price to fall</i>	<i>Demand increases</i> as tastes and preferences of consumers eventually change in favor of the product relative to substitutes
Decrease in supply causes <i>price to rise</i>	<i>Demand decreases</i> as tastes and preferences of consumers eventually change away from the product and toward the substitutes

The previous events occurred more than 35 years ago. After a series of ups and downs associated with such world events as the U.S. problem with Iran in the late 1970s and the 1990 Gulf War, the price of oil settled at around \$20 per barrel throughout the second half of the 1990s. Sometimes, the average price even fell into the teens. In mid-2003, the price of oil began to climb to about \$50 per barrel. At first, American consumers appeared to continue with “business as usual.” An article in the *New York Times* at that time quoted an oncologist from Wisconsin as saying, “I don’t like gas being this expensive. My driving is integral to my job. It’s integral to my pleasure in life. I won’t cut back on that. I might cut back on other things.” (He drove an SUV that got 18 miles to the gallon.) Another SUV owner, a nurse from Massachusetts, was also quoted. “I have to have a car,” she explained. “I have a little girl I have to drive everywhere. In the suburbs you have to drive.”³

Throughout the rest of the last decade, oil prices ranged from about \$30 to \$90 a barrel. In 2011, the average price was about \$87.⁴ At the time this textbook was written, the price of oil had climbed to over \$100 a barrel. Among the factors cited for this increase are the effect of speculation; uncertainty in the Middle East; and the growing demand for oil by China, India, and other growth economies. (See this chapter’s section on “Global Applications.”)

Using Supply and Demand in Forecasting

Chapter 5 is devoted to the subject of demand estimation and forecasting. But we should point out here that a fundamental part of economic forecasting is to understand the nature of the determinants of supply and demand. Forecasting supply is somewhat easier than forecasting demand, particularly if factors such as natural disasters (e.g., droughts or floods in the market for food products), international conflicts, or political crises (e.g., turmoil in the Middle East) are not involved. For example, in the market for manufactured goods, one essentially needs to assess the current and future capacity of the producers. This could be done by counting the number of

³“Laissez-Faire My Gas Guzzler, Already: Never Mind the Price, Just Fill ‘er Up,” *New York Times*, September 7, 2004.

⁴http://inflationdata.com/inflation/inflation_rate/historical_oil_prices_table.asp

factories in operation and their utilization rates, or by looking at the overall level of inventories in the industry.

The more difficult challenge is to estimate demand. As you see in Chapter 5, economists try to understand and forecast demand by using the statistical analysis of data collected either from historical records or from a cross section of economic entities (e.g., consumers, households, companies, states, regions, or countries). Much effort is made to understand the quantitative relationship between changes in the determinants of demand and the amount of demand of a particular good or service.

Although companies work very hard to try to quantify their understanding of the demand for their products, it is important to at least have a good understanding of the forces that affect this demand. For example:

1. Makers of PCs and notebooks are facing slower growth in the demand for their products as the demand for tablet PCs and smart phones (e.g., substitute products) continue to increase.
2. Makers of jump drives are starting to experience the same decrease in demand (as are the makers of CDs and DVDs) as more digital storage and transfer of data is done through the Internet.

Using the above examples as a guide, what impact do you think cloud computing will have on the demand for stand-alone applications such as Microsoft Office or storage devices for computers?

SUPPLY, DEMAND, AND PRICE: THE MANAGERIAL CHALLENGE

A critical factor that managers must consider when making decisions such as the pricing of products and the allocation of a company's scarce resource is the market environment in which their company is competing. This chapter focuses on the mechanics of supply and demand in a highly competitive market. In the extreme case, the forces of supply and demand are the sole determinants of market price. This type of market is called "perfect competition." (For complete details about this market, see Chapter 8.) Managers operating in perfectly competitive markets are simply "price takers" trying to earn a profit by making decisions about the allocation of resources based on their short- and long-run assessment of the movements of supply, demand, and prices.

There are other types of competitive markets in which firms act as "price makers" by exercising varying degrees of control over the price of their product. We refer to this type of market as "imperfectly competitive" and the control of market price as **market power**. This power to strongly influence market price may stem from these firms' ability to differentiate their product through advertising, brand name, or special features or add-on services. Also, many oligopolistic firms hold extremely large shares of the market, and their sheer size enables them to dictate prices. (Microsoft and its products for the PC immediately come to mind as an example.) Nonetheless, supply and demand do establish the overall framework in which prices are set. For example, regardless of how strong the market power of a firm, it would be extremely difficult for it to raise prices in the face of falling or sluggish market demand.

GLOBAL APPLICATION: THE BRIC COUNTRIES AND THE SUPPLY AND DEMAND FOR OIL

By now, you must have seen many photos or YouTube videos of traffic jams in the major cities of China. Perhaps you have personally experienced a Beijing traffic jam. If you have, then you will not need any statistics to tell you that, along with the amazing growth of the BRIC countries over the past decade has come their increasing

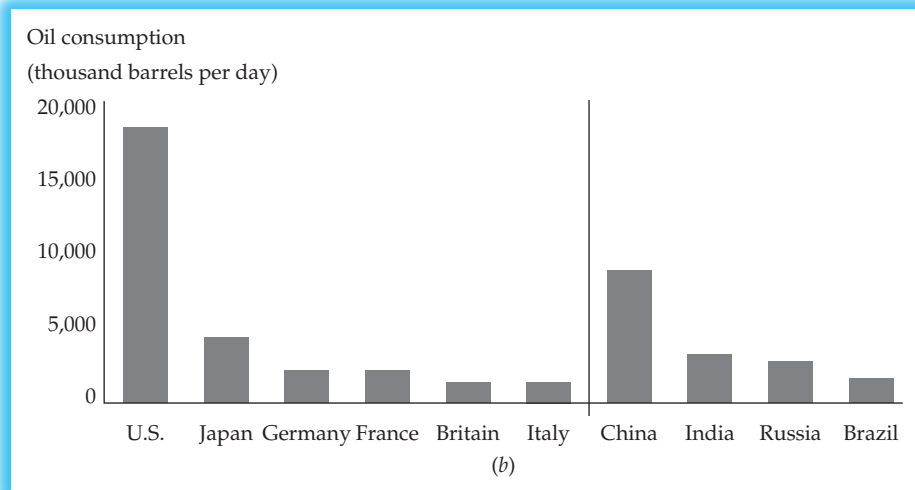
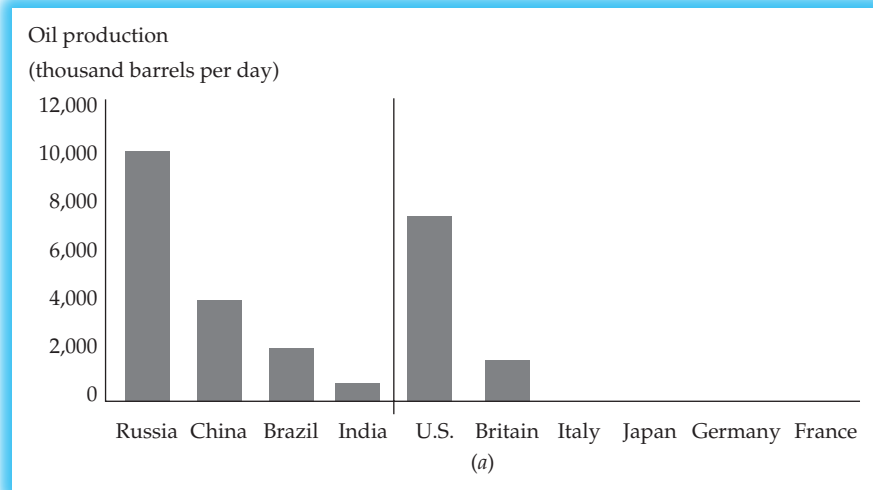


Figure 3.9 Production and Consumption of Oil in the BRICs and Other Countries

Source: Charts © 2011 Deloitte Global Services Limited (“DGS”), used with permission. Charts incorporate preexisting information from BP Statistical Review of World Energy, June 2010 edition (based on 2009 year-end data). DGS has no connection or affiliation with BP or the BP Statistical Review of World Energy publication. Please see the copyright page for the full attribution and disclaimer notices pertaining to these charts.

appetite for oil as well as for coal and natural gas. The graphs in Figure 3.9 give you a general idea of the current supply and demand conditions for oil. Notice that one of the BRICs, Russia, is the leading producer of oil among the BRIC and G6 countries. Of course, the United States is the leading consumer. But notice how China has reached the point where its daily oil consumption is about half that of the United States. Indeed, the growing demand for gasoline by Chinese drivers was cited as one of the reasons for the marked increase in gasoline prices in the United States in the first half of 2012. More important, oil consumption by China and India is expected to quadruple by 2030.⁵

⁵<http://the-diplomat.com/china-power/2011/11/07/china-brics-and-the-environment/>

The Solution

Rather than present her findings in a formal report using overhead slides, Nicole believed she would be more effective by simply providing her boss with an informal briefing. In fact, Nicole noticed a backlash among senior executives regarding the excessive use of huge “decks” of slides. They actually seemed to detract from the presenter’s message. She knew she could send Bob all the details in a complete deck after she made her key points to him in this informal discussion.

“Bob, my staff has carefully gone over all available information about the potential market for tea. Right now the supply of the standard tea in a bag that the majority of Americans continue to drink is provided by three dominant players: Lipton (owned by our archrival, Unilever), Bigelow (with its leading brand, Constant Comment), and Twinings (with its British teatime image). But the real growth of this market is in the hundreds of varieties of green teas, exotic black teas, and even white and red teas. And right now, the supply for these varieties is being met by hundreds of specialty companies, many of whom distribute their product online and through retail stores. More important, these suppliers of ‘designer teas’ are able to charge a higher price than the standard teas sold by the large incumbents because of the strength and nature of the demand for this product. And it’s not just the successful startups such as The Republic of Tea and Adagio Teas that are making an impact in this market. In 1999, Starbucks began offering Tazo, a brand that was started 10 years before by a small company in Portland, Oregon. Once Starbucks began selling this brand, tea sales per store tripled. In fact, there is a company based in Atlanta called Teavana that has started a chain of tea houses.

“I believe that the explosive demand for the new teas can be explained simply by using the traditional economic factors that drive the demand for any product. Let me summarize these factors:

1. *The impact of income on demand:* The new designer teas appeal to consumers with higher incomes. (After all, who would be more willing to pay a premium price for a cup of tea?) One way that we got an insight into this aspect of demand for expensive teas is by looking at the phenomenal growth of the retail food chain Whole Foods. Whole Foods (sometimes called jokingly ‘Whole Paycheck’) caters to exactly the same type of consumer that would buy designer teas. Indeed, next time you visit this store, check out its display of tea and coffee. You’ll be amazed at the variety (and prices) of their stock. One of the key reasons for the success of Whole Foods is an increase in the proportion of U.S. households earning \$100,000 or more from 18 percent 10 years ago to 22 percent in 2004. This increase in the income segment that desires and can afford this product has clearly helped boost the overall market demand for designer teas, exotic coffees, and most of the other products sold at Whole Foods.⁶
2. *The impact of tastes and preferences on demand:* Our preliminary marketing studies indicate that the people who buy this product actually consider it a ‘necessary’ part of their life. In other words, designer tea is not simply a product, but part of a tea drinker’s lifestyle. Moreover, many people in our survey tell us that the preparation and the drinking of tea counteracts the day-to-day stresses of the high-tech, 24/7 work life that we all now seem to lead. Tea is perceived to have a more calming effect than any other beverage, particularly coffee, and even the process of brewing tea becomes a soothing ritual that breaks up the hectic pace of the work day.
3. *Prices of related products:* We do not know for certain, but we believe that coffee is considered a closer substitute for tea than either bottled water or carbonated beverages (particularly those without caffeine). Starbucks’ great success with its Tazo brand of tea indicates that many of the same people who demand fancy coffee seem to desire designer teas. In fact, tea has even been dubbed ‘the new coffee.’⁷ We all know what has been happening to the price of coffee. No doubt the recent increase in coffee prices has helped increase the demand for its close substitute, tea.”

“Okay, Nicole, I really like this idea. But I’m not fully convinced, and I also think we should see how our new initiatives in the energy drink business turn out before we try something else. Have your staff send me more details, and I would be particularly interested if you were able to provide me with a quantitative analysis of the impact of those factors you mentioned on the demand for designer teas. But to be honest, you have really got me thinking seriously about this product. In fact, I wonder how much we would have to pay for The Republic of Tea? But maybe that’s jumping too far ahead. Let’s see the analysis first.”

⁶John Gapper, “Comment: Organic Foods Stores Are on a Natural High,” *Financial Times*, September 16, 2004.

⁷<http://www.wired.com/business/2009/04/tech-millionair/>

SUMMARY

This chapter presents the basic elements of supply and demand. We begin by introducing the *law of demand* and the *law of supply* and the nonprice factors that affect demand and supply. The law of demand states that, other factors held constant, the quantity demanded is inversely related to price. The law of supply states that, other factors held constant, the quantity supplied is directly related to price. Other factors that affect demand are (1) tastes and preferences, (2) income, (3) prices of related products, (4) number of buyers, and (5) future expectations. Other factors that affect supply are (1) costs, (2) technology, (3) prices of other products that sellers can supply, (4) number of sellers, (5) future expectations, and (6) weather conditions. Both numerical and graphical examples of supply and demand and how they interrelate to determine the equilibrium price and quantity were presented. Appendix 3A presents the same material in algebraic terms.

We study how price serves a short-run rationing function and a long-run guiding function in the marketplace. Price serves a rationing function when it increases or decreases to clear the market of a shortage or surplus caused by a change in market conditions (i.e., a shifting of the supply or demand curve). Price changes serve as a guiding function signaling producers or consumers to put more or less of their resources in the affected markets.

In explaining the rationing and guiding functions of price, we note the particular way in which economists define the short run and the long run. We also discuss how comparative statics analysis is used to explain the rationing and guiding functions of the price. This technique, involving the comparison of equilibrium points before and after changes in the market have occurred, is a standard way of analyzing problems and is used throughout this text.

IMPORTANT CONCEPTS

Change in demand. The result of a change in one or more of the nonprice determinants of demand, graphically represented by a *shift* in the demand curve (rightward for an increase in demand and leftward for a decrease in demand). (p. 40)

Change in supply. The result of a change in one or more of the nonprice determinants of supply, graphically represented by a *shift* in the supply curve (rightward for an increase and leftward for a decrease). (p. 43)

Change in the quantity demanded. The result of a change in the price of a good or service, graphically represented by a *movement along* a particular demand curve. (p. 40)

Change in the quantity supplied. The result of a change in the price of a good or service, graphically represented by a *movement along* a particular supply curve. (p. 43)

Demand. Quantities of a good or service that people are ready to buy at various prices, other factors besides the price held constant. Demand can be expressed as a numerical schedule, as a demand curve on a graph, or as an algebraic equation. (p. 39)

Equilibrium price. The price that equates the quantity demanded with the quantity supplied; the price that clears the market of any shortage or surplus. (p. 45)

Equilibrium quantity. The amount that people are ready to buy and sell at the equilibrium price level. (p. 45)

Guiding function of price. Also referred to as the *allocating function of price*, the movement of resources into or out of markets as a result of changes in the equilibrium market price. This is considered to be a long-run function. On the supply side of the market, sellers may enter or leave the market or may vary all their factors of production. On the demand side, consumers may change their tastes or preferences or find long-lasting alternatives to a particular good or service. (p. 50)

Law of demand. The quantity demanded depends *inversely* on price. (p. 39)

Long run. A time period in which new sellers may enter a market or sellers already in a market may leave. This time period is sufficient for both old and new sellers to vary *all* their factors of production. From the standpoint of consumers, the long run provides time enough to respond to price changes by actually changing their tastes or preferences or their use of alternative goods and services. For example, suppose bad weather in Brazil results in an increase in the price of coffee. In the short run, people are expected to buy less coffee because of the higher price. However, in the long run, they may buy even less coffee because the higher price will have prompted them to drink more tea on a regular basis. (p. 50)

Market demand. The sum of all individual demands for a good or service. (p. 39)

Market power. The power to set the market price. (p. 53)

Nonprice determinants of demand. (1) Tastes and preferences, (2) income, (3) prices of related products (i.e., substitutes or complements), (4) future expectations, (5) number of buyers. (p. 41)

Nonprice determinants of supply. (1) Costs, (2) technology, (3) prices of other products that may be produced by a firm, (4) future expectations, (5) number of sellers, (6) weather conditions. (p. 43)

Quantity demanded. The amount that people are ready to buy at a given price. (p. 39)

Quantity supplied. The amount that people are ready to sell at a given price. (p. 42)

Rationing function of price. The increase or decrease in price to clear the market of any shortage or surplus. This is considered to be a short-run function because both buyers and sellers are expected to respond only to price changes. (p. 48)

Shortage. A condition that exists in the market when the quantity demanded exceeds the quantity supplied at a price *below* the equilibrium or market-clearing price. (p. 45)

Short run. A time period in which only those sellers already in the market may respond to a change in market price by using more or less of their variable resources. From the standpoint of consumers, the short run is a period in which they respond only to price changes. As a result of a change in price, consumers may change their tastes or preferences or their use of alternative goods or services. However, in economic analysis, these related changes are considered long-run phenomena. (p. 50)

Supply. Quantities of a good or service that people are ready to *sell* at various prices, other factors besides price held constant. Supply can be expressed as a numerical schedule, as a supply curve on a graph, or as an algebraic equation. (p. 42)

Surplus. A condition that exists in the market when the quantity supplied exceeds the quantity demanded at a price that lies *above* the equilibrium or market-clearing price. (p. 45)

QUESTIONS

1. Define *demand*. Define *supply*. In your answers, explain the difference between *demand* and *quantity demanded* and between *supply* and *quantity supplied*.
2. List the key nonprice factors that influence demand and supply.
3. In defining demand and supply, why do you think economists focus on price while holding constant other factors that might have an impact on the behavior of buyers and sellers?
4. Define comparative statics analysis. How does it compare with sensitivity analysis or what-if analysis used in finance, accounting, and statistics?
5. Define the *rationing function* of price. Why is it necessary for price to serve this function in the market economy?
6. Define the *guiding* or *allocating function* of price.
7. Discuss the differences between the short run and the long run from the perspective of producers and from the perspective of consumers.
8. Explain the difference between shortages and scarcity. In answering this question, you should consider the difference between the short run and the long run in economic analysis.
9. Why do you think it is important for managers to understand the mechanics of supply and demand both in the short run and in the long run? Give examples of companies whose business was either helped or hurt by changes in supply or demand in the markets in which they were competing.
10. "If Congress levies an additional tax on luxury items, the prices of these items will rise. However, this will cause demand to decrease, and as a result the prices will fall back down, perhaps even to their original levels." Do you agree with this statement? Explain.
11. Overheard at the water cooler in the corporate headquarters of a large manufacturing company: "The competition is really threatening us with their new product line. I think we should consider offering discounts on our current line in order to stimulate demand." In this statement, is the term *demand* being used in a manner consistent with economic theory? Explain. Illustrate your answer using a line drawn to represent the demand for this firm's product line.

12. Briefly list and elaborate on the factors that will be affecting the demand for the following products in the next several years. Do you think these factors will cause the demand to increase or decrease?
 - a. Convenience foods (sold in food shops and supermarkets)
 - b. Products purchased on the Internet
 - c. Fax machines
 - d. Digital cameras
 - e. DVDs rented from retail outlets
 - f. Pay-per-view television programming
 - g. Airline travel within the United States; airline travel within Europe
 - h. Gasoline
13. Briefly list and elaborate on the factors that will be affecting the supply of the following products in the next several years. Do you think these factors will cause the supply to increase or decrease?
 - a. Crude oil
 - b. Beef
 - c. Computer memory chips
 - d. Hotel rooms
 - e. Fast food outlets in emerging markets
 - f. Credit cards issued by financial institutions
 - g. Laptop computers
 - h. PC servers

NOTE

In the chapter we emphasized tabular and graphical displays of data but many economic models rely on simple algebraic representations of the same ideas. All but one of the questions below rely on linear demand and supply curves (note economists often call something a curve even when it is a line). The easiest way to make this algebraic leap is to recall that lines are most readily thought of using the slope intercept form, $y = mx + b$. When we talk about demand, we describe quantity as a function of price (and perhaps other things as well). Therefore, y is Q and x is P so we have $Q = mP + b$. The intercept, b , represents how much will be demanded if $P = 0$. The slope, m , describes how much demand changes for each unit change in P . Note that $m < 0$ for demand (this is the Law of demand).

The only other unusual aspect of this situation is that economists switch the axes from those you learned back in algebra and geometry. Economists place the dependent variable, Q , on the horizontal axis and the independent variable, P , on the vertical axis. One can obtain the more standard geometric interpretation by solving for P to obtain inverse demand or inverse supply. In this orientation, you will once again be able to see the slope-intercept form at work. (It is worth noting as an economic aside that inverse demand is interpreted as a marginal valuation schedule.)

These topics are discussed more completely in the Appendix to Chapter 3, but even without formally reading the appendix, you may be able to attack these algebraic problems by remembering the link between the graph of a line and its algebraic representation discussed above.

PROBLEMS

1. The following function describes the demand condition for a company that makes caps featuring names of college and professional teams in a variety of sports.

$$Q = 2,000 - 100P$$

where Q is cap sales and P is price.

- a. How many caps could be sold at \$12 each?
- b. What should the price be in order for the company to sell 1,000 caps?
- c. At what price would cap sales equal zero?

2. Consider the following supply and demand curves for a certain product.

$$Q_S = 25,000P$$

$$Q_D = 50,000 - 10,000P$$

- a. Plot the demand and supply curves.
 b. What are the equilibrium price and equilibrium quantity for the industry? Determine the answer both algebraically and graphically. (Round to the nearest cent.)
3. The following relations describe the supply and demand for posters.

$$Q_D = 65,000 - 10,000P$$

$$Q_S = -35,000 + 15,000P$$

where Q is the quantity and P is the price of a poster, in dollars.

- a. Complete the following table.

Price	Q_S	Q_D	Surplus or Shortage
\$6.00			
5.00			
4.00			
3.00			
2.00			
1.00			

- b. What is the equilibrium price?
4. The following relations describe monthly demand and supply for a computer support service catering to small businesses.

$$Q_D = 3,000 - 10P$$

$$Q_S = -1,000 + 10P$$

where Q is the number of businesses that need services and P is the monthly fee, in dollars.

- a. At what average monthly fee would demand equal zero?
 b. At what average monthly fee would supply equal zero?
 c. Plot the supply and demand curves.
 d. What is the equilibrium price/output level?
 e. Suppose demand increases and leads to a new demand curve:

$$Q_D = 3,500 - 10P$$

What is the effect on supply? What are the new equilibrium P and Q ?

- f. Suppose new suppliers enter the market due to the increase in demand so the new supply curve is $Q = -500 + 10P$. What are the new equilibrium price and equilibrium quantity?
 g. Show these changes on the graph.
5. The ABC marketing consulting firm found that a particular brand of tablet PCs has the following demand curve for a certain region:

$$Q = 10,000 - 200P + 0.03Pop + 0.6I + 0.2A$$

where Q is the quantity per month, P is price (\$), Pop is population, I is disposable income per household (\$), and A is advertising expenditure (\$).

- a. Determine the demand curve for the company in a market in which $P = 300$, $Pop = 1,000,000$, $I = 30,000$, and $A = 15,000$.
 b. Calculate the quantity demanded at prices of \$200, \$175, \$150, and \$125.
 c. Calculate the price necessary to sell 45,000 units.
6. Joy's Frozen Yogurt shops have enjoyed rapid growth in northeastern states in recent years. From the analysis of Joy's various outlets, it was found that the demand curve follows this pattern:

$$Q = 200 - 300P + 120I + 65T - 250A_c + 400A_j$$

- where Q = number of cups served per week
 P = average price paid for each cup
 I = per capita income in the given market (thousands)
 T = average outdoor temperature
 A_c = competition's monthly advertising expenditures (thousands)
 A_j = Joy's own monthly advertising expenditures (thousands)

One of the outlets has the following conditions: $P = 1.50$, $I = 10$, $T = 60$, $A_c = 15$, $A_j = 10$.

- Estimate the number of cups served per week by this outlet. Also determine the outlet's demand curve.
 - What would be the effect of a \$5,000 increase in the competitor's advertising expenditure? Illustrate the effect on the outlet's demand curve.
 - What would Joy's advertising expenditure have to be to counteract this effect?
- Illustrate the example of the world sugar market with supply and demand diagrams. Be sure to show how the relative shifts in supply and demand have led to the reduction in the world price of sugar.
 - Over the years, the market demand for "long-playing records made of polyvinyl has fallen considerably as new technologies replaced the old "LP." Yet, LPs are still available for sale and they sell at price points higher (in some cases much higher) than CDs. According to economic theory, when demand falls, the price of a product should fall. Explain this apparent contraction between theory and fact.
 - Suppose a firm has the following demand equation:

$$Q = 1,000 - 3,000P + 10A$$

- where Q = quantity demanded
 P = product price (in dollars)
 A = advertising expenditure (in dollars)

Assume for the following questions that $P = \$3$ and $A = \$2,000$.

- Suppose the firm dropped the price to \$2.50. Would this be beneficial? Explain. Illustrate your answer with the use of a demand schedule and demand curve.
 - Suppose the firm raised the price to \$4.00 while increasing its advertising expenditure by \$100. Would this be beneficial? Explain. Illustrate your answer with the use of a demand schedule and a demand curve. (*Hint*: First construct the schedule and the curve assuming $A = \$2,000$. Then construct the new schedule and curve assuming $A = \$2,100$.)
- A travel company has hired a management consulting company to analyze demand in twenty-six regional markets for one of its major products: a guided tour to a particular country. The consultant uses data to estimate the following equation (the estimation technique is discussed in detail in Chapter 5):

$$Q = 1,500 - 4P + 5A + 10I + 3P_X$$

- where Q = amount of the product demanded
 P = price of the product in dollars
 A = advertising expenditures in *thousands* of dollars
 I = income in *thousands* of dollars
 P_X = price of some other travel products offered by a competing travel company

- Calculate the amount demanded for this product using the following data:

$$\begin{aligned} P &= \$400 \\ A &= \$20,000 \\ I &= \$15,000 \\ P_X &= \$500 \end{aligned}$$

- Suppose the competitor reduced the price of its travel product to \$400 to match the price of this firm's product. How much would this firm have to increase its advertising in order

to counteract the drop in its competitor's price? Would it be worth it for them to do so? Explain.

- c. What other variables might be important in helping estimate the demand for this travel product?
11. Following are three sample equations. Plot them on a graph in which Q is on the vertical axis and P is on the horizontal axis. Then transform these equations so P is expressed in terms of Q and plot these transformed equations on a graph in which P is on the vertical axis and Q is on the horizontal axis.
- $Q = 250 - 10P$
 - $Q = 1,300 - 140P$
 - $Q = 45 - 0.5P$
12. Use the following equation to derive a demand schedule and a demand curve. What types of products might exhibit this type of nonlinear demand curve? Explain.

$$Q = 100P^{-0.3}$$

13. **Background Information:** Restaurants have traditionally used bottom-end wines to sell by the glass (BTG) at reasonably low prices per glass. In recent years, there has been a growing trend within the restaurant industry to extend BTG programs to higher-end wines. In doing so, they have greatly expanded the sales of higher-end wines. The typical markup on a bottle of wine is 100 percent or more, meaning that a bottle of wine that sells at retail for \$25 in a wine shop can be priced at \$100 in a restaurant. To simplify this discussion, suppose that wine is offered at \$10 per glass. The size of the serving can be adjusted depending on the price per bottle of the wine being poured. For example, a 4-ounce pour can be offered for wines in the \$40 to \$50 range and a 2-ounce pour can be offered for wines in the \$80 to \$100 range. Because a bottle has 25.4 ounces, both examples provide more revenue from BTG sales than from full-bottle sales! Such programs have been quite successful and produce little loss due to unsold wine in the bottle because, with proper storage, a partial bottle can be poured the next evening with no notable change in taste.

The Problem: Suppose you manage a restaurant with a BTG program. You sell 200 glasses per week at \$10 per glass. As an experiment, you once raised the price during the week to \$11 a glass and found that you sold 20 glasses less than before the price change. Suppose you assume this information as the basis for a full demand curve for wine BTG at your restaurant.

- Obtain an algebraic and graphical depiction of your restaurant's BTG demand.
- How many glasses do you expect to sell at \$15 per glass?
- How would your answers change if the response to a \$1 price change is 10 glasses per week or 25 glasses per week?

We now demonstrate how Excel can be used to answer this problem. This allows you to easily conduct a what-if or sensitivity analysis. It also prepares you to use Excel to analyze and estimate more complex demand relationships such as those presented in Chapters 4 and 5.

Note: This problem requires three important assumptions.

- You know a point on the BTG demand curve (e.g., $P = \$10$ and $Q = 200$).
- You know the slope of the BTG demand curve at this point is $\Delta P / \Delta Q = .05$.
- You know that the slope does not change (i.e., demand for wine BTG is linear).

For a review of slopes and equations, go to the Mathematical Appendix on the Companion Website or see Appendix 3A.

Excel Sheet 3.1 shows a screenshot of cells A1:K28 from the Single Demand worksheet of the Demand Excel App. It shows one way to analyze this problem in Excel. The graph is shown next to calculations, and equations are shown next to critical cells, in the worksheet. This is simply done to clarify how each number was obtained: By clicking on a cell, you can check the equations for yourself. This example provides a detailed roadmap for creating a graph in Excel. You are encouraged to open Excel and do this problem yourself.

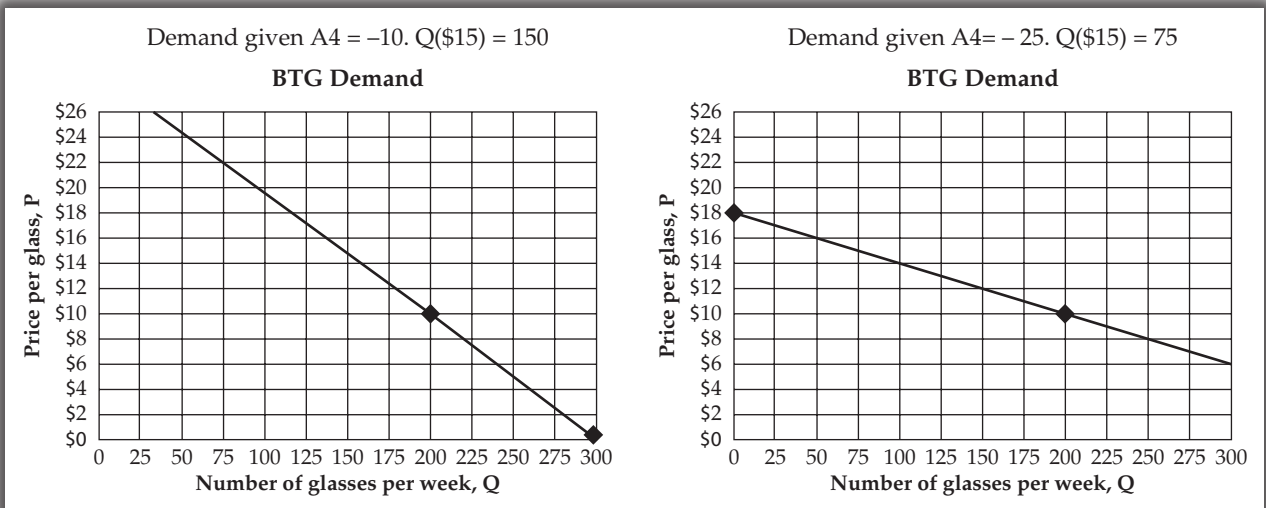
Notice that the negative slope of the demand curve is represented by the negative sign in cell A4 rather than by writing the equation as $P = b - mQ$. Both practices work as long as



	A	B	C	D	E	F	G	H	I	J	K
1	DATA (colored cells)		Answer for B given:			<p style="text-align: center;">BTG Demand</p>					
2	200	Q	New Price in \$		15						
3	10	P in \$	Quantity at this price								
4	-20	$\Delta Q/\Delta P$	100	$=A2+A4*(E2-A3)$							
5	Calculations		C4 uses point-slope formula								
6	-0.05	m, slope ($=1/A4$)									
7	\$ 20.00	b, P Intercept ($= P_0-mQ_0 = A3-A6*A2$)									
8	400	Q intercept ($= -b/m = -A7/A6$)									
9	<p>Creating the graph. The easiest way to graph a line is to use the endpoints. This version includes the demand point to increase clarity.</p>										
12	Equations in A14:B16										
13	Q	P	Q	P	(Q,P) point						
14	0	\$ 20	0	$=A7$	P axis						
15	200	\$ 10	$=A2$	$=A3$	Bundle						
16	400	\$ -	$=A8$	0	Q axis						
17	This graph was obtained by following these steps.										
18	1. Highlight A13:B16 then click Insert, Scatter, Scatter with Straight Lines and Markers.										
19	2. Add axis labels & change the title by clicking on Layout 1 from the Chart Layouts options on the Chart Tools ribbon.										
20	3. Change each title by clicking on the title then type the title in the fx area (above the spreadsheet at D).										
21	a Three titles were created and font sizes were adjusted using the A^ arrow on the Home ribbon.										
22	b The P legend at the right was removed by clicking on it then pressing delete on the keyboard.										
23	c The size of the graph was changed by clicking on the graph and dragging to the preferred size.										
24	4. Change axis formatting by right clicking on either axis then click Format Axis and follow the Format Axis menu.										
25	a This Q axis has a fixed maximum set at 300 with major units set at 20.										
26	b This P axis has a fixed maximum set at 26 with major units set at 2.										
27	5. Vertical major unit gridlines have been added by right clicking on the horizontal axis and click Add Major Gridlines										
28	Part B is answered, and a more general solution is provided, in cell C4 using the point-slope formula for a line.										

Excel Sheet 3.1

Source: Demand Excel App, Single Demand worksheet.



Alternate Figures from Excel Sheet 3.1 based on Changing Demand Responsiveness in Cell A4

you follow through with the mathematics consistently. The benefit of not forcing the minus sign into the equation is that a supply curve can readily be modeled with this graph by entering a positive slope in cell A4. Of course, you also want to adjust the labels appropriately.

One quick benefit of doing the problem in Excel is that the new scenarios in part c are analyzed by changing a single cell, A4, on Excel Sheet 3.1 (the figures below are from Excel Sheet 3.1 based on new values for A4). Additional scenarios can be examined by varying any of the four pieces of data on which the problem is based (in cells A2:A4 and E2 on Excel Sheet 1).

Transferring Your Excel Work to Other Programs: Once you have a graph you want to use, you can easily insert it into another program by highlighting the graph in Excel and then using the Copy command on Excel's Home ribbon. To paste in a Word or PowerPoint document, open that software and use Paste Special, Enhanced Metafile, rather than Paste. From this point, the graph can be resized or positioned as needed by right-clicking on the image; it can be adjusted as needed using the Size and Format Picture menus.

Here is a second, related problem: A successful advertising campaign has increased BTG demand by 25 percent at any price relative to pre-advertising levels. Pre-advertising demand is assumed to be the same as was given in the initial 1Demand scenario (200 glasses are sold at a price of \$10 per glass, and each dollar price change is associated with a twenty-glass-per-week change in sales).

The postadvertising demand implies that 250 glasses are sold at a price of \$10 per glass, and each dollar price change is associated with a twenty-five-glass-per-week change in sales. This is confirmed by having the P intercept the same for both demand curves. Excel Sheet 3.2 shows how this is analyzed in Excel.

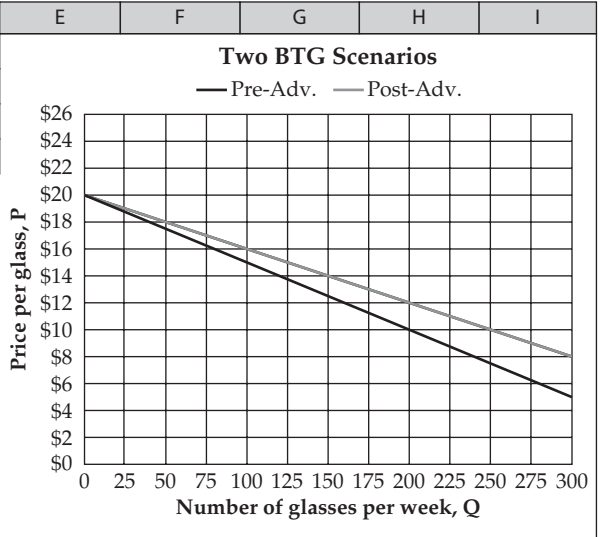
One final scenario: Try doing this problem yourself using the worksheet you created or the 2Demands worksheet. Consider an alternative postadvertising scenario. Rather than selling 25 percent more glasses at any price, you find that if you charge 25 percent more per glass, then you are able to maintain the same number of BTG glasses sold. Adjust D2:D4 to model this new scenario and Copy and Paste (using Paste-Special as described above) your answer graph to a fresh sheet of paper; include answers to a and b:

a. Provide a verbal description of demand in this instance by filling in the values of C and D in the following sentence:

Postadvertising demand shows that 200 glasses are sold at a price of C per glass, and each dollar price change is associated with a D-glass-per-week change in sales.

b. What is the equation for the postadvertising demand curve in slope-intercept form?

	A	B	C	D
1	2 demands	Pre-Adv.	Scenario	Post-Adv.
2	Quantity	200	Q	250
3	Price	10	P in \$	10
4	Change in Q	-20	$\Delta Q/\Delta P$	-25
5	Doing repetitive calculations. The slope and			
6	intercept calculations need to be done for each			
7	demand curve. Once one set of calculations is done,			
8	the other is replicated using Copy and Paste. In this			
9	instance, highlight B12:B14, click Copy, then click on			
10	cell D12 and click Paste.			
11	Equations for B12:B14 are in C12:C14.			
12	slope, m	-0.05	=1/B4	-0.04
13	P intercept, b	\$ 20.00	=B3-B6*B2	\$ 20.00
14	Q intercept	400	=-B13/B12	500
15	Creating the graph. This version graphs the line from			
16	the endpoints in rows 13:14. See 1-5 below.			
17	Q	Pre-Adv.	Q	Post-Adv.
18	0	\$ 20	0	\$ 20
19	400	\$ -	500	\$ -



1. Set up the endpoints for the Pre-Demand curve as shown in A17:B19, using B13:B14.
2. To create the post-advertising demand endpoints, highlight A17:B19, click Copy, click C17 and click Paste.
- 22 Type Post-Adv. in cell D17. (This is another example of having Excel do repetitive calculations.)
3. Highlight A17:B19, click Insert, Scatter, Scatter with Straight Lines. In this instance, Excel assumes you have the data oriented by columns. To obtain the Pre-demand curve, click Switch Row/Column in the Data area of the Chart Tools ribbon. Ignore formatting for the moment.
4. To obtain Post-Demand, Right click on the graph (or use the Chart Tools ribbon) and click Select Data. Click Add. The Edit Series popup menu results.
 - a. Click D17 in the Series Name area. Hit Tab.
 - b. Highlight C18:C19 in the Series X Values area. Hit Tab.
 - c. Hit Delete (to remove =(1)). Highlight D18:D19 in the Series Y Values area, hit Tab, Enter.
5. To format the graph, add labels and adjust axes and vertical grid lines following steps similar to 2-5 from the Single Demand worksheet (except 3b) shown in Excel 3.1.

Excel Sheet 3.2

Source: Demand Excel App, Multiple Demands worksheet

The Mathematics of Supply and Demand



This appendix presents the short-run analysis of supply and demand using algebraic equations and graphs. As you will see, the mechanics of supply and demand can be very concisely expressed in algebraic equations. Furthermore, viewing the demand function in terms of an equation will better prepare you for the next two chapters on demand elasticity and estimation.

The demand function for a good or service can be expressed mathematically as:

$$Q_D = f(P, X_1, \dots, X_n)$$

where Q_D = quantity demanded

P = price

X_1, \dots, X_n = other factors believed to affect the quantity demanded

Using pizza once again as our example, let us assume that price and the nonprice factors affect the demand for pizza in the following way:

$$Q_D = -100P + 1.5P_{hd} - 5P_{sd} + 20A + 15Pop \quad (3A.1)$$

where Q_D = quantity demanded for pizza (pies)

P_{hd} = price of hot dogs (cents)

P_{sd} = price of soft drinks (cents)

A = advertising expenditures (thousands of dollars)

Pop = percentage of the population aged 10 to 35

Suppose we hold constant all factors affecting the quantity demanded for pizza except price by assuming the values of these nonprice factors to be

$$P_{hd} = 100 \text{ (\$1.00 or 100 cents)}$$

$$P_{sd} = 75 \text{ (\$.75 or 75 cents)}$$

$$A = 20 \text{ (\$20,000)}$$

$$Pop = 35 \text{ (35 percent)}$$

Substituting these values into Equation (3A.1) gives us

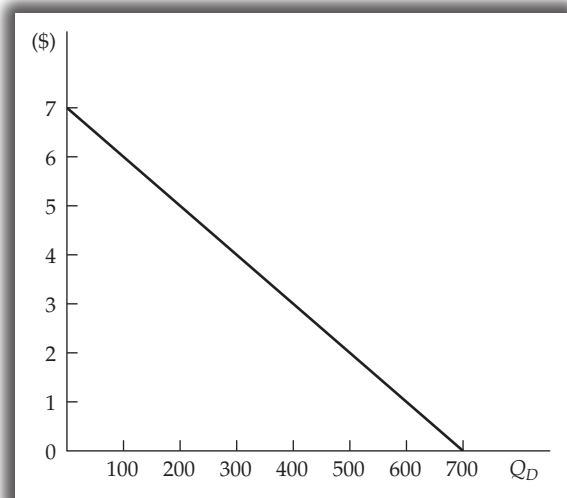
$$Q_D = -100P + 1.5(100) - 5(75) + 20(20) + 15(35)$$

$$Q_D = 700 - 100P \quad (3A.2)$$

All the values of the nonprice variables are now included in the constant term, 700. Plotting this equation gives us the demand curve shown in Figure 3A.1.

Those familiar with the graphical presentation of algebraic equations may be puzzled about the way in which economists present the supply and demand equations in graphical form. As a rule, the dependent variable is placed on the vertical or Y axis, and the independent variable is placed on the horizontal or X axis. Given this format, one would expect Q , the dependent variable, to be placed on the vertical axis and P , the independent variable, to be placed on the horizontal axis. However, in this chapter, as well as in Chapter 4, Q is placed on the horizontal axis, and P on the vertical axis. It seems that the originator of these diagrams, Professor Alfred Marshall, first presented them in this manner.⁸

Figure 3A.1 Demand Curve



⁸Alfred Marshall, *Principles of Economics*, 8th ed., Philadelphia: Porcupine Press, 1920, reprinted 1982, p. 288.

Regardless of Marshall’s original reasons for reversing the axes, let us simply state that in the analysis of cost, revenue, and profit, the quantity of output is the independent variable. Thus, placing Q on the horizontal axis in the analysis of supply and demand simply prepares us for its subsequent designation as an independent variable.

In the meantime, an adjustment must be made in linking the supply and demand equations to their graphs in order to conform to mathematical convention. In supply-and-demand analysis, whenever an equation such as $Q_D = 700 - 100P$ is plotted on a graph, we must do one of two things. If we want to be consistent with mathematical convention, we must place Q_D on the vertical axis and the P on the horizontal axis. This is shown in Figure 3A.2a. If we want to follow the usual format in economics, we must rearrange the terms in the equation so P is expressed in terms of Q_D . This orientation, in which P is a function of Q_D , is called inverse demand.

$$\begin{aligned} Q_D &= 700 - 100P \\ 100P &= 700 - Q_D \\ P &= \frac{700 - Q_D}{100} \\ P &= 7 - 0.01Q_D \end{aligned}$$

As such, P is now the dependent variable and can be plotted on the vertical axis. Q_D is now the independent variable and can be plotted on the horizontal axis. This is illustrated in Figure 3A.2b.

Let us review this point by assuming one of the nonprice factors affecting the quantity demanded for pizza has changed. In particular, suppose the price of hot dogs increases to \$1.20. In Equation (3A.2), this would increase the constant or “Y intercept” term from 700 to 730. This, in effect, would cause the demand curve to shift from its original position to the new one shown in Figure 3A.2a. Figure 3A.2b shows the effect of an increase in the price of hot dogs on the transformed demand equation. In this case, the constant term or “Y intercept” increases from 7 to 7.3 and is also shown by a rightward shift in the demand curve.

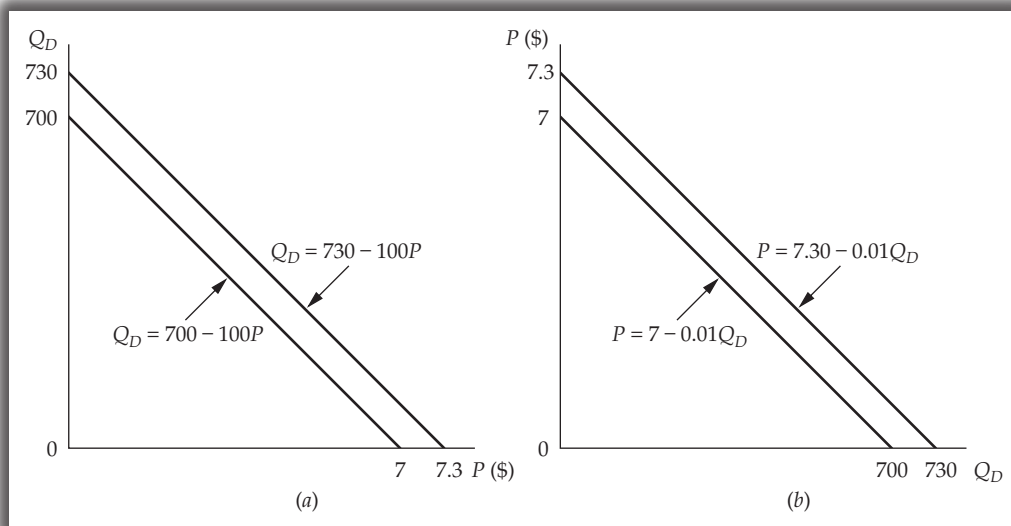
We now focus on the supply by assuming the supply equation is the same as the curve used in Table 3.4 and Figure 3.3. This equation can be expressed as

$$Q_S = -100 + 100P \tag{3A.3}$$

Once the supply and demand equations are given, there are several ways to find the equilibrium price and quantity. One way is to solve for the supply and demand equations simultaneously. This is done by first setting up the two equations in the following way:

$$\begin{aligned} Q_D &= 700 - 100P \\ Q_S &= -100 + 100P \end{aligned}$$

Figure 3A.2 Transforming the Demand Curve



We can then eliminate P by adding the equations together. This gives us

$$\begin{aligned} 2Q &= 700 - 100 \\ &= 600 \\ Q^* &= 300 \end{aligned} \quad (3A.4)$$

To find the equilibrium price (P^*), we simply return to either the demand or the supply equation, insert the value for the equilibrium quantity, (300), and solve for P . Using Equation (3A.2), this would give us

$$\begin{aligned} 300 &= 700 - 100P \\ 100P &= 700 - 300 \\ P^* &= 4 \end{aligned}$$

Notice that when we added the supply and demand equations, we no longer made the distinction between Q_s and Q_d because in equilibrium $Q_s = Q_d$. In fact, this brings us to the other way in which the supply and demand equations can be used to determine the equilibrium price and quantity. By definition, market equilibrium occurs when the quantity supplied is equal to the quantity demanded. Thus, we can set Equation (3A.2) equal to Equation (3A.3) and solve for the unknown P . That is,

$$\begin{aligned} 700 - 100P &= -100 + 100P \\ 200P &= 800 \\ P^* &= 4 \end{aligned}$$

By inserting the value of 4 into either equation, we obtain the equilibrium quantity, 300. This solution was depicted graphically as Figure 3.4, p. 45.

You now have three ways to view the basic elements of supply and demand. First, there are supply and demand schedules, as shown in Tables 3.1, 3.2, 3.3, and 3.4, in which the equilibrium price is found by matching the quantity supplied with the quantity demanded. Second, there are supply and demand diagrams, as presented in most of the figures in this chapter, in which price and quantity are determined by the intersection of the supply and demand curves. Finally, there are supply and demand equations that enable one to find equilibrium price and quantity by solving for the unknowns in the two equations. For teaching purposes, the use of graphs is favored. But regardless of the manner in which the concepts of supply and demand are presented, there remains the challenge for business decision makers to ascertain the actual supply and demand data for their particular industries and organizations.

Demand Elasticity

Learning Objectives

Upon completion of this chapter, readers should be able to:

- Define and measure elasticity.
- Apply the concepts of price elasticity, cross-elasticity, and income elasticity.
- Understand the determinants of elasticity.
- Show how elasticity affects revenue.

The Situation

Henry Caulfield is the owner-operator of a local Gas 'n Go gas station and convenience store. Henry chose to locate his store in an area at least 10 minutes away by car from the nearest supermarket or grocery store. For the most part, Henry's business has been quite successful.

Then one day he noticed that a new grocery store was opening just one block away. A month later he noticed that a new convenience store had opened for business less than a 3-minute drive away. Henry realized that to maintain the status quo in the face of this new competition, he would have to make some tough decisions about his pricing and promotion policies and the mix of items carried in his store.

The items that he carried were typical of those found in retail establishments of this type, with beer, cigarettes, hot coffee, and soft drinks accounting for about 75 percent of total sales. Soft drinks were by far the best-selling item in his store. Essentially, the retail price of soft drinks was based on the wholesale price plus a markup of about 400 percent. Henry recognized that this markup was considerably higher than the one used by a supermarket, but he believed that people were willing to pay more for the convenience.

On certain occasions, Henry would offer a particular brand of soft drink at a substantial discount. Regardless of whether he lost or made money on soft drinks by this action, he found that it helped to attract additional customers into his store, and gasoline sales actually increased. Several people from an adjacent town told him that they waited until they were in the vicinity of his station to fill their tanks because of his discount on soda. Given the public's responsiveness to special discounts on soft drinks and the ability of this product to promote other products, Henry decided to use the pricing of soft drinks as his main weapon against his new competition. Instead of offering a temporary discount, he decided to reduce the price of soft drinks permanently. However, after a month, despite the lower soft drink prices, there was a noticeable decline in his revenue from soft drinks. Henry realized that he would have to reassess his competitive tactic.

THE ECONOMIC CONCEPT OF ELASTICITY

In Chapter 3, we study the idea of demand and discuss the movement along a demand curve (i.e., change in quantity demanded). The demand curve sloped downward to the right; this means, of course, that the lower the price, the greater the quantity of the product consumed. We are now going to discuss the question of how sensitive the change in quantity demanded is to a change in price. The measurement of this sensitivity in percentage terms is called the **price elasticity of demand**. Henry Caulfield made implicit use of this concept when he decided to lower his soft drink prices to compete with the new stores in his area. But this is only one of the elasticity measures with which we concern ourselves in this chapter. We also cover the concepts of income elasticity, cross-elasticity, and supply elasticity.

In most general terms, we can define **elasticity** as a percentage relationship between two variables, that is, the percentage change in one variable relative to a percentage change in another. In different terms, we divide one percentage by the other:

$$\text{Elasticity} = \frac{\text{Percent change in A}}{\text{Percent change in B}}$$

The result of this division is **elasticity**. It is then our task to interpret it and determine the effects of the change. The meaning of the size and the sign of elasticity (that may be negative or positive) is the focus of our inquiry for the remainder of this chapter. Let us first turn to the most frequently encountered elasticity concept, the price elasticity of demand.

THE PRICE ELASTICITY OF DEMAND

When Henry Caulfield contemplated lowering his price to counteract his new competition, he was dealing with price elasticity of demand. He was determining whether by lowering his prices he would raise his unit sales sufficiently to increase his total revenue.¹

When we speak of the price elasticity of demand, we are dealing with the sensitivity of quantities bought to a change in the producer's price. Thus, this concept describes an action that is within the producer's (or, in this case, the dealer's) control. Other elasticities discussed later are outside the producer's control and may evoke other actions on the producer's part to counteract them.

The price elasticity of demand is defined as a percentage change in quantity demanded caused by a 1 percent change in price. Let us develop this concept mathematically. We can write the expression, "percentage change in quantity demanded" as

$$\frac{\Delta \text{Quantity demanded}}{\text{Initial quantity demanded}}$$

where Δ (delta) signifies an absolute change. The second part of this relationship, "percentage change in price," can be written as

$$\frac{\Delta \text{Price}}{\text{Initial price}}$$

¹More important, he was wondering whether he would actually increase his profits with this action. However, we are not yet in a position to deal with this question.

Dividing the first expression by the second, we arrive at the expression for the price elasticity of demand:

$$\frac{\Delta \text{Quantity}}{\text{Quantity}} \div \frac{\Delta \text{Price}}{\text{Price}} = \frac{\% \Delta \text{Quantity}}{\% \Delta \text{Price}}$$

This is the general expression. We turn now to the actual computation of elasticities, and we describe two methods of obtaining the price elasticity of demand.



Module 4A

Measurement of Price Elasticity

Let us begin with **arc elasticity**, the method most commonly used in economics textbooks. The formula for this indicator is

$$E_p = \frac{Q_2 - Q_1}{(Q_1 + Q_2)/2} \div \frac{P_2 - P_1}{(P_1 + P_2)/2}$$

where E_p = Arc price elasticity
 Q_1 = Original quantity demanded
 Q_2 = New quantity demanded
 P_1 = Original price
 P_2 = New price

The numerator of this fraction, $(Q_2 - Q_1)/[(Q_1 + Q_2)/2]$, indicates the percentage change in the quantity demanded. The denominator, $(P_2 - P_1)/[(P_1 + P_2)/2]$, indicates the percentage change in the price.

Notice that the change in each variable is divided by the *average* of its beginning and ending values. For example, if the price of a product rises from \$11 to \$12, causing a fall in the quantity demanded from 14 to 12, the formula gives the following price elasticity coefficient:

$$\begin{aligned} E_p &= \frac{12 - 14}{(14 + 12)/2} \div \frac{12 - 11}{(11 + 12)/2} \\ &= \frac{-2}{13} \div \frac{1}{11.5} \\ &= \frac{-2}{13} \times \frac{11.5}{1} \\ &= \frac{-23}{13} \\ &= -1.77 \end{aligned}$$

The reason the arc elasticity formula employs the average of the beginning and ending values can be clearly seen. If we had used the beginning values, the coefficient would be

$$\begin{aligned} E_p &= \frac{12 - 14}{14} \div \frac{12 - 11}{11} \\ &= \frac{-2}{14} \div \frac{1}{11} \\ &= \frac{-1}{7} \times \frac{11}{1} \\ &= \frac{-11}{7} \\ &= -1.57 \end{aligned}$$

However, suppose the price fell from \$12 to \$11, causing the quantity demanded to rise from 12 to 14 units. Using the beginning values would result in a coefficient of -2 (readers can make this calculation themselves to obtain the answer). Thus, the *same* unit change in price and quantity gives *different* values of elasticity, depending on whether the price increases or decreases.² By using the average of the beginning and ending values, we avoid this ambiguity. The price elasticity coefficient is the same whether price increases or decreases.

An additional source of ambiguity arises in the computation of elasticity when we consider changes over different ranges of price and quantity. For example, suppose the values of price and quantity provided in the preceding analysis are part of the hypothetical demand schedule shown in Table 4.1.

The numbers in this schedule indicate a linear relationship between quantity demanded and price, with a unit change in price resulting in a two-unit change in quantity over the entire range of the schedule.³ Suppose we compute the arc elasticity for a price change from \$12 to \$10 rather than between \$12 and \$11. Using the arc elasticity formula gives

$$\begin{aligned} E_p &= \frac{12 - 16}{(16 + 12)/2} \div \frac{12 - 10}{(10 + 12)/2} \\ &= \frac{-4}{14} \div \frac{2}{11} \\ &= \frac{-4}{14} \times \frac{11}{2} \\ &= \frac{-44}{28} \\ &= -1.57 \end{aligned}$$

Table 4.1 Hypothetical Demand Schedule

Price	Quantity
\$18	0
17	2
16	4
15	6
14	8
13	10
12	12
11	14
10	16
9	18
8	20
7	22
6	24
5	26

²The reason for this ambiguity is simply that the base number differs between a percentage increase and a decrease between two numbers. A good example of this can be found in the retail trade. Suppose a company buys a dress wholesale for \$100 and marks it up 100 percent, thereby establishing the retail price at \$200. Suppose in a clearance sale the company decides to sell it at cost. This would represent a 50 percent markdown (i.e., from \$200 down to \$100).

³The algebraic expression of this demand equation is $Q = 36 - 2P$ and the inverse demand equation is $P = 18 - 0.5Q$.



Notice that the coefficient is different from the previously computed value. In fact, for any given value of price, the arc elasticity coefficient will vary depending on the new price's distance from the original price.⁴

To adjust for the ambiguity inherent in the use of the arc formula, economists recommend the use of **point elasticity**, the second of the two ways to compute the elasticity coefficient. This method of computation is expressed as follows (we use the Greek letter ϵ when we are referring specifically to point elasticity):

$$\epsilon_p = \frac{\Delta Q}{\Delta P} \times \frac{P_1}{Q_1}$$

When demand is linear, $\Delta Q/\Delta P = m$, the slope of the demand equation $Q = b + mP$. Since demand is downward sloping, we know that $m < 0$.

The equation implied in Table 4.1 is $Q = 36 - 2P$. Thus, the point elasticity coefficient at \$12 and 12 units is:

$$\begin{aligned}\epsilon_p &= -2 \times \frac{12}{12} \\ &= -2\end{aligned}$$

When demand is nonlinear, the calculation of $\Delta Q/\Delta P$ is somewhat more complicated because the slope of a curve changes. This slope is readily obtained using the calculus concept of derivative. In this instance,

$$\epsilon = dQ/dP \times P_1/Q_1$$

The derivative of Q with respect to P (i.e., dQ/dP) is simply the instantaneous version of slope. If you have not seen derivatives before you can learn about them in the Mathematical Appendix on the Companion Website.

To determine dQ/dP , we need an algebraic expression for demand. For example, consider the following demand curve:

$$Q = 100 - P^2$$

In this instance, $dQ/dP = -2P$.

Assuming $P_1 = 5$, then $Q_1 = 75$. Then the point elasticity is

$$\begin{aligned}\epsilon_p &= -2P \times \frac{5}{75} \\ &= -2(5) \times \frac{5}{75} \\ &= \frac{-50}{75} \\ &= -0.67\end{aligned}$$

Although it is very convenient to use linear demand curves, in reality the shape of the demand curve may be different. In the preceding illustration, price decreases will bring forth smaller increases in quantity. Such a demand curve would take on a concave

⁴Interested readers should try computing the arc elasticity for changes between \$12 and \$9, between \$12 and \$8, and so on. They will find that the arc elasticity decreases as the change in price increases.

shape. In contrast, the demand curve may be convex. (The equation for such a curve would be, for instance, $Q = 10/P^2$.) An example of such a demand curve would occur when, for instance, some quantities would be purchased even at a very high price; in such a case, the demand curve would become almost vertical near the price axis.

So far we have discussed both linear and nonlinear demand curves whose elasticity changes as we move along the curve. However, we could encounter a demand curve whose elasticity is constant over its relevant range. Such a curve would be described by the following equation:

$$Q = aP^{-b}$$

a being a constant, and $-b$ representing the price elasticity of demand. This nonlinear equation can be converted to linear by expressing it in logarithms:

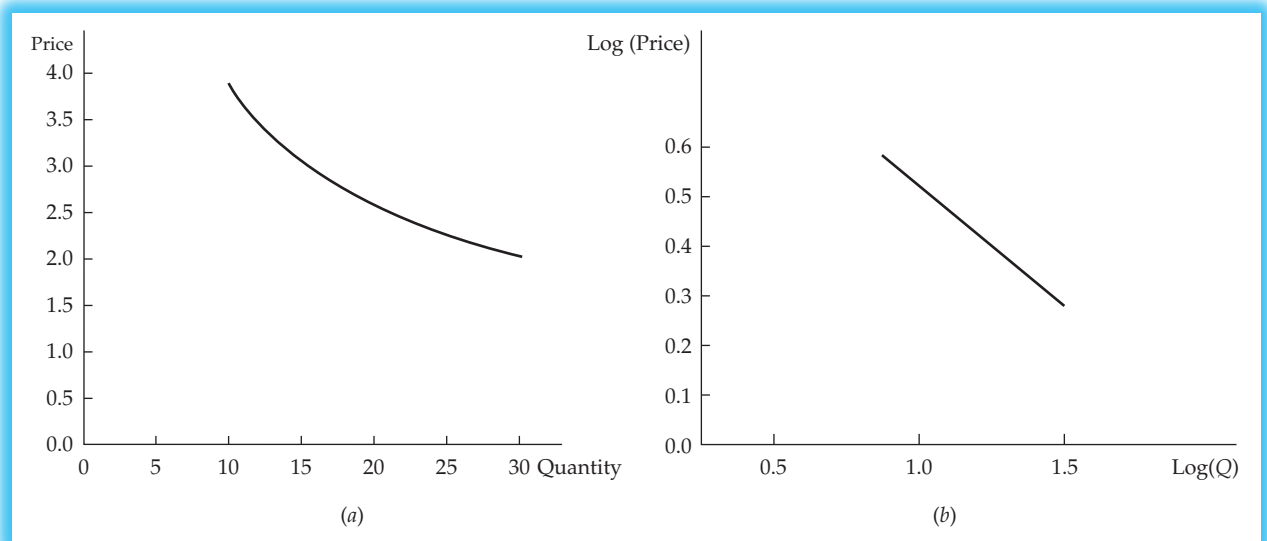
$$\log(Q) = \log(a) - b \cdot \log(P)$$

Such a demand curve would plot as a straight line on double log (or log-log) graph paper, and its elasticity ($-b$) would be the same at any point on the curve. For example, let the demand equation be $Q = 100P^{-1.7}$. From this equation, we can generate the following demand schedule:

Quantity	10	12.5	15	17.5	20	22.5	25	27.5	30
Price	3.875	3.398	3.052	2.788	2.577	2.405	2.260	2.137	2.030
Log(Q)	1.0	1.10	1.18	1.24	1.30	1.35	1.40	1.44	1.48
Log(P)	0.59	0.53	0.48	0.45	0.41	0.38	0.35	0.33	0.31

The two graphs in Figure 4.1 illustrate this demand curve. Figure 4.1a shows the curve using the normal arithmetic scales. In Figure 4.1b the scales are transformed into logarithms, and the demand curve plots as a straight line, signifying that the elasticity is constant throughout.⁵

Figure 4.1 Constant Elasticity Demand



⁵We encounter constant elasticity again in the appendix to Chapter 5 and in Chapter 6 when we discuss the Cobb-Douglas production function. The CED Excel App on the Companion Website provides further insight into constant elasticity demand (CED) functions by allowing you to vary the elasticity of the CED function.

The concept of point elasticity, as well as the use of calculus, is found to be of particular importance in Chapter 5, when the estimation of demand equations is discussed. However, for now, using discrete changes and arc elasticity would appear to be the more realistic thing for Henry Caulfield (and most other practical businesspeople) to do. He may not be familiar with calculus, but this certainly does not detract from his business acumen. He is dealing with a concrete problem: how much he will sell if he lowers his price by a discrete quantity (i.e., a certain number of cents). Arc elasticity is perfectly suitable for this problem.

We must realize, however, that in actual business situations, the effects of elasticity may be calculated in simple percentage terms, using the starting price and quantity as base numbers. For instance, a businessperson may state that it is expected that a 10 percent decrease in price will result in a 20 percent increase in quantity, implying that elasticity is -2 ; as an example, a decrease in price from \$10 to \$9 will increase quantity sold from 1,000 to 1,200 units. Of course, this would suggest that raising the price from \$9 to \$10 (an 11.1 percent increase) would decrease quantity from 1,200 to 1,000 units (a decrease of 16.7 percent). The elasticity in this case is -1.5 . This asymmetry creates the same problem discussed earlier, which was solved by employing arc elasticity. In actual situations, however, we must be realistic and understand that a movement down along the demand curve may not bring about the same results as a movement up along the demand curve. Even more important, we should remember that mathematical refinement may not be the most essential thing. What is critical to the businessperson faced with a decision is whether a decrease in price will create sufficiently more quantity to improve profits.

Economists, in their neat way, have created categories of elasticity:

1. *Relative elasticity of demand:*

$$E_p > 1 \text{ (in absolute terms)}^6$$

This occurs when a 1 percent change in price causes a change in quantity demanded greater than 1 percent. The coefficient calculated earlier, 1.77, is a case of relatively elastic demand.

2. *Relative inelasticity of demand:*

$$0 < E_p < 1 \text{ (in absolute terms)}$$

Here the percentage change in price is greater than the corresponding change in quantity. For example, in Table 4.1 as price is lowered from 8 to 7, quantity rises from 20 to 22, giving a coefficient of 0.71.

3. *Unitary elasticity of demand:*

$$E_p = 1 \text{ (in absolute terms)}$$

A 1 percent change in price results in a 1 percent change in quantity in the opposite direction.

These are three common measures of price elasticity. But there are also two limiting cases at the extremes of the elasticity scale:

1. *Perfect elasticity:*

$$E_p = \infty \text{ (in absolute terms)}$$

⁶Because a demand curve has a negative slope, price elasticity will be negative. However, it is often more efficient to ignore the minus sign and to discuss the elasticity in absolute terms. Later in this chapter, we discuss cross-price elasticity and income elasticity. Here absolute terms cannot be used because the existence of a negative or positive sign gives the result a significantly different meaning.

In this case, there is only one possible price, and at that price an unlimited quantity can be sold. The demand curve for $E_p = \infty$ is a horizontal line. We encounter a demand curve with this shape later, when we discuss perfect competition.

2. *Perfect inelasticity:*

$$E_p = 0$$

Under this condition, the quantity demanded remains the same regardless of price. Such a demand curve may exist for certain products within a particular price range. An example may be the case of salt. Today's price of salt is about 44 cents per pound. If this price were to rise to 54 cents (a significant percentage increase), or fall to 34 cents (a significant percentage decrease), it is very doubtful that the consumption of salt would change.

Both of these extreme cases, although possible under certain conditions, will seldom be observed in real life. Still, the two limits should be well understood by every student of economics.

The Determinants of Elasticity

Now that we have described what elasticity is, let us look into the reasons that the demand for some goods and services is elastic, whereas for others it is inelastic. In other words, what determines elasticity? As we look into this question, we must remember that the elasticity for a particular product may differ at different prices. Thus, although the demand elasticity for salt is very low—possibly zero—in the vicinity of its current price, it may not be so inelastic at \$5 or \$10 per pound.

It is often said—and many use this as a rule of thumb—that demand is inelastic for goods considered to be necessities, and it is elastic for luxury products. For example, the demand for furs, gems, and expensive automobiles is probably more elastic than the demand for milk, shoes, and electricity.

Unfortunately, the luxury/necessity dichotomy is ambiguous. Demand for expensive automobiles may be elastic, but if we consider the demand for Mercedes autos, we will probably find that within the prevailing price range, a movement up or down of several thousand dollars would make relatively little difference to those people who are in the market for this particular kind of car. The probable reason for such inconsistencies is relatively simple: one person's luxury is another person's necessity.

Several important factors that influence demand elasticity are outlined in Figure 4.2 and discussed in the next few pages.

Probably the most important determinant of elasticity is ease of substitution. This argument cuts both ways: if there are many good substitutes for the product in

- ▶ Ease of substitution
- ▶ Proportion of total expenditures
- ▶ Durability of product
 - ▶ Possibility of postponing purchase
 - ▶ Possibility of repair
 - ▶ Used product market
- ▶ Length of time period

Figure 4.2 Factors Affecting Demand Elasticity

question, elasticity will be high; conversely, if this commodity is a good substitute for others, its demand elasticity will also be high. The broader the definition of a commodity, the lower its price elasticity will tend to be because there is less opportunity for substitution. For instance, the demand elasticity for beer or bread will tend to be less than that for a particular brand of beer or for white bread. There are fewer substitutes for bread in general (particularly if we include in this definition other baked products, such as rolls and bagels) than there are for white bread or, even further, a specific brand of white bread. If the price of bread rises (relative to other products), we may consume somewhat less bread than before. However, if the price of brand A white bread rises while other white bread prices remain the same, then one would expect the quantity of brand A demanded to drop significantly as consumers switch to other brands.

Henry's convenience store was once the only game in town, so to speak. His nearest competitor was relatively distant. Now customers can substitute for Henry's merchandise by walking one block to the grocery store. Because Henry was most likely selling the same soft drink brands as his close competitor, the substitution effect is extremely strong.

Another major determinant of demand elasticity is the proportion of total expenditures spent on a product. Here we can go back to our salt example. The reason for the low elasticity of demand for salt is that the proportion of a consumer unit's (e.g., an individual, a family, and so on) income spent on salt is extremely small. A hefty price increase (e.g., from 44 cents to 54 cents per pound) would probably cause a shrug of the shoulder and would affect consumption of salt very little.

The spending on soft drinks by a typical individual or a family constitutes a larger portion of income than spending on salt. However, in most circumstances, spending on soda still represents a relatively small percentage of a family's income. Thus, we would not expect a change in price to affect the quantity demanded significantly. Still, in households where large quantities of soft drinks are consumed, a price change could have some effect on quantities sold, although it would probably require a substantial change in price to affect purchases significantly.

However, for a product such as a large appliance, the situation may be entirely different. To most families, a clothes washer represents more than a trivial expense, and a price change could have an important impact on purchases. Thus, we expect the demand elasticity for a clothes washer to be considerably greater than that for salt or soft drinks. There is another possible reason for the relatively high demand elasticity of a clothes washer. An appliance purchase may be postponable because there is commonly a choice among buying, repairing, and using a Laundromat. Faced with a higher purchase price, a consumer may choose to repair the old appliance or go to a Laundromat.⁷

Despite the entry of new firms into Henry's market, the geographic size of the market is limited to a relatively small local area. As markets broaden, more and more product substitution becomes possible. Advances in modes of transportation and communication and decreases in their cost have increased the size of markets over time. Thus, the number of substitutes competing for consumers' dollars has increased.

⁷The choice between buying and maintaining becomes even more pronounced for the purchase of an automobile. Here, of course, the price and the proportion of a person's income are considerably higher. In this case, there is a third possibility for the consumer: the purchase of a used car. Thus, recalling what we said about substitutability, we can say that the elasticity of demand for cars in general is lower than the demand elasticity for new cars. Of course, it is also possible to purchase a used washing machine, but there is no organized market for these as there is for cars. We will have more to say about the used car market in Chapter 11.

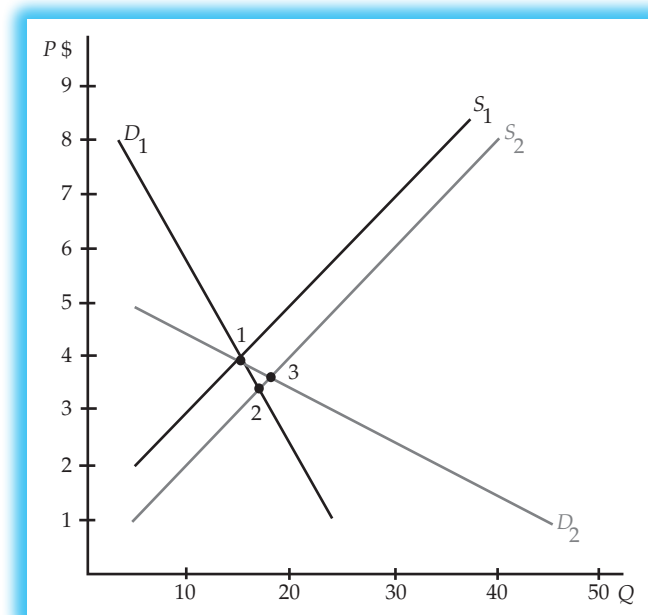
Markets have not only widened within national borders, they have more frequently crossed borders, increasing the importance of international trade. Although advances in transportation and communications have been instrumental, an extremely important trend toward freer trade through international trade agreements has undermined artificial barriers (tariffs and quotas). This has increased competition globally and thus has increased demand elasticities facing firms. The consumer is the ultimate beneficiary of such trends.

The Effect of Elasticity on Price and Quantity

The response to a change in supply is determined, to a great extent, by the price elasticity of demand. This is of great importance to a businessperson who must be able to adjust its production and prices to a change in economic conditions. Except in the case of perfectly elastic or perfectly inelastic demand, both price and quantity produced will be affected. The extent of the change of each is determined by the price elasticity of demand.

In Figure 4.3, we have charted a demand curve D_1 whose equation is $Q = 27 - 3P$. The equation for the supply curve S_1 is $Q = -5 + 5P$. The resulting equilibrium price is \$4 and quantity is 15 units (point 1).⁸ If we chart another, considerably more elastic, demand curve, D_2 , of the form $Q = 55 - 10P$, with the same supply curve, the equilibrium point will be the same, $Q = 15$, and $P = 4$. Now, let us say that supply increases, and the supply curve S_2 is of the form $S = 0 + 5P$. With the new supply curve, the equilibrium points for the two demand curves will no longer be the same. With D_1 the price will be \$3.375 and quantity 16.875 units (point 2). However, with D_2 the result will be $P = 3.667$ and $Q = 18.333$ (point 3). The more elastic demand curve, D_2 , has resulted in a smaller price decrease but a larger quantity increase than D_1 , the less

Figure 4.3 Effect of Elasticity on Price and Quantity



⁸The reader can use Excel Module 3A, which can be accessed on the Companion Website, to calculate the equilibrium point and draw a graph.

elastic curve. The reader should try this exercise with the same demand curves, but with a decrease in supply to, say, $Q = -10 + P$. In this case, the more elastic demand curve will result in a larger decrease in quantity and a smaller increase in price. Thus, we can see that it is extremely important for a manager to have a good idea regarding the sensitivity of demand to price changes. Making the wrong decision in raising or lowering price in response to changing supply conditions could be disastrous to the manager's company.

The Elasticity of Derived Demand

This section of the elasticity discussion represents a small digression, albeit an important one. So far we have discussed demand elasticity for a final product, that is, a product purchased for consumption, such as soft drinks, a clothes washer, salt, white bread, or beer.

We are now going to look briefly at the demand for items that go into the production of a final commodity, such as materials, machinery, and labor. The demand for such components of a final product is called **derived demand**. In other words, these components are not demanded for their own sake but because there is a demand for the final product requiring them.

The great British economist Alfred Marshall, about whom previous references are made, described four principles governing the elasticity of the derived demand curve.⁹ According to Marshall, the derived demand curve will be more inelastic:

1. The more essential is the component in question
2. The more inelastic is the demand curve for the final product
3. The smaller is the fraction of total cost going to this component
4. The more inelastic is the supply curve of cooperating factors

An example illustrates this concept. Let us consider demand for residential housing (the final product) and the derived demand for one class of labor employed in construction, electricians. After all, the demand for electricians does not exist for its own sake but is due to the demand for housing. Probably all of Marshall's principles apply in this case, but two of them are particularly important. The first is essentiality: you just cannot build a house without employing electricians. Second, the cost of electrical work is probably a relatively small percentage of the entire cost of the house.

Suppose the electricians demand and obtain a substantial wage increase. A contractor may try to cut a few corners with regard to electrical work, but most of it must still be done. Thus, the employment of electricians will not decrease much. The implication here is that the elasticity of demand for electricians is low. Assume the work of electricians involves 10 percent of the total cost of construction (this cost is probably overstated). A 10 percent wage increase for electricians represents a 1 percent increase in the total cost of construction. This small addition to the total cost will most likely not trigger a price increase and thus will not affect the employment of electricians to any significant extent. If we also consider the probability that the demand for housing is somewhat inelastic and that the supply elasticity of cooperating factors (i.e., other crafts employed on the project) is rather low, we can conclude that the demand elasticity for electricians is relatively low.

These conclusions tend to hold in the short run much more than in the long run. Over a short period of time, employment of electricians will not drop very much. However, given a longer adjustment period, elasticity of demand will rise as people

⁹Alfred Marshall, *Principles of Economics*, 8th ed., Philadelphia: Porcupine Press, reprinted 1982, pp. 319–20.

find ways to substitute for the expensive factor, both on the production side and on the consumption side.¹⁰

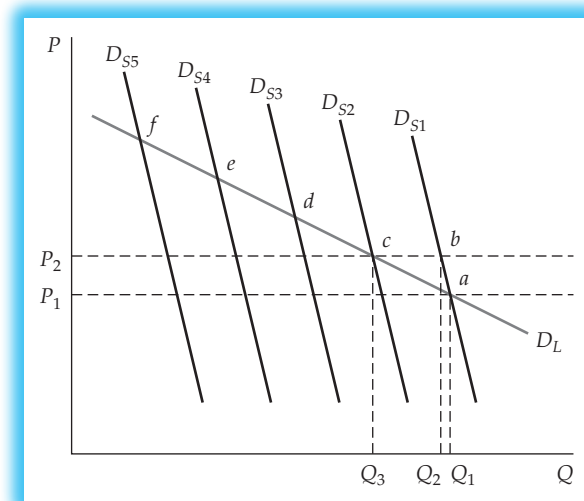
Elasticity in the Short Run and in the Long Run

A long-run demand curve will generally be more elastic than a short-run curve. Here “short run” is defined as an amount of time that does not permit a full adjustment by consumers to a price change. In the shortest of runs, no adjustment may be possible, and the demand curve over the relevant range may be almost perfectly inelastic. As the time period lengthens, consumers will find ways to adjust to the price change by using substitutes (if the price has risen), by substituting the good in question for another (if the price has fallen), or by shifting consumption to or from this particular product (i.e., by consuming more or less of other commodities).

A good example is the case of energy costs. When heating oil prices shot up in the 1970s, the immediate response by consumers was not great. As time passed, however, consumers adjusted their oil consumption. They became used to lower temperatures around the house and at work. They began to dress more warmly indoors. (Would this result in a higher demand for sweaters? We look at this particular idea, known as cross-elasticity, later in this chapter.) Over a still longer period, consumers (including one of the authors) converted their homes from oil heat to gas heat. Not only was there conversion, but more newly built homes were equipped with gas heat. In addition, homes up for resale would advertise gas heating to attract potential buyers, and gas-heated homes commanded a premium price. How can we demonstrate this phenomenon graphically?

We can represent this relationship between the short run and the long run using a series of short-run demand curves intersected by the long-run demand curve, as illustrated in Figure 4.4. Each of the short-run demand curves (D_{S1} to D_{S5}) is rather inelastic. Assume the original position is point a , which represents a price of P_1 and a quantity of Q_1 . If the price rises to P_2 , consumers will, in the short run, decrease the quantity demanded to point b , at quantity Q_2 , a relatively small difference in quantity.

Figure 4.4 Short-Run Versus Long-Run Elasticity



¹⁰A very interesting analysis of short- and long-run elasticity effects on the economic power of labor unions can be found in Milton Friedman, *Price Theory: A Provisional Text*, Hawthorne, NY: Aldine, 1962, pp. 155–59.

As time passes, during which consumers adjust to using substitutes, a new short-run demand curve, D_2 , will result, and demand will take place at point c , at Q_3 , which represents a much larger decrease in quantity. Thus, we can connect points a and c to illustrate the change in quantity demanded in the long run. Then, for constantly increasing prices, we can generate new short-run demand curves, D_{S3} , D_{S4} , and D_{S5} , and connect points d , e , and f to create a long-run demand curve.

In other words, in the short run, price changes increase or decrease the quantity demanded very little, up or down each short-run curve. However, over time, adjustment permits movement to another short-run curve, and a long-run demand curve is created. The long-run demand curve formed from a point on each short-run curve is obviously far more elastic.

The lengths of the short run and the long run depend on how quickly an adjustment can be made. In the case of heating oil, the long run represented several years. But let us return to the case of Henry Caulfield. Once Henry cuts his price (if he decides to do so), the news will probably spread quite quickly around the community. The adjustment from the short run to the long run will probably be a matter of days, or weeks at most.



Module 4B
Module 4C

Demand Elasticity and Revenue

Demand elasticity in itself is a rather interesting concept. However, if all it meant was the responsiveness of quantity to price change, it could be easily dismissed. But there is an aspect of demand elasticity that is terribly important to Henry Caulfield or any other businessperson in the throes of a pricing decision (in either direction).

There is a relationship between the price elasticity of demand and revenue received. A decrease in price would decrease revenue if nothing else were to happen. But because demand curves tend to be downward sloping, a decrease in price will increase the quantity purchased, and this will increase receipts. Which of the two tendencies is stronger? Remember that elasticity is defined as the percentage change in quantity divided by the percentage change in price. If the former is larger (and, therefore, the coefficient will be greater than 1 in absolute terms), then the quantity effect is stronger and will more than offset the opposite price effect.

What does that entail for revenue? If price decreases and, in percentage terms, quantity rises more than price has dropped, then total revenue will increase. We summarize the rules describing the relationship between elasticity and total revenue (TR) in Table 4.2.

Let us return to the example of a straight-line demand curve and see what happens to revenue. You will remember that elasticity on such a curve decreases as we move down and to the right. Total revenue and arc elasticity at each price interval are calculated in Table 4.3. Figures 4.5 and 4.6 show graphically the relationship between elasticity and revenue. It is obvious that as price decreases, revenue rises when demand is elastic, falls when it is inelastic, and reaches its peak (i.e., is level) when elasticity of demand equals -1.

Table 4.2 The Relationship Between Price Elasticity and Total Revenue (TR)

	Demand		
	Elastic	Unitary Elastic	Inelastic
Price increase	TR↓	$\overline{\text{TR}}$	TR↑
Price decrease	TR↑	$\overline{\text{TR}}$	TR↓

Table 4.3 Demand Schedule Showing Total Revenue and Elasticity Values

Price	Quantity	Elasticity		Total Revenue
		Arc, E_p	Point, ϵ_p	
\$18	0		$-\infty$	\$0
17	2	-35.00	-17.00	34
16	4	-11.00	-8.00	64
15	6	-6.20	-5.00	90
14	8	-4.10	-3.50	112
13	10	-3.00	-2.60	130
12	12	-2.30	-2.00	144
11	14	-1.77	-1.57	154
10	16	-1.40	-1.25	160
9	18	-1.12	-1.00	162
8	20	-0.89	-0.80	160
7	22	-0.71	-0.64	154
6	24	-0.57	-0.50	144
5	26	-0.44	-0.38	130
4	28	-0.33	-0.29	112
3	30	-0.24	-0.20	90
2	32	-0.16	-0.13	64
1	34	-0.09	-0.06	34
0	36	-0.03	0	0

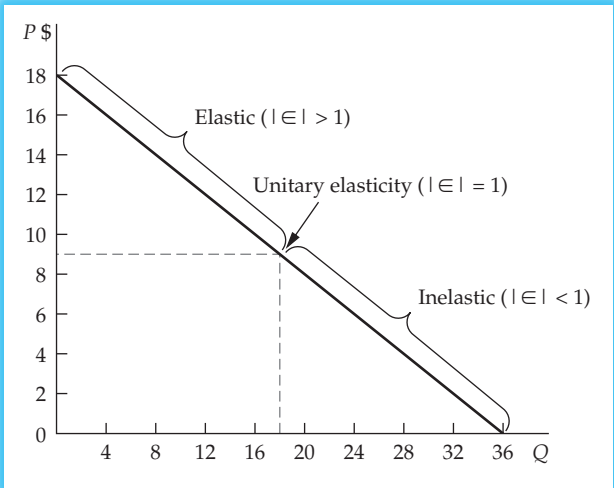
At this point, we can formally introduce a term we use a great deal throughout this book: **marginal revenue (MR)**. This concept can be defined as the change in total revenue as quantity changes by one unit.

Arc marginal revenue, $MR = \Delta TR / \Delta Q$ for discrete change in Q

This can also be defined at a point using derivatives as

Point marginal revenue, $MR = dTR / dQ$

Figure 4.5 The Elasticity–Demand Relationship



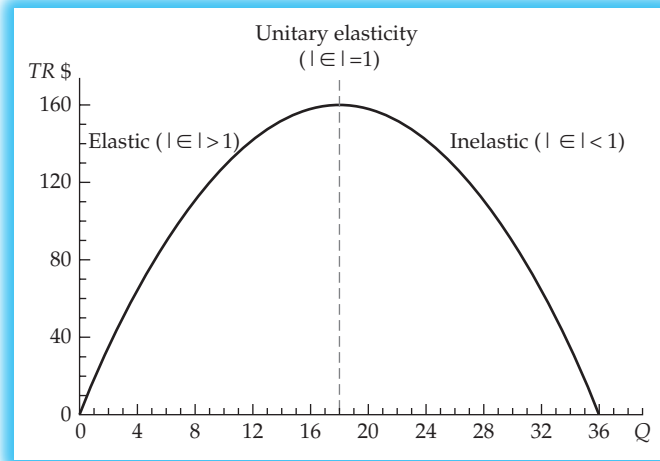


Figure 4.6 The Effect of Elasticity on Total Revenue

Now we can add to the previous table a *marginal revenue* column. This is shown in Table 4.4. Marginal revenue is positive as total revenue rises (and the demand curve is elastic). When total revenue reaches its peak (elasticity equals 1), marginal revenue reaches zero. This is exactly true when using point marginal revenue and point elasticity, and approximately true when using arc marginal revenue and arc elasticity.

Figure 4.7 shows the mathematical and graphical relationship between the demand curve and marginal revenue (MR). It turns out that when the demand curve is described by a straight line, the marginal revenue curve is twice as steep as the demand curve. Under these circumstances, the marginal revenue curve can be drawn by bisecting the distance between the *Y*-axis (vertical axis) and the demand curve. Of course,

Table 4.4 Demand Schedule Showing Marginal Revenue and Elasticity

Price	Quantity	Elasticity		Total Revenue	Marginal Revenue	
		Arc, E_P	Point, ϵ_P		Arc	Point
\$14	8		-3.5	\$112		\$10
13	10	-3.0	-2.6	130	\$9	8
12	12	-2.3	-2.0	144	7	6
11	14	-1.8	-1.6	154	5	4
10	16	-1.4	-1.3	160	3	2
9	18	-1.1	-1.0	162	1	0
8	20	-0.89	-0.80	160	-1	-2
7	22	-0.71	-0.64	154	-3	-4
6	24	-0.57	-0.50	144	-5	-6
5	26	-0.44	-0.38	130	-7	-8
4	28	-0.33	-0.29	112	-9	-10

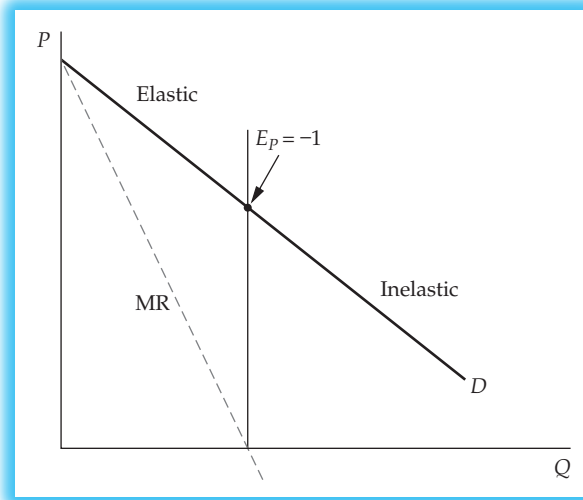


Figure 4.7 The Relationship between Demand and Marginal Revenue

at the point where marginal revenue crosses the X -axis (horizontal axis), the demand curve is unitarily elastic (and total revenue reaches its maximum).¹¹

All of this is going through Henry's mind as he is deciding how to counteract his competition. For him to benefit from decreasing his price, the demand curve for soft drinks from his store must be elastic. A price cut leading to a decrease in revenue would be self-defeating (or even disastrous). But this is not his only concern. If Henry is a profit maximizer, then it is profit, not revenue, that concerns him. If the demand for his product is elastic, revenue will increase. But as he sells more units, his total cost will, of course, rise. Will the increase in revenue more than offset the additional cost? That is the question uppermost in Henry's mind. Unfortunately, we are not ready to answer at this point. We first have to study production functions and cost functions, and then link demand and cost in Chapters 8 and 9.

The Mathematics of Elasticity and Revenue

As stated in the discussion of arc elasticity, elasticity changes along the demand curve (except in the case of constant elasticity discussed previously). So does point elasticity, of course. Thus, the point at which elasticity is measured must be specified. For example, the point elasticity at $Q = 10$ and $P = 13$ is

$$-2 \times 13 \div 10 = -2.6$$

In Table 4.3, the arc elasticity was shown to be -3 in the interval between $P = 13$ and $P = 14$ and -2.3 in the next interval. Point elasticity, as should be expected, is somewhere between the two arc elasticities. At $Q = 20$ and $P = 8$, the result is

$$-2 \times 8 \div 20 = -0.8$$

We are now in the inelastic section of the demand curve.

¹¹In mathematical terms, the path from the linear demand curve to the marginal revenue curve can be traced as follows:

$$\begin{aligned} \text{Inverse demand curve:} & \quad P = a - bQ \\ \text{Total revenue:} & \quad PQ = aQ - bQ^2 \\ \text{Marginal revenue:} & \quad dTR/dQ = a - 2bQ \end{aligned}$$

The Elasticity Excel App on the Companion Website continues the discussion of the relation between elasticity, marginal revenue and total revenue along a linear demand curve and allows you to create your own version of the tables and figures provided in the text.



We also mentioned the use of calculus in reference to point marginal revenue, which was defined as dTR/dQ . For this purpose, we reverse our demand function so that quantity becomes the independent variable.

$$\begin{aligned} \text{Inverse demand:} & \quad P = 18 - 0.5Q \\ \text{Total revenue:} & \quad TR = PQ = 18Q - 0.5Q^2 \\ \text{Marginal revenue:} & \quad MR = dTR/dQ = 18 - (1 \times Q) \end{aligned}$$

At $Q = 10$ and $TR = 130$, marginal revenue is

$$18 - (1 \times 10) = 8$$

Again, because we are measuring TR at a point rather than over an interval, the result ($MR = 8$) will be somewhat different from the arc marginal revenue obtained in Table 4.4, where discrete differences were used in the calculation.

At $Q = 20$, $MR = 18 - (1 \times 20) = -2$. We are now in the area of negative marginal revenue (relative inelasticity on the demand curve), or at a quantity higher than the one that produces maximum revenue.

If we want to find the point where revenue is maximized, we look for the point at which $MR = 0$. Thus,

$$\begin{aligned} 18 - (1 \times Q) &= 0 \\ 18 &= Q \end{aligned}$$

Empirical Elasticities

In Chapter 5, we explain how economists estimate demand curves and elasticities from industry and product data. However, it may be of interest at this point to reinforce the meaning of price elasticity by mentioning briefly the results of some studies published in recent years:

A study of the demand for coffee estimated the price elasticity to be -0.2 in the short run and -0.33 in the long run.

A study of the demand for kitchen and other household appliances stated that the elasticity was -0.63 .

Meals (excluding alcoholic beverages) purchased at restaurants have a high demand elasticity of -2.27 .

A study of air travel demand over the North Atlantic found price elasticity to be -1.2 . Furthermore, the elasticity for first-class travel was, as one would expect, considerably lower than that for economy travel, -0.4 and -1.8 , respectively. A recent study estimates the price elasticity of passenger travel for the U.S. airline industry to be a high -1.98 .

The price elasticity of beer has been estimated at -0.84 and of wine at -0.55 .

The price elasticity of white pan bread in Chicago appeared to be -0.69 , while for premium white pan bread elasticity was measured at 1.01 . This is consistent with the idea that there are more substitutes for premium white pan bread brands than for all white pan bread.

In the short run, demand elasticity for cigarettes in the United States appeared to be -0.4 , but it was higher, although still quite low, in the long run at -0.6 . A study of the effect of tax increases on the consumption of cigarettes in Canada found price elasticity to be between -0.45 and -0.47 . This would indicate that the taxing authority would be able to increase its revenue when taxes are increased. Another study looked at the price elasticity of cigarettes in China and Russia and obtained extremely low elasticities, between 0 and -0.15 .

A study of wine imports into the United States estimated the price elasticity of Chilean wines to be about -1.6 , while the elasticity of French and Italian wines was less than -0.3 . The price elasticity of crude oil was calculated for 23 countries for the period of 1979 to 2000. In the short run, oil demand appears to be very price inelastic. The estimate for the United States was -0.06 . Long-run elasticities were found to be higher; for the United States, it was -0.45 . The averages for all 23 countries were -0.05 in the short run and -0.18 in the long run.

A survey of many studies dealing with gasoline demand found that in the short run elasticity is about -0.3 , while in the long run it ranges from -0.6 to -0.8 . Thus demand is price inelastic in both cases but much more so in the short run.

A recent study of elasticities for Internet services found the price elasticity for subscribers to be between -0.6 and -0.7 . These results were based on data from countries that are members of the Organization for Economic Cooperation and Development (OECD).

A study of gasoline demand for four low-income developing Asian countries estimated short-run price elasticities to be between -0.12 and -0.17 and long-run elasticities to be -1.18 . For seven middle-income countries the numbers were about -0.3 in the short run and -0.58 in the long run.

The price elasticity of demand for private care in nursing facilities in Texas was estimated to be -0.69 .

A study of demand for private and public higher education in the United States showed that price elasticity for private colleges was -1 while it was -1.7 for public colleges.

A British Columbia study estimated that a 10.0 percent increase in the minimum price of alcoholic beverages decreased consumption of all alcoholic drinks by 3.4 percent. For specific beverages, spirits and liqueurs decreased by 6.8 percent, wine by 8.9 percent, and beer by 1.7 percent. In all cases, the results exhibit demand to be price inelastic.

Finally, a review of a large number of studies of price elasticity for sixteen categories of food and beverages showed elasticities of less than 0. Among the groups were soft drinks (-0.79), fruit (-0.70), vegetables (-0.58), and eggs (-0.27). All categories showed demand to be price inelastic.¹²

¹²A. A. Okunade, "Functional Forms and Habit Effects in the U.S. Demand for Coffee," *Applied Economics*, 24, (1992), pp. 1203–12; H. S. Houthakker and L. D. Taylor, *Consumer Demand in the United States: Analysis and Projections*, 2nd ed., Cambridge, MA: Harvard University Press, 1970, pp. 63, 81; J. M. Cigliano, "Price and Income Elasticities for Airline Travel," *Business Economics*, September 1980, pp. 17–21; D. Heien and G. Pompelli, "The Demand for Alcoholic Beverages: Economic and Demographic Effects," *Southern Economic Journal*, January 1989, pp. 759–69; Bahram Adrangi and Kambiz Raffiee, "New Evidence on Fare and Income Elasticity of the U.S. Airline Industry," *Atlantic Economic Journal*, Vol. 28, Issue 4, December 2000, p. 493; Gregory J. Werden, "Expert Report in United States v. Interstate Bakeries Cop. and Continental Baking Co.," *International Journal of the Economics of Business*, July 2000, pp. 139–48; Ping Zhan, Corinne Husten, and Gary Giovino, "Effect of the Tobacco Price Support Program on Cigarette Consumption in the United States: An Updated Model," *American Journal of Public Health*, Vol. 90, Issue 5, May 2000, pp. 746–800; Jonathan Gruber, Anindya Sen, and Mark Stabile, "Estimating Price Elasticities When There Is Smuggling," *Journal of Health Economics*, Vol. 22, Issue 5, September 1, 2003, pp. 821ff; Peter M. Lance, John S. Akin, William H. Dow, and Chung-Ping Loh, "Is Cigarette Smoking Sensitive to Price? Evidence from Russia and China," *Journal of Health Economics*, Vol. 23, Issue 1, January 1, 2004, pp. 137ff; James L. Seale Jr., Mary A. Merchant, and Alberto Basso, "Imports versus Domestic Production: A Demand System Analysis of the U.S. Red Wine Market," *Review of Agricultural Economics*, Vol. 25, Issue 1, Spring–Summer 2003, pp. 187–202; John C. B. Cooper, *Price Elasticity of Demand for Crude Oil: Estimates for 23 Countries*, OPEC Review, March 2003; Daniel J. Graham and Stephen Glaister, "The Demand for Automobile Fuel," *Journal of Transport Economics and Policy*, Vol. 36, Part 1, January 2002, pp. 1–26; Rajeev K. Goel, Edward T. Hsieh, Michael A. Nelson and Rati Ram, "Demand Elasticities for Internet Services," *Applied Economics*, Vol. 38, Issue 9, May 20, 2006, pp. 75–80; Robert McRae, "Gasoline Demand in Developing Asian Countries," *Energy Journal*, 1994, Vol. 15, Issue 1, pp. 143–155; Kris Joseph Knox, Eric C. Blankmeyer, and J. R. Stutzman, "Private-Pay Demand for Nursing Facilities in a Market with Excess Capacity," *Atlantic Economic Journal*, 2006, 34, pp. 75–83; Mohamad Hamadeh and Roy Khoueir, "Demand Elasticities for Higher Education in the United States," *International Journal of Business and Economic Perspectives*, Vol. 5, Number 2, Fall 2010, pp. 60–67; Tim Stockwell, M. Christopher Auld, Jinhui Zhao, and Gina Martin, "Does Minimum Wage Pricing Reduce Alcohol Consumption?" *Addiction*, Vol. 107, Issue 5, May 2012, pp. 912–920; Tatiana Andreyeva, Michael W. Long and Kelly D. Brownell, "The Impact of Food Prices on Consumption: A Systematic Review of Research on the Price Elasticity of Demand for Food," *American Journal of Public Health*, Vol. 100, No. 2, February 2010, pp. 216–222.

THE CROSS-PRICE ELASTICITY OF DEMAND

The previous discussion dealt with the influence of a price change on the quantity demanded of the product subject to the price change. **Cross-price elasticity** deals with the impact (again, in percentage terms) on the quantity demanded of a particular product created by a price change in a related product (while everything else remains constant). What is the meaning of “related” products? In economics, we talk of two types of relationships: **substitute good** and **complementary good**.

In Henry Caulfield’s case, we are dealing with substitutes. The sodas sold by the new grocery store are substitutes for those sold by Henry. They probably are the same products (same brands) but are sold by different suppliers, and one supplier can be considered a substitute for the other. Of course, there are also substitutes on Henry’s own shelves—he stocks different brands of cola, for example.

Much of the time when we consider cross-price elasticity we are dealing with similar products (not just different brands of the same product) in a more general sense. Thus, chicken and beef can be considered to be substitutes; a change in the price of chicken will have an effect on the consumption of beef. Other instances of substitutes come to mind easily: coffee and energy drinks, Xbox and Playstation, butter and margarine, aluminum and steel, and glass and plastic.

Complements are products that are consumed or used together. Henry sells potato chips, pretzels, and other “munchies” that are consumed together with soft drinks. Other cases of complementary products are peanut butter and jam, ipods and earbuds, tennis rackets and tennis balls, and smartphones and apps. Henry also has found that when he offered substantial discounts on soft drinks, his sales of gasoline increased. From Henry’s perspective, the soft drink cross-price elasticity of demand for gasoline suggests that soft drinks are complementary to gasoline.

The definition of cross-price elasticity is a measure of the percentage change in quantity demanded of product A resulting from a 1 percent change in the price of product B. The general equation can be written as

$$E_{A,B} = \frac{\Delta Q_A}{Q_A} \div \frac{\Delta P_B}{P_B}$$

Again, we run into a little problem regarding the denominator in this expression, and arc elasticity comes to the rescue:¹³

$$E_{A,B} = \frac{(Q_{2A} - Q_{1A})}{(Q_{2A} + Q_{1A})/2} \div \frac{(P_{2B} - P_{1B})}{(P_{2B} + P_{1B})/2}$$

What about cross-elasticity coefficients? First, let us look at the sign. A decrease in the price of the supermarket’s soft drinks will cause the quantity of soft drinks sold by Caulfield to decrease. Of course, if the supermarket raises prices, Caulfield’s sales will rise. Thus, the sign of cross-elasticity for substitutes is positive. In contrast, the

¹³The following equation, obtained by arithmetic manipulation, may be easier for some readers:

$$E_{A,B} = \frac{(Q_{2A} - Q_{1A})}{(P_{2B} - P_{1B})} \times \frac{(P_{2B} + P_{1B})/2}{(Q_{2A} + Q_{1A})/2}$$

Or, if point elasticity is of interest, the use of calculus gives

$$\epsilon_{A,B} = \frac{\partial Q_A}{\partial P_B} \times \frac{P_B}{Q_A}$$

The calculations would proceed similarly to those for demand elasticity.

coefficient sign for cross-elasticity of complements is negative. For instance, a decrease in Henry's soft-drink prices is likely to increase his snack food sales.

To measure the strength of the elasticity coefficient, we employ a more arbitrary definition than for demand elasticity. As a rule of thumb in business, two products are considered good substitutes or complements when the coefficient is larger than 0.5 (in absolute terms, because the coefficient for complements is negative).

Empirical Elasticities

Again, it should be useful to briefly mention some study results:

A study of the residential demand for electric energy found the cross-elasticity with respect to prices of gas energy to be low, about +0.13.

Aluminum's cross-elasticity of demand with respect to prices of steel was estimated at about +2.0, and even somewhat higher with respect to copper.

The cross-elasticity of demand for beef with respect to pork prices was calculated to be about +0.25. With respect to prices of chicken, it was about +0.12. Both numbers indicate that the products are substitutes, but in this study, the elasticity coefficients were relatively low.

The cross-elasticity for domestic and imported cigarettes in Taiwan is a positive 2.78, indicating that they are substitutes.

A study of demand elasticity of kerosene in Indonesia estimated its cross-elasticity with electricity was only 0.097 in the short run and 0.261 in the long run. The authors ascribe the low value (during the study period 1957–1992) to the lack of access to electricity in a majority of households.

A study of red wine production in the United States and imports from other countries estimated several cross-price elasticities. Price increases in several of the imported wines have a positive effect on the quantity demanded of U.S. red wines. Red wines from Chile appeared to have a significant effect; a 1 percent price rise increased the quantity demanded for U.S. wines by 1.2 percent. Wines from other countries also had a positive, but smaller, effect: the cross-price elasticity for Australian wines was 0.4, while it was 0.3 for French wines and only 0.04 for wines from Spain.

A study of the impact of used books on new books at Amazon.com showed that used books are not a particularly good substitute for new books. The cross-price elasticity was only 0.088.

Another recent study estimates cross-price elasticities for various types of meats in Jordan. It concluded that beef and poultry are substitutes for mutton, with coefficients of 0.66 and 1.88, respectively. However, the cross-price elasticity for fish was only 0.33.¹⁴

INCOME ELASTICITY

Before the arrival of his competitors, Henry Caulfield saw his sales grow, not only as the number of households in the area increased, but also as the household income level in the area rose. This represents quantity of sales as a function of (i.e., influenced by)

¹⁴R. Halvorsen, "Residential Demand for Electric Energy," *Review of Economics and Statistics*, 57 (February 1975), pp. 12–18; Merton J. Peck, *Market Control in the Aluminum Industry*, Cambridge, MA: Harvard University Press, 1961, pp. 31–34; Daniel B. Suits, "Agriculture," in Walter Adams, *The Structure of American Industry*, 8th ed., New York: Macmillan, 1990, p. 11; Chee-Ruey Hsieh, Teh Wei Hu, and Chien-Fue Jeff Lin, "The Demand for Cigarettes in Taiwan: Domestic Versus Imported Cigarettes," *Contemporary Economic Policy*, Vol. 17, No. 2, April 1999, pp. 223–234; Rajindar K. Koshal and Manjulika Koshal, "Demand for Kerosene in Developing Countries: A Case of Indonesia," *Journal of Asian Economics*, Vol. 10, Issue 2, Summer 1999, pp. 329–336; James L. Seale Jr., Mary A. Marchant, and Alberto Basso, "Imported versus Domestic Production: A Demand System Analysis of the U.S. Red Wine Market," *Review of Agricultural Economics*, Vol. 25, Issue 1, Spring–Summer 2003, pp. 187–202; Anindya Ghose, Michael D. Smith, and Rahul Telang, "Internet Exchanges for Used Books, and Empirical Analysis of Product Cannibalization and Welfare Impact," *Information Systems Research*, Vol. 17, No. 1, March 2006, pp. 3–19; Amer S. Jabarin, "Estimation of Meat Demand System in Jordan," *International Journal of Consumer Studies*, Vol. 29, No. 3, May 2006, pp. 232–238.

consumers' income. As a measure of the sensitivity of this relationship, economists use the term **income elasticity** of demand. The general expression for this elasticity is

$$E_Y = \% \Delta Q \div \% \Delta Y$$

where Y represents income.¹⁵ The definition of income elasticity is a measure of the percentage change in quantity consumed resulting from a 1 percent change in income.

As before, we turn to arc elasticity for the actual calculation of income elasticity:¹⁶

$$E_Y = \frac{(Q_2 - Q_1)}{(Q_2 + Q_1)/2} \div \frac{(Y_2 - Y_1)}{(Y_2 + Y_1)/2}$$

Income elasticity can be either positive or negative. For most products, one would expect income elasticity to be positive. After all, given a rise in income, a person will spend more. When income elasticity is positive, we say that a good is normal. When it is zero, we say that demand is independent of income, and when it is negative, we call it an inferior good.

If the quantity bought of a product were to decrease absolutely as the result of an increase in income, that product is called an inferior good. Although such a result may at first seem implausible, a little reflection shows that such a condition may exist. Some products are demanded by consumers whose incomes are low. But as incomes rise, they shift consumption to goods more commensurate with their new economic status. What types of products would these be? The usual examples used by economists are potatoes, pork and beans, and canned luncheon meat. We refer to these commodities as **inferior goods**.

Income elasticity +1 represents another dividing line. As income rises, people can increase consumption of products (and services) proportionally, less than proportionally, or more than proportionally to the income rise. If the expenditure on product A goes up by 10 percent when income goes up by 10 percent, then the income elasticity equals 1. That is, the proportion of the consumer's income spent on this commodity remains the same before and after the change in income.¹⁷ Suppose a consumer's annual income is \$30,000 and spending on clothing is \$2,700 per year. If this person's annual income rises by 10 percent to \$33,000, and he or she then spends \$2,970 annually on clothing—also a 10 percent increase—the proportion of total income spent on clothing remains at 9 percent.

If the income elasticity coefficient is greater or less than 1, the fraction of income spent on the product in question changes more or less than proportionally with income. Products whose income elasticities exceed +1, taking larger portions of consumers' income as incomes increase, are often referred to as "superior" goods or as luxuries.

¹⁵Again, in terms of calculus, this expression could be written as

$$\epsilon_y = \partial Q / \partial Y \times Y \div Q$$

This, of course, expresses the point elasticity. Also, as the reader can see partial derivatives have been used. As in all other cases of elasticity, the assumption is that only the effect of income on quantity is being measured, with all other possible variables in the demand relationship (e.g., price, price of related products, interest rates, advertising) held constant.

¹⁶As in the case of the other elasticities, the equation can be rewritten in several different forms. The student can select the equation that is most convenient:

$$E_Y = [(Q_2 - Q_1) \div (Q_2 + Q_1)] \div [(Y_2 - Y_1) \div (Y_2 + Y_1)]$$

$$E_Y = [(Q_2 - Q_1) \div (Q_2 + Q_1)] \times [(Y_2 + Y_1) \div (Y_2 - Y_1)]$$

$$E_Y = [(Q_2 - Q_1) \div (Y_2 - Y_1)] \times [(Y_2 + Y_1) \div (Q_2 + Q_1)]$$

¹⁷Remember that, by definition, the price of the product remains the same. Thus it does not matter whether we measure an increase in quantity of the product or in expenditures on the product.

Again, a brief review of empirical studies that have estimated income elasticities illustrates this concept:

Short-run income elasticity for food expenditure has been estimated to be about 0.5 and the elasticity of restaurant meals 1.6. The results show that as incomes rise, spending for food eaten at home increases at a slower rate than income; thus takes up a smaller portion of income. In contrast, the expenditure on restaurant meals rises substantially more rapidly as incomes rise, thus becoming a higher proportion of income. The short-run income elasticity for jewelry and watches appeared to be 1.0; however, the elasticity in the long run was estimated at 1.6. Apparently, consumers take some time to adjust their demand.

The income elasticity for air travel between the United States and Europe was a relatively high 1.9. Recently, the income elasticity for passenger travel with the U.S. airline industry has been estimated at unity.

A study of food demand in Spain, which divided food into six categories, found that only meat had an income elasticity of greater than one, 1.54, thus classifying it as a superior good. The elasticity of fish was 0.81, while fats and oils were the lowest at 0.35. A survey of studies of demand for gasoline found that short-run income elasticity tends to range between 0.35 and 0.55, while in the long run demand is more elastic, in the range between 1.1 and 1.3.

Income elasticities of food expenditures of the Hispanic population in the United States were estimated for three categories: total food, food eaten at home, and food eaten away from home. The results all showed elasticities of less than 1—the proportion of income spent on food decreases as income increases. For the three categories, income elasticities were 0.29, 0.21, and 0.49, respectively. As expected, food eaten away from home had the highest elasticity.

The coefficients of income elasticity for passenger railway transportation in Slovenia turned out to be less than 1, in the range from 0.43 to 0.93.

The income elasticity for subscribers to the Internet was estimated to be in excess of 1, thus making it a superior commodity. This study was based on data from OECD countries. Another study examined the income elasticities for gasoline demand in 11 Asian countries. Separate elasticities were calculated for 7 middle-income countries and 4 low-income countries. In the middle-income countries, short-run elasticities were found to be between 0.57 and 0.67, and long run elasticities were 1.7. For low-income countries the results were 0.74 to 0.79 and 1.18, respectively.

The income elasticity for housing in Spain was estimated to be between 0.7 and 0.95. This was lower than the results of some previous studies.

The income elasticity for private college education in the United States was 0.84, while it was 1.6 for public colleges.¹⁸

¹⁸Houthakker and Taylor, *Consumer Demand in the United States*, pp. 62–63, 72. Actually, this study measured the relationship between spending on food and total expenditures, rather than income. Because the proportion of total expenses to income is relatively stable, this substitution does not significantly change results. J. M. Cigliano, “Price and Income Elasticities for Airline Travel,” *Business Economics*, September 1980, pp. 17–21; Bahram Adrangi and Kambiz Raffiee, “New Evidence on Fare and Income Elasticity of the U.S. Airline Industry,” *Atlantic Economic Journal*, Vol. 28, Issue 4, December 2000, p. 493; A. Garcia and J. M. Gil, “Spanish Food Demand: A Dynamic Approach,” *Applied Economics*, Vol. 30, Issue 10, October 1998, pp. 1399–1405; Daniel J. Graham and Stephen Glaister, “The Demand for Automobile Fuel,” *Journal of Transport Economics and Policy*, Vol. 36, Part 1, January 2002, pp. 1–26; Bruno A. Lanfranco, Glen C. W. Ames, and Chung I. Huang, “Food Expenditure Patterns of the Hispanic Population in the United States,” *Agribusiness*, Vol. 18, Issue 2, 2002, pp. 197–211; Jani Bekö, “Some Evidence on Elasticities of Demand for Services of Public Railway Passenger Transportation in Slovenia,” *Eastern European Economics*, Vol. 42, No. 2, March–April 2004, pp. 63–85; Rajeev K. Hsieh, Edward T. Nelson, and Michael A. Ram, “Demand Elasticities for Internet Services,” *Applied Economics*, Volume 38, Issue 9; Robert McRae, “Gasoline Demand in Developing Asian Countries,” *Energy Journal*, 1994, Vol. 15, issue 1, pp. 143–55; Daniel Fernández-Kranz and Mark T. Hon, “A Cross-Section Analysis of the Income Elasticity of Housing Demand in Spain,” *Journal of Real Estate Finance and Economics*, 2006, 32, pp. 449–470; Mohamad Hamadeh and Roy Khoueiri, “Demand Elasticity for Higher Education in the United States,” *International Journal of Business and Economics Perspectives*, Vol. 5, Number 2, Fall 2010, pp. 60–67.

OTHER ELASTICITY MEASURES

We have covered the three most common elasticity measures, but there are others. Elasticity is encountered every time a change in some variable affects quantities. For instance, one thing that Henry could do to counteract his competition is to advertise his products. He might speculate how an increase in advertising expenses would affect his total sales. Thus, **advertising elasticity** can be defined as the percentage change in quantity relative to a 1 percent change in advertising expenses. Advertising elasticity has been studied in economic literature. A few years ago, an investigation was conducted into the effects of advertising for beef, poultry, and pork. The results of both generic and brand advertising were measured. Generic advertising did not appear to have any statistically significant results. In contrast, brand advertising had significant, although rather small effects on the consumption of these meats. Thus, a 10 percent increase in advertising increased the consumption of beef by 0.07 percent, the consumption of pork by 0.33 percent, and of poultry by 0.5 percent.

Another study was conducted to examine the effect of advertising on milk consumption in five cities in New York State. The elasticities ranged from 0.014 to 0.088; again, this effect appeared to be small but positive.¹⁹

Another variable that could have a significant impact on demand—particularly for durable goods—is the interest rate. No one would deny that shifts in mortgage interest rates can cause significant changes in the demand for residential (or nonresidential) construction. Also, the special loan rates offered by automobile manufacturers to customers beginning in the 1980s appear to stimulate car sales, and sales tend to sag when the low rates are terminated.

Elasticity could also be calculated in relation to population size. What is the effect on sales of changes in population? For instance, we could calculate the elasticity of demand for baby carriages as a result of population increases due to the baby boom (and the children of the baby boomers). Or we could investigate the effect of a change in the number of adults (population older than the age of 18) on the annual purchases of automobiles (again, as always, holding all other variables constant). Of course, in Henry Caulfield's case, changes in his community's population will affect his sales. The degree to which sales will be affected is measured by elasticity.

These are just a few examples of possible elasticity calculations. In Chapter 5, when we discuss demand estimation, we see that economists use many variables to explain changes in demand. Although price elasticity, cross-price elasticity, and income elasticity are the ones most frequently measured, elasticities can be obtained for a large variety of variables.

ELASTICITY OF SUPPLY

Before we close this chapter, it is useful to devote a little space to the **price elasticity of supply**. The price elasticity of supply measures the percentage change in quantity supplied as a result of a 1 percent change in price. In other words, this elasticity is a measure of the responsiveness of quantities produced by suppliers to a change in price. In Chapter 3, a supply schedule and a supply curve are developed, and the curve slopes upward and to the right. Thus, the arc coefficient of supply elasticity,

$$E_S = \frac{(Q_2 - Q_1)}{(Q_2 + Q_1)/2} \div \frac{(P_2 - P_1)}{(P_2 + P_1)/2}$$

is a positive number: quantity and price move in the same direction.

¹⁹Gary W. Brewster and Ted C. Schroeder, "The Impacts of Brand and Generic Advertising on Meat Demand," *American Journal of Agricultural Economics*, Vol. 77, November 1995, pp. 969–79; John Lenz, Harry A. Kaiser, and Chanjin Chung, "Economic Analysis of Generic Milk Advertising Impacts on Markets in New York State," *Agribusiness*, Vol. 14, No. 1, January/February 1998, pp. 73–83.

The interpretation of the coefficient is the same as for the case of demand elasticity. The higher the coefficient, the more quantity supplied will change (in percentage terms) in response to a change in price.

Again, as in the case of demand elasticity, it is important for a manager to know how supply elasticity affects price and quantity when demand changes. When the supply curve is more elastic, the effect of a change in demand will be greater on quantity than on price of the product. In contrast, with a supply curve of low elasticity, a change in demand will have a greater effect on price than quantity.

GLOBAL APPLICATION: PRICE ELASTICITIES IN ASIA

In our examples of price elasticity studies we mentioned several that were done in countries other than the United States. Here we describe briefly a group of studies spanning a large number of consumer products conducted by ACNielsen in Asia during 2001–2002.

The study looked at elasticities in various product categories, by brands and among countries. The numbers we quote below are in absolute terms. When overall categories were studied, price elasticities were relatively low, with 70 percent being less than 2.0 and 35 percent below 1.0. The average was 1.5. However, when individual brands were examined, price elasticities were considerably higher. The average for Asian countries was 2.3. For example, food and beverage brands had price elasticity of 3.0 and household and personal care items' elasticity was 2.2. These findings conform to our previous discussion of determinants of elasticity: the broader the definition of a commodity, the lower will be the price elasticity.

There appear to be substantial differences in price elasticities among the different countries of Asia. For instance, consumers in Malaysia and Hong Kong were considerably more responsive to price changes—well above the Asian average of 2.3—while the Philippines and Korea exhibited elasticities of less than 1.5.²⁰

Another study examined price and income elasticities for imports and exports in five Asian countries—India, Japan, Philippines, Sri Lanka, and Thailand. The author used long-term time series to make his estimates. (You will learn more about time series regressions in Chapter 5.) He found the following elasticities:

	Import		Export	
	Price	Income	Price	Income
India	-0.51	-0.11	-0.55	0.45
Japan	-0.91	0.84	-0.80	2.84
Philippines	-0.17	0.57	-1.06	0.89
Sri Lanka	-0.48	-0.39	0.48	0.90
Thailand	-0.20	0.90	-3.76	0.82

The implications of these findings are as follows:

1. If imports are price inelastic (as they are in almost all cases), a rise in import prices will lead to an increase in the import bill.
2. If imports are income elastic, an increase in incomes will entail a more than proportionate increase in imports.

²⁰*Pricing for Success*, www.2.acnielsen.com/pubs/2003_q3_ap_pricing.html.

3. If exports are price inelastic, export earnings will rise as prices increase.
4. If exports are income elastic, an increase in world incomes will lead to a greater than proportionate increase in exports.²¹

The Solution

Henry Caulfield is no stranger to the economic concept of price elasticity. He has a bachelor's degree in business administration and was doing well as the regional manager of a large supermarket chain when he decided to leave his job and open his own business. Indeed, it was his understanding of price elasticity that prompted him to reduce the soft drink prices as a way of competing against the two new stores in his area. When he had offered special discounts on soft drinks in the past, he noticed that people were very responsive. In fact, Henry had kept a record of the relationship between price and sales, a part of which is shown in Table 4.5. The "special" price was offered as part of the store's "Fourth of July Celebration" sale.

Table 4.5 Sales Data for 2-Liter Bottles of Soft Drinks

Average Price	Average Weekly Sales	Total Revenue
Regular Price: \$1.89	1,050	\$1,985
Special Price: \$.89	2,450	2,181

The data indicate an elastic demand for soft drinks at Henry's store. When demand is price elastic, a reduction in price causes total revenue to increase. This was exactly what had happened when Henry had offered his "Fourth of July Celebration" special. He was now puzzled because the permanent price reduction did not seem to be having the same positive effect on his total revenue.

Then, in a flash, it dawned on him. One of the most important aspects of demand elasticity—and, for that matter, of any aspect of economic analysis—is the assumption that certain factors are held constant in the examination of the impact of one variable on another. In this case, it was assumed that other factors besides price did not have an impact (or at least not much of an impact) on quantity when Henry had offered the special holiday price for his soft drink. What other factors besides price might now be taken into account?

To begin with, last summer Henry did not have any close competitors. Therefore, when he offered his discount, there was no other store nearby to match this price reduction. Obviously, the two new stores were not going to stand by idly watching potential customers go to Henry because he had the lowest price for soft drinks. Therefore, the demand for soft drinks at Henry's store was much less elastic than he believed because he was unable to take away their business. To make matters worse, this "price war" among the three stores might have actually reduced their total soft drink revenues. This is because when all three stores dropped their price, they might well have brought the quantity demanded into the inelastic range of their combined demand curves. (We assume here that the three stores constitute the entire local market for soft drinks.)

Regardless of the reaction of his competitors and the possible impact that their price cuts might have had on the degree of price elasticity, there was one simple fact that Henry had completely overlooked. Last year's discount took place in summer, a time when the seasonal demand for this product increases anyway. Thus, when Henry cut the price, the demand for his product had already started to increase and his increased revenue may have been caused by the fact that during this time the demand curve was moving to the right.

One final factor had to be considered. In the past, his discounts on soft drinks were "specials" and, therefore, temporary in nature. Consumers knew that they had to take advantage of these specials during a

(continued)

²¹Dipendra Sinha, "A Note on Trade Elasticities in Asian Countries," *The International Trade Journal*, 15:2, pp. 221–237.

(continued)

designated period. Because they now realized that the price of soft drinks in Henry's store was permanently lowered, they were in no hurry to buy the product. In other words, Henry had failed to take "future expectations" into account.

Thus, to be able to measure elasticity, Henry would have to separate the effects on unit sales of price from all the other nonprice determinants of demand. Because he had not done this, he had overestimated the degree of responsiveness by his customers to his price reduction. As a result, unfortunately, the reduction in soft drink prices did not provide a solution for Henry. But at least he now understood why it did not.²² In addition, this analysis reminded Henry never to take for granted that "other factors remain constant." In the real world, conditions are changing all the time, and it is important to factor these changes into the analysis. As a small consolation, Henry realized that the entry of additional suppliers into the market was all part of the economics of running a successful business. After all, if people did not think he was making any money, they would probably not be as willing to start a competing enterprise.

SUMMARY

This chapter deals with the important concept of elasticity. In the most general terms, elasticity is defined as the sensitivity of one variable to another or, more specifically, the percentage change in one variable caused by a 1 percent change in another. Several forms of elasticity connected with the demand curve were discussed.

The first was price elasticity of demand: the percentage change in the quantity demanded of a product caused by a percentage change in its own price. Because demand curves slope downward and to the right, the sign of price elasticity is negative. If the coefficient is less than -1 (or greater than 1 in absolute terms), demand is said to be elastic. In contrast, the elasticity coefficient can indicate inelasticity or unitary elasticity.

Elasticity is also tied to total revenue. When demand is elastic, revenue rises as quantity demanded increases; revenue reaches its peak at the point of unitary elasticity and descends as quantity rises on the demand curve's inelastic sector. From the concept of revenue, we develop marginal revenue as the change in revenue when quantity changes by one unit. Marginal revenue is positive at quantities where demand is elastic and becomes negative when the demand curve becomes inelastic.

Next, we explain cross-price elasticity, the relationship between the demand for one product and the price of another. Products can be substitutes, and their cross-price elasticity is then positive; cross-elasticity is negative for products that are complements.

The third major elasticity concept, income elasticity, measures the sensitivity of demand for a product to changes in the income of the population. Goods and services are defined as superior (or luxuries), normal, and inferior, depending on the responsiveness of spending on a product relative to percentage changes in income.

The examples calculated in the chapter use the method of arc elasticity, which measures changes in both variables over discrete intervals, as well as point elasticity, which deals with change over an infinitely small interval and consequently may require knowledge of elementary calculus.

²²Our example may help you understand an apparent paradox in the pricing of soft drinks in supermarkets. Very often, substantial discounts on soft drinks are offered at all supermarkets during the summer (i.e., "summer specials"). These discounts either may be offered by the soft drink companies to the supermarkets, which then pass them on to customers, or they may be initiated by the supermarkets themselves. Why should they do this at a time when demand is high? After all, economic theory states that an increase in demand causes prices to rise, other factors held constant. What probably happens is that one of the major soft drink companies decides to take market share away from the other major producers by cutting its price. The others quickly follow. The same is true among supermarkets. These price wars can and do occur at any time during the year. It is just that a "summer special" is a good reason to have a sale.

Several other subtopics appear in this chapter:

- Other elasticities, such as advertising and interest elasticity.
- Derived demand, which is the demand for inputs to a final product, and the price elasticity of derived demand.
- Supply elasticity, the measure of the sensitivity of quantities produced to the price charged by the producers.

In Chapter 5, which discusses methods of estimating demand functions, elasticity concepts are employed again, and they reappear in various guises in many of the chapters that follow.

IMPORTANT CONCEPTS

Advertising elasticity. The percentage change in quantity demanded caused by a 1 percent change in advertising expenses. (p. 90)

Arc elasticity. Elasticity that is measured over a discrete interval of a demand (or a supply) curve. (p. 70)

Complementary good. A product consumed in conjunction with another. Two goods are complementary if the quantity demanded of one increases when the price of the other decreases. (p. 86)

Cross-price elasticity. The percentage change in the quantity consumed of one product as a result of a 1 percent change in the price of a related product. (p. 86)

Derived demand. The demand for products or factors that are not directly consumed but go into the production of a final product. The demand for such a product or factor exists because there is demand for the final product. (p. 78)

Elasticity. The percentage change in one variable divided by the percentage change in the other variable. (p. 69)

Income elasticity. The percentage change in quantity demanded caused by a 1 percent change in income. (p. 88)

Inferior good. A product whose consumption decreases as income increases (i.e., its income elasticity is negative). (p. 88)

Marginal revenue. The change in total revenue resulting from changing quantity by one unit. (p. 81)

Point elasticity. Elasticity measured at a given point of a demand (or a supply) curve. (p. 72)

Price elasticity of demand. The percentage change in quantity demanded caused by a 1 percent change in price. (p. 69)

Price elasticity of supply. The percentage change in quantity supplied as a result of a 1 percent change in price. (p. 90)

Substitute good. A product that is similar to another and can be consumed in place of it. Two goods are substitutes if the quantity consumed of one increases when the price of the other increases. (p. 86)

QUESTIONS

1. State the general meaning of *elasticity* as it applies to economics. Define the *price elasticity of demand*.
2. Explain the difference between *point elasticity* and *arc elasticity*. What problem can arise in the calculation of the latter, and how is it usually dealt with? In actual business situations, would you expect arc elasticity to be the more useful concept? Why or why not?
3. It has often been said that craft unions (electricians, carpenters, etc.) possess considerably greater power to raise wages than do industrial unions (automobile workers, steel workers, etc.). How would you explain this phenomenon in terms of demand elasticity?
4. Discuss the relative price elasticity of the following products:
 - a. Mayonnaise
 - b. A specific brand of mayonnaise
 - c. Chevrolet automobiles
 - d. Jaguar automobiles

- e. Washing machines
 - f. Air travel (vacation)
 - g. Beer
 - h. Diamond rings
5. What would you expect to happen to spending on food at home and spending on food in restaurants during a decline in economic activity? How would income elasticity of demand help explain these changes?
 6. Would you expect the cross-elasticity coefficients between each of the following pairs of products to be positive or negative? Why?
 - a. Personal computers and software
 - b. Electricity and natural gas
 - c. Apples and oranges
 - d. Bread and DVRs
 7. Why is it unlikely that a firm would sell at a price and quantity where its demand curve is price inelastic?
 8. Which products would exhibit a higher elasticity with respect to interest rates, automobiles or small appliances? Why?
 9. The immediate effect of gasoline price increases in the aftermath of the Persian Gulf crisis in August 1990 on gasoline consumption was not very significant. Would you expect the consumption of gasoline to be more severely affected if these higher prices remained in effect for a year or more? Why or why not?
 10. Suppose the federal tax on gasoline increased by 5 cents per gallon. Do you think that such an increase, reflected in the price of gasoline, would have a significant impact on gasoline consumption?
 11. Why do you think that whenever governments (federal and state) want to increase revenues, they usually propose an increase in taxes on cigarettes and alcohol?
 12. Could a straight-line demand curve ever have the same elasticity on all its points?
 13. If a demand curve facing a firm is horizontal or nearly so, what does it say about this firm's competition?
 14. A company faced by an elastic demand curve will always benefit by decreasing price. True or false? Explain.
 15. Discuss the income elasticities of the following consumer products:
 - a. Margarine
 - b. Fine jewelry
 - c. Living room furniture
 - d. Whole lobsters
 16. If the income elasticity of tomatoes is estimated to approximate +0.25, what would you expect to happen to the consumption of tomatoes as personal income rises?
 17. (Read the "The Market for Used Automobiles" section in Appendix 4A before answering the question.) When prices of used cars dropped about 10 percent in October 2001, their sales increased by 4.5 percent. Does this mean that the demand elasticity for used cars is 0.45?
 18. The U.S. Postal Service has been raising postal rates on a regular basis. The service had been losing money. One of the reasons is increased competition from companies such as United Parcel Service and Federal Express. Another reason is the use of faxes and e-mail, as well as electronic bill payment. With this decrease in demand for postal services, why do you think that the Postal Service is seeking a rate increase?
 19. A Canadian apparel company, Roots, agreed to provide the U.S. Olympic team at the 2002 Winter Olympics with various types of clothing, including berets, for free, and further, to turn over a portion of its profits on sales of this clothing to the U.S. Olympic Committee. The beret became an instant success, and Roots sold a large number of them. What type of elasticity does this arrangement represent?

PROBLEMS

- The Acme Paper Company lowers its price of envelopes (1,000 count) from \$6 to \$5.40. If its sales increase by 20 percent following the price decrease, what is the elasticity coefficient?
- The demand function for a cola-type soft drink in general is $Q = 20 - 2P$, where Q stands for quantity and P stands for price.
 - Calculate point elasticities at prices of 5 and 9. Is the demand curve elastic or inelastic at these points?
 - Calculate arc elasticity at the interval between $P = 5$ and $P = 6$.
 - At which price would a change in price and quantity result in approximately no change in total revenue? Why?
- ABC Sports, a store that sells various types of sports clothing and other sports items, is planning to introduce a new design of Arizona Diamondbacks' baseball caps. A consultant has estimated the demand curve to be

$$Q = 2,000 - 100P$$

where Q is cap sales and P is price.

- How many caps could ABC sell at \$6 each?
 - How much would the price have to be to sell 1,800 caps?
 - Suppose ABC were to use the caps as a promotion. How many caps could ABC give away free?
 - At what price would no caps be sold?
 - Calculate the point price elasticity of demand at a price of \$6.
- The equation for a demand curve has been estimated to be $Q = 100 - 10P + 0.5Y$, where Q is quantity, P is price, and Y is income. Assume $P = 7$ and $Y = 50$.
 - Interpret the equation.
 - At a price of 7, what is price elasticity?
 - At an income level of 50, what is income elasticity?
 - Now assume income is 70. What is the price elasticity at $P = 8$?
 - Mr. Smith has the following demand equation for a certain product: $Q = 30 - 2P$.
 - At a price of \$7, what is the point elasticity?
 - Between prices of \$5 and \$6, what is the arc elasticity?
 - If the market is made up of 100 individuals with demand curves identical to Mr. Smith's, what will be the point and arc elasticity for the conditions specified in parts *a* and *b*?
 - The Teenager Company makes and sells skateboards at an average price of \$70 each. During the past year, they sold 4,000 of these skateboards. The company believes that the price elasticity for this product is about -2.5. If it decreases the price to \$63, what should be the quantity sold? Will revenue increase? Why?
 - The ABC Company manufactures digital clock radios and sells on average 3,000 units monthly at \$25 each to retail stores. Its closest competitor produces a similar type of radio that sells for \$28.
 - If the demand for ABC's product has an elasticity coefficient of -3, how many will it sell per month if the price is lowered to \$22?
 - The competitor decreases its price to \$24. If cross-price elasticity between the two radios is 0.3, what will ABC's monthly sales be?
 - The Mesa Redbirds football team plays in a stadium with a seating capacity of 80,000. However, during the past season, attendance averaged only 50,000. The average ticket price was \$30. If price elasticity is -4, what price would the team have to charge in order to fill the stadium? If the price were to be decreased to \$27 and the average attendance increased to 60,000, what is the price elasticity?
 - The Efficient Software Store had been selling a spreadsheet program at a rate of 100 per month and a graphics program at the rate of 50 per month. In September 2012, Efficient's supplier lowered the price for the spreadsheet program, and Efficient passed on the savings

to customers by lowering its retail price from \$400 to \$350. The store manager then noticed that not only had sales of the spreadsheet program risen to 120, but also the sales of the graphics program increased to 56 per month. Explain what has happened. Use both arc price elasticity and arc cross-elasticity measures in your answer.

10. Given the demand equation $Q = 1,500 - 200P$, calculate all the numbers necessary to fill in the following table:

<i>P</i>	<i>Q</i>	Elasticity		Total Revenue	Marginal Revenue
		Point	Arc		
\$7.00					
6.50					
6.00					
5.50					
5.00					
4.50					
4.00					
3.50					
3.00					
2.50					

11. Would you expect cross-price elasticity between the following pairs of products to be positive, negative, or zero?
- Television sets and DVRs
 - Rye bread and whole wheat bread
 - Construction of residential housing and furniture
 - Breakfast cereal and men’s shirts
- Explain the relationship between each pair of products.
12. In order to attract more customers on Mondays (a slow day), Alex’s Pizza Shop in Austin decided to reduce the price of their pizza rolls from \$3.50 to \$2.50. As a result, Monday sales increased from 70 to 130. Also, Alex’s sales of soft drinks rose from 40 to 90.
- Calculate the arc price elasticity of demand for the pizza rolls.
 - Calculate the arc cross-price elasticity of demand between soft drink sales and pizza rolls prices.
13. According to a study, the price elasticity of shoes in the United States is 0.7, and the income elasticity is 0.9.
- Would you suggest that the Brown Shoe Company cut its prices to increase its revenue?
 - What would be expected to happen to the total quantity of shoes sold in the United States if incomes rise by 10 percent?
14. A book store opens across the street from the University Book Store (UBS). The new store carries the same textbooks but offers a price 20 percent lower than UBS. If the cross-price elasticity is estimated to be 1.5, and UBS does not respond to its competition, how much of its sales is it going to lose?
15. A local supermarket lowers the price of its vanilla ice cream from \$3.50 per half gallon to \$3. Vanilla ice cream (unit) sales increase by 20 percent. The store manager notices that the (unit) sales of chocolate syrup increase by 10 percent.
- What is the price elasticity of vanilla ice cream?
 - Why have the sales of chocolate syrup increased, and how would you measure the effect?
 - Overall, do you think that the new pricing policy was beneficial for the supermarket?

16. The Compute Company store has been selling its special word processing software, Aceword, during the last 10 months. Monthly sales and the price for Aceword are shown in the following table. Also shown are the prices for a competitive software, Goodwrite, and estimates of monthly family income. Calculate the appropriate elasticities, keeping in mind that you can calculate an elasticity measure only when all other factors do not change.

Month	Price Aceword	Quantity Aceword	Family Income	Price Goodwrite
1	\$120	200	\$4,000	\$130
2	120	210	4,000	145
3	120	220	4,200	145
4	110	240	4,200	145
5	115	230	4,200	145
6	115	215	4,200	125
7	115	220	4,400	125
8	105	230	4,400	125
9	105	235	4,600	125
10	105	220	4,600	115

17. The demand curve for product a is given as $Q = 2000 - 20P$.
- How many units will be sold at \$10?
 - At what price would 2,000 units be sold? 0 units? 1,500?
 - Write equations for total revenue and marginal revenue (in terms of Q).
 - What will be the total revenue at a price of \$70? What will be the marginal revenue?
 - What is the point elasticity at a price of \$70?
 - If price were to decrease to \$60, what would total revenue, marginal revenue, and point elasticity be now?
 - At what price would elasticity be unitary?
18. The Transportation Authority in Anytown, USA, raised bus fares from \$1 to \$1.15 on January 1, 2011. The authority's statistics show that the number of passengers riding buses decreased from 672,000 in 2011 to 623,000 in 2012.
- How much did revenue change?
 - What is the arc elasticity for bus travel in Anytown?
 - The answer to b would be correct if all conditions (except price) remained the same between 2011 and 2012. Can you think of any other changes that would have affected the result?
19. (Read the "Newspapers and Their Price Elasticity" section in Appendix 4A before answering the question.) What is the arc demand elasticity for the *London Times*? What happened to revenue as a result of the price decrease?
20. The Distinctive Fashions Company increased its advertising budget for its leading brand of women's yoga apparel from \$10,000 in 2011 to \$15,000 in 2012. Its sales increased from 900 units to 1,050 units, while the price remained the same at \$120 per garment. Calculate the advertising elasticity of these garments. Was this a wise move by the company?
21. Manning Inc., is the leading manufacturer of garage doors. Demand for residential garage door sales depends, of course, on the rate of new house building activity, which in turn depends on changes in income per capita. During the past year, Manning sold 10,000 garage doors at an average price of \$1,500 per door. In the coming year, disposable income per

capita is expected to increase from \$32,000 to \$34,000. Without any price change, Manning expects current-year sales to rise to 12,000 units.

- a. Calculate the arc income elasticity of demand.
 - b. The company economist estimates that if the price of doors is increased by \$100, they could sell 11,500 doors. What is the arc price elasticity and what would be the company's revenue?
 - c. Should they raise the price even more?
22. When the iPhone was launched in the second quarter of 2008, it was priced at \$599, and Apple sold 270,000 units during this quarter. In mid-September of that year, Apple reduced the price to \$434, and sales rose to 1,119,000 during this next quarter. Without even doing any calculation, these figures clearly show that Apple's demand was very elastic. How reliable do you think this estimate is? What other factors do you think should be taken into account for a more complete assessment of iPhone's elasticity? (<http://theblogpaper.co.uk/article/business/31may09/price-elasticity-demand-iphone>)



Applications of Supply and Demand

Chapters 3 and 4 lay the foundation for the students' knowledge of supply and demand and elasticity. Knowing these elements is essential for further study of economics and is a prerequisite for all the chapters that follow.

Before we discuss the various building blocks that will complete the study of managerial economics, this appendix endeavors to reinforce the concepts of supply and demand and of elasticity in two ways:

1. Some specific applications of supply and demand are discussed, including the effects of price controls, and excise taxes.
2. Various actual situations as reported in the press are introduced and discussed, and it is shown that the materials we have just learned can be applied to analyze these situations.

Interference with the Price Mechanism

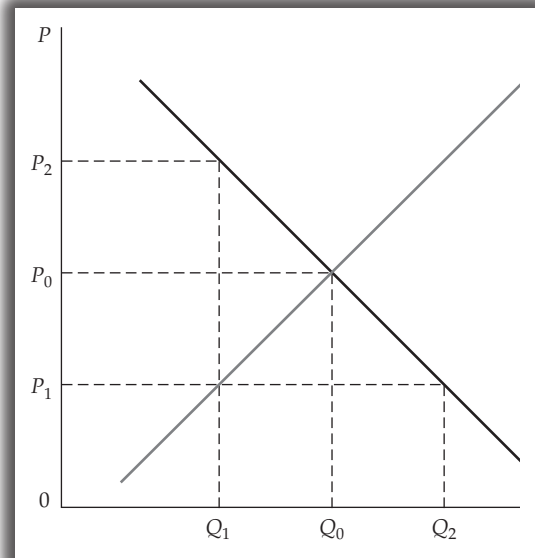
In Chapter 3, we discuss the movement toward equilibrium in both the short and long run. A change in demand or supply will call forth actions that will cause equilibrium to occur at a new supply–demand intersection. It is shown that in the short run, price changes will eliminate shortages or surpluses. In the long run, resources in the economy shift from the production of one product to another in response to changes in demand. The shift away from one equilibrium and the move to a new equilibrium will proceed when these movements are permitted to occur freely and are not impeded by any outside interference. Thus, when the supply of corn decreased and price rose so the market cleared at this new price—that is, at the new intersection of supply and demand—there was nothing inhibiting this change from taking place.

However, with present economic institutions, free movement of prices is not always allowed. At least three times in the last 70 years,²³ price controls were imposed in the United States. Prices on various products were set (or fixed at existing levels), and these products could not be sold at prices higher than those prescribed by government. Such a policy is usually

²³During World War II, the Korean War, and again in 1971.

referred to as setting a *price ceiling*. If the price ceiling for a product is set at the prevailing equilibrium level, then the ceiling will have no effect (until a change in circumstances dictated a higher price). But if the price were set below the equilibrium price,²⁴ then, as explained in Chapter 3, a shortage would result. In Figure 4A.1, the equilibrium price is P_0 and the quantity sold (and clearing the market) at this price is Q_0 . If for some reason the price winds up at P_1 under free-market conditions (i.e., no price controls), the price will rise until the equilibrium price (P_0) is again reached. But if the price is prescribed to be no higher than P_1 ,²⁵ the movement toward equilibrium

Figure 4A.1 The Effect of a Price Ceiling on Supply and Demand



²⁴It is obvious that a ceiling set above the equilibrium price would be meaningless.

²⁵Price ceilings can be enforced by the government imposing fines or even prison sentences on violators. Such punishment would have appeared rather lenient to some of our ancestors. During the times of price controls in ancient Egypt, Greece, and Rome, the death sentence was the penalty for breaking price control laws. The edict of Diocletian in A.D. 301 imposed the death sentence on those selling at prices higher than decreed, as well as on those buying at such prices (Robert L. Schuettinger and Eamonn F. Butler, *Forty Centuries of Wage and Price Controls*, Washington, DC: Heritage Foundation, 1979, p. 23).

will not take place. Only Q_1 will be supplied while Q_2 is demanded at the lower price, so a shortage of magnitude $Q_1 - Q_2$ will be established. Thus, only the consumers in the interval $0 - Q_1$ will be able to buy this particular product. What will be the result of this forced disequilibrium? Possibly consumers will try to shift their demand to other products, causing a pressure on the other products' prices. If these products are also price controlled, shortages of these other goods will occur.

There is another possible result. Because only Q_1 units of the product will be supplied at price P_1 , these units could be purchased at price P_2 along the demand curve. Consumers would be willing to pay P_2 , a price higher than the equilibrium price, P_0 , for the limited quantity Q_1 . Thus, a strong pressure on the price will be exerted, and somewhere in this process the difference between P_1 and P_2 will be paid to the suppliers.

An example of such a case was the price of automobiles after World War II. A ceiling price below the price level that would have cleared the market was imposed on new cars. This low price caused automobile manufacturers to limit their production. However, consumers were paying high prices for these cars in the way of a dealer's premium. They may also have received lower trade-in prices on their old automobiles or may have bought their new car as a "used" one because second-hand cars were not price controlled. The price they actually paid was indeed higher than it could have been if the manufacturers had charged a higher list price.²⁶ Similarly, where rent ceilings have been imposed many people end up paying a bonus to the superintendent or to a rental agent.

Another example precedes those just discussed by more than 150 years. During the Revolutionary War, the legislature of Pennsylvania imposed limits on prices of goods sold to the military and was thus instrumental in creating extreme shortages of food for George Washington's army at Valley Forge.²⁷

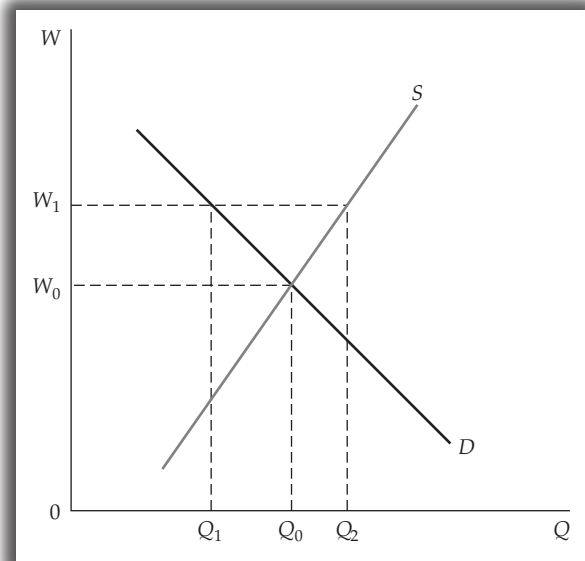
We can also present a more recent example. The African country of Zimbabwe experienced extreme hyperinflation, sometimes estimated at 10,000 percent per year. In the middle of 2007, Robert Mugabe, the country's president, issued an order slashing prices in half. In response to this order, "Zimbabwe's economy is at a halt." As reported by

the *New York Times*, "meat is virtually nonexistent," "gasoline is nearly unobtainable," and "[hospital] patients are dying for lack of basic medical supplies." Manufacturing has slowed because businesses cannot produce goods for less than their government-imposed sales prices. "It appears ... that not even an unchallenged autocrat can repeal the laws of supply and demand."²⁸

On the other side of the price control coin are price floors. In such cases, a price is established below which the product or service may not be sold. An excellent example of a price floor is the legal minimum wage. Employers are not allowed to pay their workers less than the established minimum, so they must therefore deal with the disturbance to a price equilibrium.²⁹

If the equilibrium wage (e.g., per hour) for some unskilled work were to be at level W_0 as shown in Figure 4A.2, but the law stated that a wage lower than W_1 is illegal, then a surplus of labor $Q_2 - Q_1$ would exist. In the absence of the minimum wage law, wages would drop to W_0 , and the quantity supplied and demanded of labor would meet at Q_0 . Thus all workers offering themselves for employment at that wage would be hired.

Figure 4A.2 The Effect of a Price Floor on Supply and Demand



²⁶Milton Friedman, *Price Theory: A Provisional Text*, Hawthorne, NY: Aldine, 1962, p. 18.

²⁷Schuettinger and Butler, *Forty Centuries*, p. 41.

²⁸Michael Wines, "Caps on Prices Only Deepen Zimbabweans' Misery," *New York Times*, July 28, 2007.

²⁹Wages are, of course, the price of labor, so it is quite correct to discuss minimum wages under the topic of price floors.

But if the wage cannot fall below W_1 , what will happen? The unemployed will look for work elsewhere. If the minimum wage prevails in all types of employment, they would not be able to find work. However, there are still some forms of employment in the United States that are not covered by law. A person can also become self-employed, in which case minimum wages do not apply.³⁰

The effect of increases in minimum wages on employment has been widely studied by economists for many years. In the past, most economists agreed that increases in minimum wages had a negative effect on employment, especially in the case of young and unskilled workers. Unemployment has generally been the highest among teenage workers, many of whom have dropped out of school and acquired few if any skills.³¹ However, several recent studies have questioned the traditional findings, concluding that minimum wage increases have not necessarily led to decreased employment.³² Although the new studies cast doubt on the traditional hypothesis, it is much too early to dismiss it. Much additional research is necessary to clarify the effect that increases in minimum wages have on employment.³³

³⁰At the risk of sounding facetious, the extreme case of self-employment is unemployment.

³¹Among many studies, see, for instance, C. Brown, C. Gilroy, and A. Kohen, "The Effect of the Minimum Wage on Employment and Unemployment," *Journal of Economic Literature*, 20 (June 1982), pp. 487–528; B. S. Frey, W. Pommerehne, F. Schneider, and G. Gilbert, "Consensus and Dissension among Economists: An Empirical Inquiry," *American Economic Review*, 74 (December 1984), pp. 986–94; T. G. Moore, "The Effect of Minimum Wages on Teenage Unemployment Rates," *Journal of Political Economy*, July/August 1971, pp. 897–902. A more recent study that finds negative effects on employment is D. Deere, K. M. Murphy, and F. Welch, "Employment and the 1990–1991 Minimum Wage Hike," *American Economic Review Papers and Proceedings*, 85 (May 1995), pp. 232–37.

³²See, for instance, D. Card and A. B. Krueger, "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania," *American Economic Review*, 84 (September 1994), pp. 772–84; D. Card, "Do Minimum Wages Reduce Employment? A Case Study of California, 1987–89," *Industrial and Labor Relations Review*, 46 (October 1992), pp. 38–54; L. F. Katz and A. B. Krueger, "The Effect of the Minimum Wage on the Fast-Food Industry," *Industrial and Labor Relations Review*, 46 (October 1992), pp. 6–21; *State Minimum Wages and Employment in Small Businesses*, Fiscal Policy Institute, April 21, 2004.

³³A good summary of the recent research and an analysis of various hypotheses regarding the effect of minimum wages on employment is M. Zavodny, "Why Minimum Wage Hikes May Not Reduce Employment," *Economic Review*, Federal Reserve Bank of Atlanta, 83, 2 (second quarter 1998), pp. 18–28.

Several additional points should be made regarding the impact of minimum wages. Even if an increase in the minimum does have a negative effect on employment, workers who remain employed at the higher wage will benefit. The workers can be found in the interval $0 - Q_1$.³⁴ Second, the short-run effects of an increased legal minimum are probably stronger than the long-run effects. As time passes, the wage levels in the economy will rise (either due to inflation or in real terms), and at some point the minimum wage may approach the free-market equilibrium wage. A third point also appears worth mentioning. Because an increase in the minimum wage must be passed by the legislature (the U.S. Congress in the case of a federal minimum), it is a part of the political process. Legislators may be reluctant to enact a minimum wage increase if it would increase unemployment. Thus, such legislation may be passed only if it would appear to have a minimal effect.

The Incidence of Taxes

From the viewpoint of the economics of the firm, one important example of applied analysis using supply and demand curves and elasticities is in the area of the incidence or effect of excise taxes on the prices and quantities of products.

An excise tax is a tax imposed as a specific amount per unit of product. It is also sometimes referred to as a specific tax, as compared with a sales tax, which is levied as a percent of the price of the product or service. The federal excise tax on gasoline as of this writing is 18.4 cents per gallon. The sales tax, which is usually collected by states and local communities, in the city of Phoenix, Arizona, for example, is 9.3 percent of the price of a product. Sales taxes are often referred to as ad valorem taxes. We could discuss the incidence of either ad valorem or specific taxes, but we choose the latter for our analysis. The principles and applications are similar, but a specific tax provides a simpler and more straightforward illustration.

A numerical example aids in this exposition. Table 4A.1 shows the demand and supply schedules for a particular product. The equilibrium price is \$4. At this price, 15 units will be demanded, and

³⁴It is an interesting decision, implicitly made by the U.S. Congress when it passes a law increasing the minimum wage, whether the overall welfare of the country will be increased if some part of the labor force has its wages improved while another part has its income lowered.

Table 4A.1 Demand and Supply and Tax Incidence

Unit Price	Quantity Demanded	Quantity Supplied	
		Without Tax	With Tax
\$6	5	25	20
5	10	20	15
4	15	15	10
3	20	10	5
2	25	5	0
1	30	0	

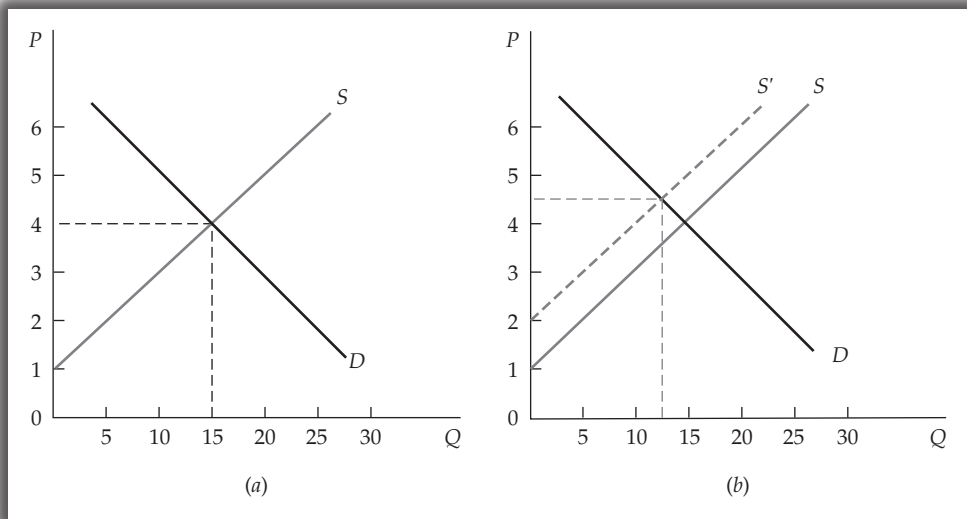
15 units will be supplied, thus clearing the market.³⁵ The demand and supply curves are shown in Figure 4A.3a, where an equilibrium at $P = 4$ and $Q = 15$ can be observed.

Now suppose the government imposes an excise tax of \$1 per unit, which it will collect from the sellers. The effect is to shift the supply curve up by the unit tax. The shift can be thought of in the following way: Before the enactment of the tax, suppliers offered to sell 20 units at \$5. But now, for the producers

to obtain \$5 per unit, these products will have to be sold at \$6 apiece (of which \$1 will be remitted to the government).³⁶ In effect, the production cost for this good has risen by \$1 per unit. The last column in Table 4A.1 shows the new supply schedule.

The important question to be asked is, What will be the market-clearing price and quantity after the imposition of the new excise tax? An easy answer would be \$1 more than before, or \$5. Certainly, the suppliers would prefer not to receive less per unit

Figure 4A.3 The Influence of an Excise Tax on Supply and Demand



³⁵An arithmetic solution can be obtained as follows: the equation of the demand curve for the schedule shown in Table 4A.1 is $Q_D = 35 - 5P$, and the equation for the supply curve is $Q_S = -5 + 5P$. Solving for $Q_D = Q_S$, we obtain

$$\begin{aligned}
 35 - 5P &= -5 + 5P \\
 40 &= 10P \\
 4 &= P
 \end{aligned}$$

³⁶The equation of this new supply curve is $Q_S = -10 + 5P$.

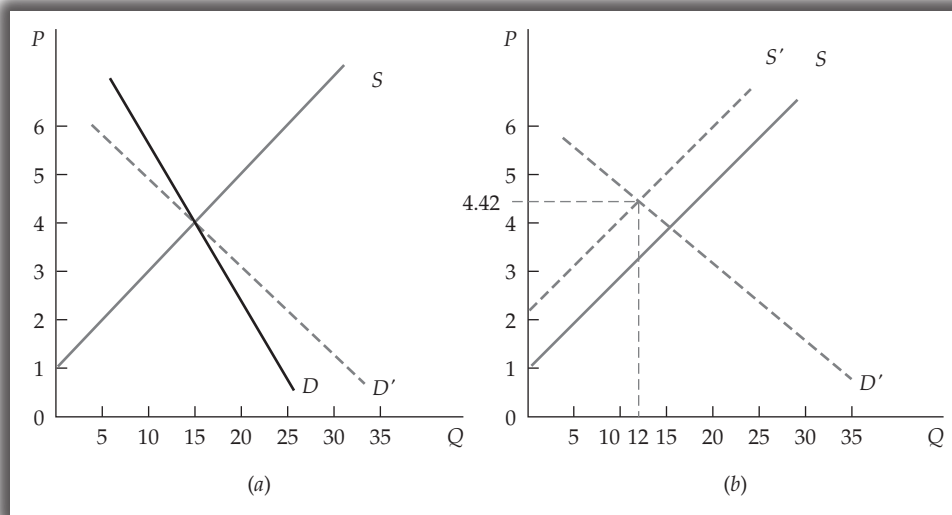


Figure 4A.4 Effect of Demand Elasticity on Equilibrium

than they had been getting before the tax. But this is not the correct answer, except in very rare cases.³⁷ The new intersection will be at \$4.50, and the quantity will be 12.5 units.³⁸

Thus, sellers will receive only \$3.50 per unit after the imposition of the tax, and consumers will be paying 50 cents more than before. In economic jargon, 50 cents of the tax has been shifted forward to consumers, and 50 cents has been shifted back to the producers. This new equilibrium is shown in Figure 4A.3b.

How the incidence of the tax is distributed between the two parties to the transaction depends on the elasticity of the supply and demand curves. The more elastic the demand curve, the larger will be the portion of the tax that the supplier has to bear. In Figure 4A.4a, we repeat the demand and supply curves previously shown and add a second demand curve, which (before the tax) also intersects the supply curve at \$4 and 15 units but at all other points is flatter (more elastic) than the original demand function. In Figure 4A.4b, the tax is added on to the supply curve. With the new demand curve, the equilibrium price is \$4.42, and the quantity demanded is just above 12 units.³⁹

³⁷This will occur where the demand curve is perfectly inelastic.

³⁸Using our equations,

$$\begin{aligned} 35 - 5P &= -10 + 5P \\ 45 &= 10P \\ 4.5 &= P \end{aligned}$$

³⁹The equation for the more elastic demand curve is $Q_D = 43 - 7P$ and the equilibrium price is

$$\begin{aligned} 43 - 7P &= -10 + 5P \\ 53 &= 12P \\ 4.4167 &= P \end{aligned}$$

The effect of the tax on the equilibrium quantity is of significance to the government unit levying the tax. It is obvious that a government setting a new (or increasing an old) excise tax is taking such action to increase its revenue. However, if the demand curve for the particular product is very elastic, the erosion of the revenue base will cut short the amount of revenue the government expects to collect. In the present case, the government would have collected \$12.50 in revenue with the original demand curve and only \$12.08 for the more elastic demand curve. Had the demand curve been perfectly inelastic (vertical), then not only would the entire tax have been shifted to the consumer, but government revenue would have been \$15 because the number of units sold would have remained at 15. Thus, a government would prefer to enact an excise tax on a product with low demand elasticity.⁴⁰

Some of the more familiar excise taxes are those on tobacco and alcohol. Because the consumption of these products is not considered desirable by today's

⁴⁰You have certainly been subjected to a tax increase in some product you consume, whether it was tobacco, gasoline, or alcohol, to mention just three products on which excise taxes are levied by both federal and local governments. You will probably recall that on the day the tax was increased, the price of, say, gasoline, rose by the precise amount of the tax. This is because, first of all, the increase may have been relatively small in comparison to the total price, so the demand curve may be quite inelastic in this relatively narrow price range. Second, as we have already learned, demand elasticity tends to be lowest in the very short run, so the tax may be completely (or almost completely) shifted to the consumer at first. But as time passes, there may be a series of small price decreases, or—and this is the more likely scenario in an inflationary environment—prices may not rise as quickly as they otherwise would have.

standards, tobacco and alcohol are frequently among the first to be selected when additional taxes are contemplated. All states, as well as the federal government, impose an excise tax on these two products. In some cases, the amount of the tax on each unit is greater than 50 percent of the total price of the product. Because of the low esteem in which these products are held by a large segment of the population, opposition to the imposition of a tax (or an additional tax) is generally not great (except by the two industries involved). These taxes are often referred to as “sin” taxes. But would it have been attractive to levy such high excise taxes on these two products had the demand curve for them been very elastic?⁴¹ Probably not, because the tax base would have eroded significantly. Therefore, the government unit that wants to achieve what is popularly known as a “revenue enhancement” will find it considerably more favorable to enact an excise tax on products whose demand elasticity in the range of the tax increase is relatively low. Tobacco and alcohol seem to fit this category well. Thus, a government unit can claim to be taxing “undesirable” commodities and at the same time help to maximize its revenue.

Among the many proposals to fight the large federal deficits of the late 1980s was the imposition of a very large tax (as much as 30 to 50 cents per gallon) on gasoline. The popular estimate has been that each cent of tax would decrease the deficit by about \$1 billion. However, such calculations may not consider what would happen to the consumption of gasoline over time. The experience with OPEC’s price increases in the 1970s and early 1980s shows that the long-run demand curve for gasoline is by no means inelastic.

The elasticity of supply is also important from the viewpoint of tax incidence. From Figure 4A.5, we can see that the effect on the price and quantity is greater the higher the elasticity of supply. If supply elasticity is relatively low, the producer will bear the greater burden of the tax.⁴²

Actual Situations

We now turn to some actual events, reported in newspapers and journals, that can easily be

⁴¹In Chapter 4, we point out that price elasticity for alcohol and cigarettes is considerably less than 1 (in absolute terms).

⁴²Again, as in the case of demand, chances are that in the very short run, supply curves will be rather inelastic because suppliers cannot immediately remove resources from this industry to another pursuit. As time progresses, however, and resources shift out of the affected industry, the supply curve will become more elastic, and both price and production will be affected to a more significant degree.

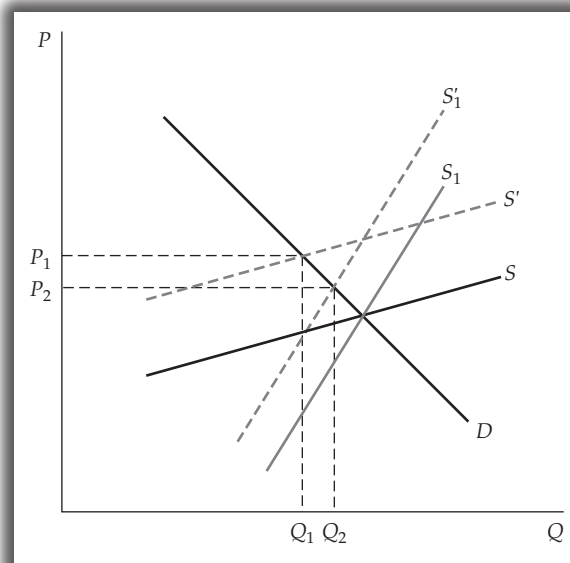


Figure 4A.5 Supply Elasticity and Tax Incidence

explained using supply-and-demand analysis. Some of these items describe events that took place a number of years ago. These older examples are included because they teach us lessons regarding supply–demand analysis that are really timeless. Together with these, a number of recent examples illustrate how the subjects discussed in Chapters 3 and 4 have current applications.

Voluntary Export Restraints

In 1981, the United States and Japan agreed that Japan would limit its exports of cars to the United States to 1,680,000 annually. The limit was later increased but was still considerably below the number that would have been sold in the United States in the absence of this quota. What was the result? The price of Japanese vehicles rose. The effect of such a limitation can be seen in Figure 4A.6a. The original quantity and price of Japanese cars sold in the United States are shown as Q_0 and P_0 . The imposition of the “voluntary” export quota on Japan at a limit less than the equilibrium quantity changed the shape of the supply curve. At the level of exports, Q_1 , the supply curve becomes vertical, and the demand curve now intersects the supply curve at P_1 —a new, higher price.

Given the restriction, the Japanese began to ship their higher-priced models to the United States to satisfy the upper portion of the market. Because the demand for Japanese cars now could

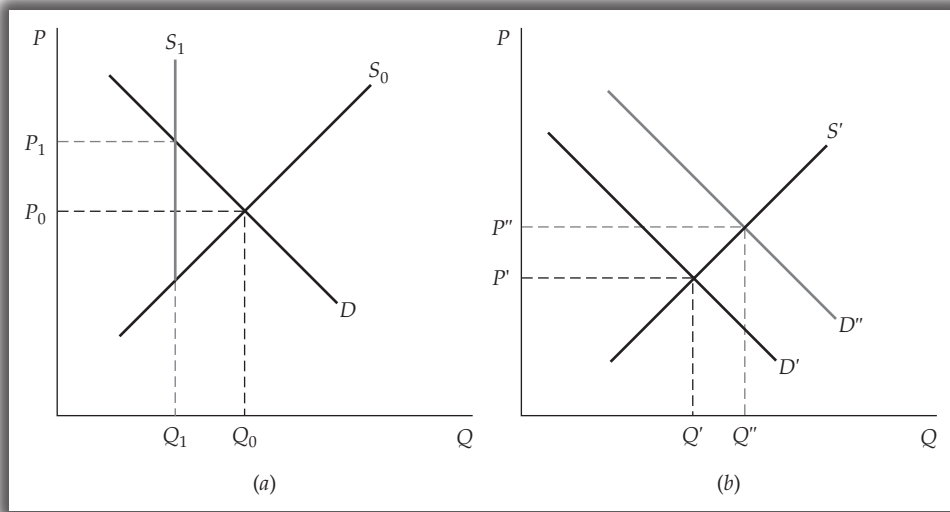


Figure 4A.6 Imposition of Voluntary Export Restraints

not be satisfied, American consumers sought to purchase domestic or other imported automobiles. Such action shifted the demand curve for the rest of the automobile market to the right, as illustrated in Figure 4A.6b, thus increasing the number and the price of vehicles purchased. It was estimated by one economist that new car prices in 1984 were \$1,500 higher than they would have been had quotas not been in effect. The additional cost to American consumers was \$13 billion. The beneficiaries of this increased cost were the big three automobile manufacturers (to the tune of about \$6 billion in higher profits), the autoworkers (about \$3 billion in overtime payments), and the American dealers of Japanese cars and the Japanese manufacturers themselves (about \$4 billion).⁴³ Subsequently, Japanese automobile manufacturers began to produce their cars in the United States. In 1985 and thereafter, the U.S. dollar weakened considerably against the Japanese yen. The prices of Japanese cars in the United States rose significantly, permitting U.S. manufacturers to continue to increase their prices and enjoy higher profits.⁴⁴ But in 1998, with the Japanese yen displaying substantial weakness, it was Japanese cars that appeared to have a price advantage.

Economic Consequences of Other Import Restrictions

A 1973 study calculated that the loss to American consumers due to import restrictions on sugar was \$586 million per year. Ilse Mintz estimated that the

restrictions raised the price of sugar by 2.57 cents per pound and decreased consumption from 23.2 to 22.4 billion pounds per year. Figure 4A.7 shows this situation. The shaded area representing consumer loss can be calculated as follows:

$$\begin{array}{r}
 0.0257 \times 22.4 \text{ billion} \quad \quad \quad \$575.7 \text{ million} \\
 0.0257 \times (23.2 - 22.4 \text{ bill.}) \times 0.5 \quad \quad \quad \underline{10.3 \text{ million}} \\
 \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \$586.0 \text{ million}
 \end{array}$$

A similar type of study performed in 1977 looked at the consequences of the imposition of import quotas on steel. The study showed that a price increase of \$11 per ton of steel and a resulting decrease in steel consumption of 0.9 million tons per year would have cost the consumer more than \$1 billion per year and would have created a gain of about \$870 million for U.S. steel producers. Thus, consumers lose more than U.S. producers gain. Part of the difference would accrue to foreign producers. Another part is due to the decrease in steel consumption, and another results due to the substitution of high-cost U.S. steel for cheaper foreign steel.⁴⁵

The Demand for Beef

From the mid-1970s into the 1980s, there was a shift away from beef demand. The result was a significant increase in the consumption of chicken and, to a lesser extent, an increase in the consumption of fish.

⁴³Yoshi Tsurumi, "They're Merely a Subsidy for Detroit," *New York Times*, December 16, 1984.

⁴⁴"Schools Brief," *Economist*, October 25, 1986, pp. 84–85.

⁴⁵Ilse Mintz, *U.S. Import Quotas: Costs and Consequences*, Washington, DC: American Enterprise Institute, 1973; Federal Trade Commission, *Staff Report on the United States Steel Industry and Its International Rivals*, November 1977. Both of these studies are discussed in Edwin Mansfield, *Microeconomics*, 5th ed., New York: W. W. Norton, 1985, pp. 100–103, 509–12.

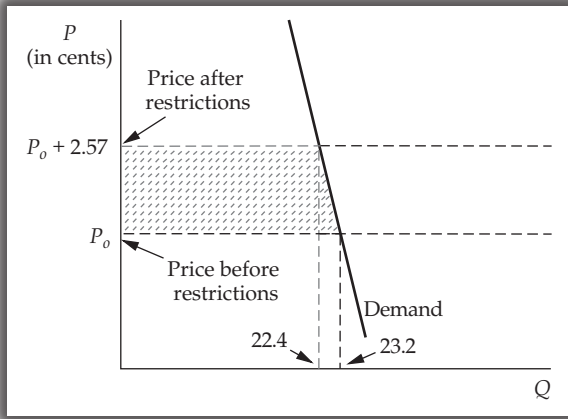


Figure 4A.7 Effect of Import Restrictions on Sugar Demand

Demand for beef had been very high into the early 1970s, and prices kept rising. However the concern about the link between red meat and cholesterol levels was instrumental in putting a check on beef demand. Also, around this same time, chicken producers developed new and convenient chicken products that appealed to the public.

During this period, the U.S. cattle population decreased substantially. As discussed previously, a change in supply is a long-run reaction to changes in demand.⁴⁶ More recently, however, the beef industry has begun to take steps that appear to have stopped the demand slide. Actually, there has been some increase in the demand for beef, as the industry has had some limited success with advertising campaigns. In 1987, cattle growers voted to establish a \$1-per-animal fund for product and nutritional research and a national advertising campaign. One possible way to combat the demand for chicken is to offer branded merchandise, a method chicken producers have been using for years. The beef industry appears to be rather slow in emulating some of the marketing techniques of the chicken producers, but, according to the research director of the beef council, “there is a sense of emergency that we have to become marketers.”⁴⁷ These trends are depicted graphically in Figure 4A.8. The original equilibrium

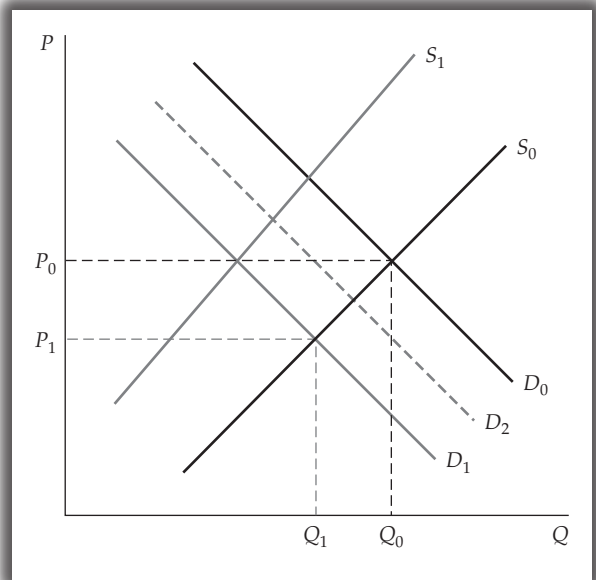
⁴⁶The discussion in this section is based on the following articles: Marj Charlier, “Beef’s Drop in Appeal Pushes Some Packers to Try New Products,” *Wall Street Journal*, August 8, 1985; Marj Charlier, “The U.S. Beef Industry Just Can’t Seem to Get the Hang of Marketing,” *Wall Street Journal*, January 4, 1989.

⁴⁷Charlier, “U.S. Beef Industry.”

of beef prices and production was at the intersection of P_0 and Q_0 . As consumer preferences shifted from beef to chicken and fish, the demand curve for beef moved to the left, from D_0 to D_1 . Quantities produced were less, and prices declined relatively.⁴⁸

Over time, resources shifted out of beef production. According to the National Cattlemen’s Association, the number of beef cattle decreased from 130 million in 1975 to about 110 million in 1985.⁴⁹ The shift of resources is shown on Figure 4A.8 as a movement of the supply curve from S_0 to S_1 . This action tended to stabilize beef prices during the period.⁵⁰ In the last few years, as just described, beef producers have begun to fight back, if not always effectively. The various campaigns mounted by the industry to convince consumers of beef’s beneficial qualities are intended to move the demand curve to the right, from D_1 and D_2 . Whether the new demand curve will be between D_1 and D_0 or even to the right of D_0 cannot be foreseen. The danger lurking in this effort to increase beef demand was that, given the low level of beef herds,⁵¹ beef prices

Figure 4A.8 The Downward Shift in Beef Demand



⁴⁸Absolute prices of both chicken and beef may have risen due to inflationary factors, which were very strong in the 1970s and early 1980s.

⁴⁹Charlier, “Beef’s Drop in Appeal.”

⁵⁰Similarly, increases in chicken-producing resources dampened the price effect of the increased chicken demand.

⁵¹Beef supplies were further depleted in 1988 because of the severe drought.

could soar, thus discouraging consumers. If the supply curve for beef in the short run is rather inelastic, this is what could happen.

What we describe was the situation prevailing during the 1980s and early 1990s. Despite some efforts by the beef industry, the demand for beef has declined by some 41 percent over the past 25 years. But more recently, the beef industry appears to have learned some lessons from the chicken producers. Now, the industry is spending hundreds of millions of dollars to “reverse a decades-long decline in red meat consumption.” Some of the large firms in the industry, such as Hormel Foods, Inc., and IBP, Inc. (now a unit of Tyson Foods, Inc.), are branding their products. One thrust in this effort is to shorten the cooking process by selling branded precooked packaged beef that can then be finished in microwave ovens in a matter of minutes. This is primarily intended to attract families with working couples that do not have the time to spend several hours preparing dinners. Although these meats tend to sell at prices considerably higher than raw meat, the two companies expected to sell about \$100 million of their product in 2002. “Within a decade, say some industry officials, microwavable red meats could easily generate \$1 billion in sales.” These companies are also producing high-quality steak and chops using their own brand labels for consumers who are not interested in microwaving their meat. In addition, some supermarket chains prefer to sell meats under their own labels, such as Kroger Company’s Cattleman’s Collection.⁵²

In 2003, it appeared that beef had made a real comeback. Demand for beef surged and prices hit records, increasing 38 percent over the prior year. These price increases were also helped by a tighter supply resulting from a ban on Canadian beef, which had accounted for 8 percent of U.S. beef consumption, due to the discovery of mad cow disease in Canada.⁵³

Demand for Various Consumer Products

It is a well-known fact that women have significantly increased their participation in the labor force since the end of World War II. After the war, men returned from the armed services, and women, who had taken the men’s place in the workforce,

returned to their traditional pursuits. In 1947, women made up about 28 percent of total employment. Women then began to enter the labor force again, and the proportion of female employment rose to 35 percent by 1965 and to 45 percent by 1987. With the change in women’s functions, a change in U.S. living patterns and consumption habits followed. According to a *Wall Street Journal* article, homes in the United States are not quite as clean as they were in the past.⁵⁴ Whereas women once spent much of their time as housewives, their acceptance into the labor force has changed this traditional role. “Sales of scouring powder, mildew removers, floor wax and dishwashing liquid slipped again last year, continuing a 10-year trend,” according to Selling Areas Marketing, Inc., a New York-based research company.⁵⁵ In contrast, sales of paper plates and aluminum baking pans have risen significantly. In addition, new time-saving household products have come to the market, and a new service business—maid services—has sprung up.

In the late 1990s, U.S. supermarkets and packaged goods manufacturers faced another shift in consumer demand. With the economy thriving and a significant number of households with working couples, there has been a vast increase in spending on restaurant and takeout meals at the expense of at-home eating. In 1997, it was estimated that away-from-home eating accounted for about 45 percent of total spending on food. The profits of packaged goods manufacturers slipped in 1998. Supermarkets were actually being hit not only by increased restaurant eating but also by competition from discount stores such as Wal-Mart. Many supermarket chains are now trying to compete by having special sections that prepare both hot and cold meals packaged for consumers who do not want to cook meals at home. Usually the profit margin on such meals is greater than on the usual items on grocery shelves, which offsets the decreases in sales of supermarkets’ regular merchandise.⁵⁶

A shift in demand patterns also occurred after the events of September 11, 2001. It appears that American consumers turned to “mundane and homey activities and purchases.” Among the beneficiaries of the change in the public’s mood were

⁵²Scott Kilman, “A Roast Is a Roast? Not in the New Game of Marketing Meat,” *Wall Street Journal*, February 20, 2002.

⁵³Scott Kilman, “U.S. Beef Enjoys a Bull Market, Enriching Ranchers and Packers,” *Wall Street Journal*, September 26, 2003.

⁵⁴Betsy Morris, “Homes Get Dirtier as Women Seek Jobs and Men Volunteer for the Easy Chores,” *Wall Street Journal*, February 12, 1985.

⁵⁵Morris, “Homes Get Dirtier.”

⁵⁶Richard Tomkins, “Home Truths for U.S. Grocers,” *Financial Times*, August 11, 1998.

stores dealing with old-fashioned crafts like quilting and knitting, and fast foods like pizza and ice cream. In contrast, there was a decline of business at fancy restaurants. Sales of DVDs and television sets as well as kitchen appliances, such as cooking ranges, were going well.⁵⁷

Newspapers and Their Price Elasticity

In July 1994, the French daily newspaper *Le Quotidien* suspended publication. To boost its revenue in the face of increasing competition and France's recession, *Le Quotidien* slashed its price from FF6 to FF4. Its circulation increased from 30,000 units to 40,000. But this resulted in a decrease of its revenue from FF180,000 per day to FF160,000. The demand curve turned out to have an arc elasticity of -0.71 . A similar action involved a decrease in the price of the *London Times* in September 1993. The price was dropped from 45 to 30 pence, while the daily prices of its competitors remained unchanged. Between August 1993 and May 1994, the *Times*' daily circulation rose from 355,000 to only 518,000, resulting in a decrease in revenue. Again, the demand curve appears to have been inelastic.⁵⁸ More recent examples, with similar results, can be found in the United States. In 2000, the *New York Post* decreased the price of its daily paper from 50 cents to 25 cents—a 50 percent decrease—resulting in a 29 percent increase in circulation. In 2001, the *Los Angeles Times* increased its price from 25 cents to 50 cents (having dropped it in 1996) and circulation fell only 5 percent.⁵⁹

There is a rather extreme example of price elasticity of demand in the newspaper industry. In 2009, the London *Evening Standard* decided to reduce its 50p price down to zero! Because P is zero, we cannot use the total revenue test to determine the degree of elasticity. However, it can be assumed that the paper's managers hoped that by making the price zero, circulation would increase to the point

where they would be able to charge more for advertising space.⁶⁰

Controlling Traffic in City Centers

In 1998, Singapore began a new system of controlling traffic in its city center. Motorists must purchase prepaid cash cards and tolls are deducted electronically from this card (which is placed on the car's dashboard). The tolls differ depending on the time of the day a car enters the city. If a car enters the central area without this card, its registration plate will be photographed and a violation notice mailed. The program appears to have been successful. Traffic has decreased by 17 percent during the enforcement period, and the speed at which cars pass through the city center has increased. Singapore is not the only city using this procedure; among others, several Norwegian cities use such cards.⁶¹ A different method to decrease traffic is used in Minneapolis. Here, one of the largest employers, American Express Financial Advisors, is subsidizing the area's bus systems in return for greatly reduced monthly bus tickets for its employees.⁶²

Although actions in both Singapore and Minneapolis aim at decreasing rush hour traffic, the two methods have very different effects on the demand for the use of automobiles. In the case of Singapore, the action has increased the cost of travel into the city center, and has caused a movement up along the demand curve, a decrease in the quantity demanded. In contrast, in Minneapolis, the result is to lower the cost of bus travel, a substitute for car commuting. The demand curve for auto travel will now shift to the left, a decrease in demand.

The effect of these two actions can be illustrated by simple graphs. Assume that the demand for automobile traffic in Singapore is as shown on Figure 4A.9a. If the quantity demanded before the imposition of the toll is at P_0 and Q_0 , then after the toll is implemented, the new equilibrium will be at the intersection of P_1 and Q_1 . This, of course, as we saw in Chapter 3, is a decrease in quantity demanded. Now, let us take a look at the Minneapolis situation. In this case, Figure 4A.9a represents the demand for bus travel. The decrease in bus fares

⁵⁷Leslie Kaufman and Julian E. Barnes, "Craving a Comfort Zone," *New York Times*, October 10, 2001; Christopher Bowe, "Consumers Turn to the Kitchen in Times of Strife," *Financial Times*, October 16, 2001; and Elliot Spagat, "Retailers Pin Hope on 'Cocooning' TV Buyers," *Wall Street Journal*, October 18, 2001.

⁵⁸Alice Rawsthorn, "Crisis in French Press May See More Casualties," *Financial Times*, July 6, 1994; R. W. Stevenson, "A Cheaper Times of London Wins Readers," *New York Times*, June 13, 1994.

⁵⁹Patricia Callahan, "Four Knight Ridder Papers Cut Prices," *Wall Street Journal*, June 11, 2002. The relatively small response to a price decrease, may not lead to a decrease in profitability. When newspapers' circulations increase, newspapers may increase their rates for advertising, thus making up for the loss.

⁶⁰<http://www.tutor2u.net/blog/index.php/gcse-economics/comments/free-london-evening-standard-price-elasticity-of-demand>

⁶¹Sheila McNulty and John Parker, "How to Stop Traffic Jams," *Financial Times*, May 9, 1998.

⁶²*Wall Street Journal*, August 27, 1998.

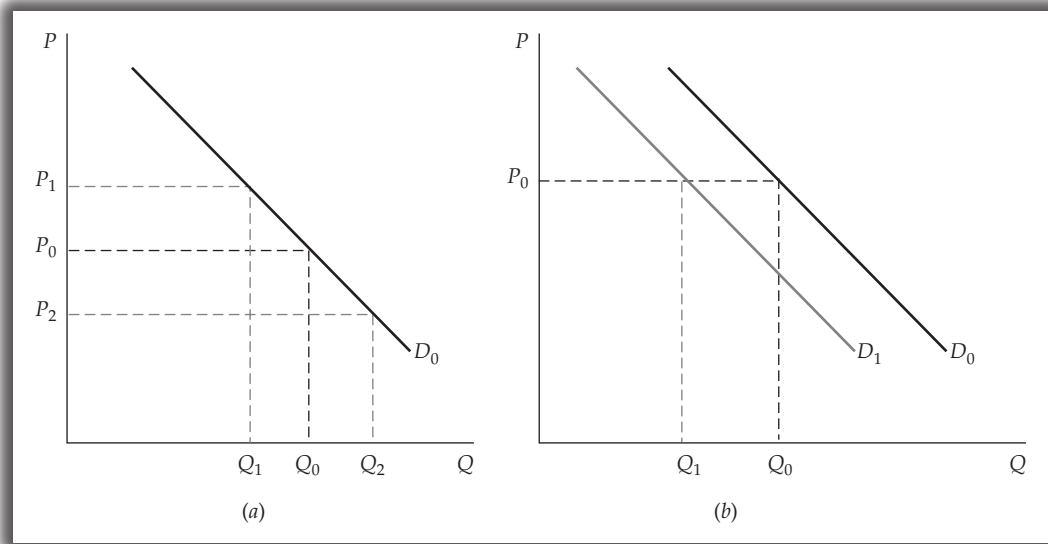


Figure 4A.9 The Effect of Two Different Policies on Automobile Travel

resulting from the financial aid given to the bus systems, will cause a move on the demand curve from P_0 (the price of fares before the decrease) to P_2 ; this is an increase in quantity demanded from Q_0 to Q_2 . Figure 4A.9b depicts the demand for automobile traffic into downtown. Demand curve D_0 represented the demand before the bus fare cuts went into effect. With the decrease in bus fares, the substitution effect will result in a new demand curve D_1 , and the demand for automobile travel will decrease from point Q_0 to Q_1 .

The city of London adopted a system very similar to that of Singapore in February 2003. A charge of £5 was levied on automobiles entering the center of London on weekdays from 7 A.M. to 6:30 P.M. Six months later, it was reported that the scheme had decreased London traffic by 16 percent⁶³ and in 2005 it was reported that congestion in London was reduced by 30 percent. Also in 2005, the congestion charge was increased to £8 and discounts were introduced for monthly or annual payments.⁶⁴ Stockholm conducted a trial of the congestion tax in 2006 and implemented it on a permanent basis on August 1, 2007. It is in effect from 6:30 A.M. to 6:30 P.M. on weekdays. The charge varies during the time of day.⁶⁵ And finally, in April 2007, the mayor

of New York City proposed an \$8 tax for cars entering the lower portions of Manhattan.⁶⁶ However, the New York State legislature voted not to enact the mayor's proposal.

The Market for Used Automobiles

In 2005, in order to improve sales, American automobile manufacturers initiated "employee pricing," giving to all buyers discounts previously granted only to employees. Consequently, many buyers who would otherwise buy used cars were now purchasing new ones. The result was a significant increase in the sales of new cars, and a corresponding increase in trade-ins. The glut of used cars drove their prices down to the lowest levels in at least two years. In the two months from May to July 2005, prices fell at least 5 percent and were expected to decline further. Dealers throughout the United States had 120,000 more used cars on their lots than they would have had if there had been no increase in new-car sales.⁶⁷

A similar situation occurred in 2001, when car dealers began to offer 0 percent financing on new cars in the aftermath of the attacks on the World Trade Center and Pentagon. In October 2001 new car sales rose by 35 percent. In addition, because of the expected decrease in travel, many car rental

⁶³"Smooth Start to London Traffic Tax," *www.cnn.com/World*, February 17, 2003; "London's Traffic Tax Paying Off," *www.news24.com*, August 17, 2003.

⁶⁴"Congestion Charge Increases to £8," *BBC News*, April 1, 2005.

⁶⁵"Stockholm Congestion Tax," *Wikipedia*, August 26, 2007.

⁶⁶Maria Newman, "Mayor Proposes a Fee for Driving to Manhattan," *New York Times*, April 22, 2007.

⁶⁷Jennifer Saranow, "Where the Real Deals Are: Luxury Used Cars," *Wall Street Journal*, August 11, 2005.

companies reduced their fleets by 20 to 30 percent. As a result the inventory of used cars increased sharply and resulted in price decreases of 10 percent and even as much as 20 percent.⁶⁸ (Try to draw demand and supply curves for new cars and used cars to illustrate this situation.)

The French Wine Industry

Over the past several years, the French wine industry has been suffering. From 1999 to 2000 alone, exports of French wines dropped 5.4 percent, and millions of liters of unsold wine were accumulating in French cellars. France's market share in the United States has dropped from 7 percent to 5 percent in 3 years. During this period, sales of producers in California, Australia, and Chile have been rising sharply in the world's important markets of North America, Northern Europe, and Asia. What are the major reasons for this shift in demand from the world's leading wine producers to a group of upstarts in the rest of the world? According to a recent article there are several:⁶⁹

1. French vintners are restricted by the government as to the types of grapes that can be produced in a particular region, and how they are cultivated.
2. They are paid by the amount of grapes they deliver, thus impairing careful growing, and resulting in inconsistent quality.
3. Predominantly, French producers are small, scattered firms (there are about 20,000 of them) that are finding it difficult to compete efficiently with large, new-world companies. Only one French company is among the world's ten largest producers.
4. The many different regional and local labels of French wines make it difficult for consumers (except for connoisseurs) to understand what they are buying, compared with the simple labeling by the new world's large wine-producing companies.
5. The dispersed nature of the French wine industry is no match for its large foreign rivals when it comes to marketing and promotion programs. For example, U.S. company E. & J. Gallo, the world's largest producer, spent \$2.5 million on marketing in England alone, while the entire Bordeaux wine-producing region spent less than half of that amount in 2000.

⁶⁸Sholem Freeman and Karen Lundegaard, "New-Car Deals Shake Up Market for Used Vehicles," *Wall Street Journal*, November 16, 2001.

⁶⁹William Echikson and others, "Wine War," *BusinessWeek*, September 3, 2001, pp. 54–60.

The problems of the French wine industry are not going away. The crisis appears to continue. Exports continue to fall, and the French population is consuming less wine. France's wine exports continued to fall; between 2005 and 2009, they declined, on the average, 1.1 percent per year in value and 9.4 percent in volume. Even so, France was still the world's leading wine producer in 2010.⁷⁰ Wine growers have met with government ministers to try to find a solution. One of the suggestions was to reclassify wine as a "natural food" rather than an alcoholic drink under French law. Such a change would permit wine makers more freedom to advertise. Advertising is now restricted due to a law passed in 1991 to decrease alcohol consumption in France. In a way, this presents the French government with a paradox. It has undertaken a program to reduce alcohol consumption and there has been some success in decreasing the per capita consumption of wine and also decreasing the number of deaths due to automobile accidents. "Overproduction, competition from New World wines and falling domestic consumption have led to steep price cuts for French wine."⁷¹

On the other hand, the problem that French champagne makers face is just the opposite. The industry continues to prosper. The global sales of champagne have been rising (some 54 percent between 1990 and 2004). However, the champagne grape growing area is limited to 84,000 acres due to a 1927 law that set the boundaries. There is a great deal of competition for any acreage that becomes available. A two-and-a-half-acre tract producing high-quality champagne may sell for \$1.2 million. There has been a push to increase the size of the champagne region, but it is not expected that this will happen in the next several years.⁷²

The Woes of the U.S. Furniture Industry

Historically, rising home sales have led to healthy sales of furniture. In 2002, record sales of new and used homes were forecasted. However, the U.S. furniture industry did not benefit from this growth. The reason for this state of affairs is intense competition, mainly from China and also from Canada. Furniture imports to the United States are reported to have

⁷⁰"France Reclaims Top Spot in Global Industry: Study," *Agence France Presse*, January 11, 2011.

⁷¹Jo Johnson, "Winemakers Call Crisis Talks with Ministers as French Exports Fall," *Financial Times*, July 22, 2004.

⁷²Sarah Nassauer, "Champagne's Land Lock," *Wall Street Journal*, August 12, 2005.

risen some 71 percent since 1999, and it is estimated that imports now comprise more than 40 percent of the U.S. market. The growth has been particularly significant in metal and wood furniture—less so in upholstered furniture.

Although cheaper imports benefit the consumer, they have had a devastating effect on the U.S. furniture industry. “To compete, some U.S. companies are joining the competition.” One example is Furniture Brands International, Inc. After closing down several of its domestic plants, it began to import products for its Lane division from Asia.⁷³

A Price Action That Misfired

In October 2003, the Universal Music Group, a unit of Vivendi Universal SA and the world’s largest music company, cut its CD prices by up to 30 percent. This action, which Universal named “JumpStart,” was taken to improve the suffering music business. The CD wholesale price was lowered to \$9.09 from \$12.98, and it was expected that retail shops would lower their prices to \$12.98, the price suggested by Universal. Universal expected that this retail price decrease would make it possible for the smaller stores to compete with mass merchants. Given the illegal downloading of music, this lower price did not appeal to consumers as much as was expected. Moreover, many music retailers did not implement the price cut or were otherwise slow to put it into effect. The smaller retailers believed that they could not afford to cut their prices, unlike some of their big rivals—such as Wal-Mart—that can afford to cut their margins and make up the decrease with sales of other, higher-price products. Another problem created by Universal was that it ceased making “co-op advertising” payments to retail stores. These payments actually represented a subsidy to help retailers from sustaining losses. Universal had calculated that, given the price cuts, it needed a 21 percent increase in volume to break even. However, increases amounted to only 8 percent to 13 percent in most weeks. As a result, Universal’s revenue decreased by 11 percent during the first quarter of 2004. In April 2004, Universal partially backtracked and raised its wholesale prices; however, they remained lower than before JumpStart.⁷⁴

⁷³Jon E. Hilsenrath and Peter Wonacott, “Imports Hammer Furniture Makers,” *Wall Street Journal*, September 20, 2002.

⁷⁴Ethan Smith, “Why a Grand Plan to Cut CD Prices Went Off the Track,” *Wall Street Journal*, June 4, 2004.

Verizon and Demand Elasticity

In 2005 the federal government announced that it would discontinue a surcharge, called the Universal Service Fund fee, that it had levied on Verizon’s DSL service. Verizon could have passed on this price break to its customers and possibly attracted more customers at the lower price. Instead, Verizon decided to replace the government fee with a supplier surcharge of about the same amount. How could this be explained? The answer is price elasticity of demand. Apparently Verizon assumed that the increase in number of customers would not bring in as much revenue as raising the price. In other words, the assumption was that at the current price the demand was price inelastic. Of course, this could be a short-run occurrence. In the long run, customers may find alternative services to be a better deal.⁷⁵

Tobacco Production Is on the Rise

Tobacco production had been subsidized by the U.S. government since 1938. The law guaranteed a minimum price and at the same time allotted quotas to tobacco farmers. In 2004 subsidies were discontinued. As a consequence, the acreage devoted to tobacco growing in the United States dropped from 408 thousand acres to 297 thousand acres, a decrease of 27 percent. But since 2005, acreage has risen some 20 percent to 355 thousand acres. The unsubsidized tobacco is now cheaper (about \$1.60 per pound compared to \$1.98 before subsidies were ended), and production in the United States has become competitive as an export. China, Russia, and Mexico have become major importers as demand for cigarettes is still growing in those countries. Tobacco companies are recruiting new tobacco farmers and are offering them financial assistance. Tobacco is now being produced in areas such as southern Illinois where it had not been produced in any significant quantities since World War I. In Pennsylvania, production has more than doubled since 2004. A farmer in Illinois has estimated that, even at this lower price, he is netting about \$1,800 per acre of tobacco compared to \$250 from his corn crop, even though corn is now selling at record prices due to the increased demand for ethanol. Thus, in the words of Professor David Orden of Virginia Tech, “we are finding out that farming can be done without subsidies.”⁷⁶

⁷⁵Andrew Cassel, “The Economy, Verizon’s Lesson about Elasticity of Demand,” *Philadelphia Inquirer*, August 23, 2006.

⁷⁶Lauren Etter, “U.S. Farmers Rediscover the Allure of Tobacco,” *Wall Street Journal*, September 18, 2007.

The Demand for Corn

The ethanol boom pushed corn prices to over \$4 per bushel in early 2007. At the end of March, the U.S. Department of Agriculture reported that farmers were expected to plant some 90.5 million acres of corn in 2007, a 15 percent increase over 2006 and the highest acreage since 1944. As a result, the acreage for other products, particularly soybeans and cotton, would be changed by 11 percent and 20 percent, respectively. On that day, corn's future prices dropped by 5 percent. By September 2007 the price of corn had decreased to \$3.65 per bushel and was expected to drop further to \$3.20 by the end of the year as the demand for ethanol was lower than previously predicted. In the same month the Department of Agriculture raised its forecast of corn production to 13.3 billion bushels, 26 percent over the prior year, and expected that stocks of corn will be double those in 2006.⁷⁷

The Coffee Glut in Brazil

A very large crop of coffee in Brazil, the world's leading coffee bean grower, as well as in other countries, caused a steep decline in the price of coffee in 2001.

Brazil attempted to decrease the supply of coffee by using a "global retention scheme." Brazil was going to keep more than 2 million bags (at 60 kilograms per bag) off the market. However, this plan collapsed in September 2001 when other coffee-producing countries refused to cooperate. In November, Brazil's policy makers planned to introduce another program, under which the government would finance 70 percent of the value of coffee that producers would hold back for a year.

⁷⁷Lauren Etter, "Hot Commodity on Wall Street: USDA's Crop-Production Report," *Wall Street Journal*, March 24–25, 2007; Andrew Martin "Farmers to Plant Largest Amount of Cotton Since '44," *New York Times*, March 31, 2007; Prasenjit Bhat-tacharya, "S Corn Prices Forecast Down to \$3.20 Per Bushel by End '07," *Dow Jones Commodities Service*, September 4, 2007; Bill Thompson, "USDA Sees Higher US Corn Production; Weaker Ethanol Demand," *Dow Jones Commodities Service*, September 12, 2007. However, contrary to these forecasts, the price of corn stood at approximately \$8.00 per bushel in June 2008.

One of the problems that producers in Brazil face is the rather low productivity and thus high cost of coffee production. According to one authority, "at current prices, the break-even point is at 45 bags per hectare (2.471 acres). But in the past year, productivity has averaged 17 bags around here." Many coffee producers have had to seek alternative products, such as soybeans, passion fruit, and grains. It has also been estimated that 10 percent of the coffee-growing area has been abandoned since the end of the year 2000, and it is expected that additional land will be left idle by 2002 harvest time.⁷⁸

The Consumption of Cigarettes

Articles on prices and consumption of cigarettes appear frequently in the press. They generally conclude that the demand for cigarettes is price inelastic. Thus, increases in prices will generally result in increases in revenue. In addition, when cigarette prices rise due to increases in taxes, governments will gain revenue.

In the United States, a 25 percent increase in cigarette prices in 2009 caused consumption to fall 8 percent, an elasticity of -0.34 .

In Ireland, a study showed that price elasticity was -0.32 . A €1 increase in prices would result in about 30,000 fewer smokers.

In 2010, a sharp tax increase in Japan caused Japan Tobacco Inc. to estimate a price elasticity of about -0.7 .

And finally, when retail prices of cigarettes in France rose 39 percent from 2002 to 2004, price elasticity was estimated to be about -0.8 .⁷⁹

⁷⁸Thierry Ogier, "There's an Awful Lot of Coffee in Brazil," *Financial Times*, November 9, 2001.

⁷⁹"Tobacco," *Financial Times*, June 17, 2010; Aoife Carr, "A €1 Rise in Cigarettes Would See 30,000 Quit," *The Irish Times*, October 19, 2011; "JT Looks Overseas to Counter Tax Boost in Home Market," *Nikkei Weekly*, October 4, 2010.

SUMMARY

As the preceding examples show, it is extremely important that a manager or entrepreneur understand such cause-and-effect relationships. Demand and supply curves move continually in an economic society. It is management's task to identify such movements, understand their consequences, and design strategies and tactics to minimize adverse results and take advantage of new opportunities. Knowledge of economic interrelationships is a tool that will assist decision makers in taking these important actions.

Demand Estimation and Forecasting

Learning Objectives

Upon completion of this chapter, readers should be able to:

- Specify the components of a regression model that can be used to estimate a demand equation.
- Interpret the regression results (i.e., explain the quantitative impact that changes in the determinants have on the quantity demanded).
- Explain the meaning of R^2 .
- Evaluate the statistical significance of the regression coefficients using the t -test and the statistical significance of R^2 using the F -test.
- Recognize the challenges of obtaining reliable cross-sectional and time series data on consumer behavior that can be used in regression models of demand.
- Understand the importance of forecasting in business.
- Describe six different forecasting techniques.
- Show how to carry out least squares projections, and decompose them into trends, seasonal, cyclical, and irregular movements.
- Explain basic smoothing methods of forecasting, such as the moving average and exponential smoothing.

The Situation

Frank Robinson, recently brought into Global Foods, Inc., to build a forecasting department, finds his new position to be very challenging and quite interesting. However, he also knows that forecasting, even in the relatively stable soft drink industry, can be a thankless task. From various forecast requests on his desk, he pulls out the one for the company's lemon-lime soda, Citronade, a brand recently purchased from an older, established company. He has been asked to estimate sales for the next year, and the deadline for his report is nearing. He has annual sales data for the last 11 years, and he also has sales data by quarter. In an industry where sales show considerable increases during the summer months, a forecast that estimates sales for the seasons of the year is very important.¹

(continued)

¹Sales forecasts for a real company are usually made on a monthly, not a quarterly, basis. But for the purposes of this example, quarterly data are used. This is done to economize on the quantity of data used. The methods employed in the analysis of the data are identical whether quarterly or monthly figures are employed. In fact, in the appendix of this chapter, we provide an expanded example using thirteen 4-week time periods per year.

(continued)

Frank first looks at the annual data, and he quickly computes year-to-year changes. These numbers are shown in Table 5.1. He notices that, although sales were up in each of the years, the percentage growth from year to year appears to have a declining trend. He will have to consider this phenomenon when he makes his forecast.

Quarterly data are shown in Table 5.2. As he looks over these numbers, Frank realizes that he has several busy days ahead of him.

Table 5.1 Sales of Citronade (in Thousands of Cases)

Year	Annual Sales	Change	Percentage Change
2002	3,892		
2003	4,203	311	8.0
2004	4,477	274	6.5
2005	4,810	333	7.4
2006	5,132	322	6.7
2007	5,407	275	5.4
2008	5,726	319	5.9
2009	6,023	297	5.2
2010	6,360	337	5.6
2011	6,641	281	4.4
2012	6,954	313	4.7

Table 5.2 Quarterly Sales of Citronade, 2002–2012 (in Thousands of Cases)

Year	1st Qtr	2nd Qtr	3rd Qtr	4th Qtr	Total
2002	842	939	1,236	875	3,892
2003	907	1,017	1,331	948	4,203
2004	953	1,103	1,406	1,015	4,477
2005	1,047	1,180	1,505	1,078	4,810
2006	1,124	1,267	1,576	1,165	5,132
2007	1,167	1,340	1,670	1,230	5,407
2008	1,255	1,403	1,766	1,302	5,726
2009	1,311	1,495	1,837	1,380	6,023
2010	1,390	1,565	1,940	1,465	6,360
2011	1,455	1,649	2,026	1,511	6,641
2012	1,536	1,714	2,103	1,601	6,954

DEMAND ESTIMATION

INTRODUCTION

Key Chapter Objectives

In the previous chapters, we analyze the demand function from a theoretical standpoint. We show how each of the determinants of demand—price and nonprice factors, such as tastes and preferences and income—affects the amount people are

willing to purchase of a particular good or service. Hypothetical numerical examples were devised to illustrate the concept of elasticity, a key way in which we measure the sensitivity of quantity demanded to changes in price and the nonprice determinants of demand. It is vital for managers to use whatever tools and data are available to explain and forecast demand. Think about how much better off the managers of Nokia would have been if they had been able to forecast the shift among consumers away from cell phones to smart phones. Also think about the challenges facing Sony and Nintendo as they try to forecast their demand. They realize that consumers are shifting away from games that are hosted on consoles to games that are made available on the Internet. This change in consumer taste has probably affected the accuracy of forecasts that they made in the past and undoubtedly will have an effect on forecasts of future sales.

This chapter presents two important statistical approaches to estimating and forecasting the demand for a product. We recognize that both topics are covered extensively in other courses offered in a typical business curriculum. Moreover, the power and ease of use of various software packages make it relatively easy for managers to generate quantitative studies of demand using regression and forecasting techniques. Therefore, this chapter is intended to be an overview of how the techniques of analysis are used in various types of studies. But these advances do not diminish the importance of obtaining good data, correctly interpreting and evaluating the analytical results, and using the results to make good management decisions. These factors are the essence of our chapter's coverage of this topic. Let us begin with a statement about the importance of obtaining good data for use in whatever type of statistical analysis managers can use.

The Critical Importance of Good Data

When computers began to be used for business operations and analysis, there was a frequently used acronym, "GIGO," which stood for "garbage in, garbage out." Readers may not recall this somewhat dated expression, but recounting it helps make an important point: Statistical analyses are only as good as the accuracy and appropriateness of the sample of information that is used.

Readily available and reliable data for use in economic studies pertaining to countries, regions, or industries (e.g., analysis and forecasting of the gross domestic product, interest rates, foreign exchange rates, industry output) are provided by official sources such as the U.S. Department of Commerce, the U.S. Department of Labor, the Organization of Economic Development, the World Bank, and the United Nations. However, data that can be used for the analysis of specific product categories (e.g., demand for pizza or smart phones) may be more difficult and costly to obtain.

The leading providers of market research services in the United States and for much of the developed countries of the world are ACNielsen and Information Resources, Inc. (IRI). A visit to either company's Web site will provide readers with a good idea of the kinds of products and services offered to market research clients, most of whom are large food and beverage and other types of fast-moving, packaged, consumer goods companies. No doubt if our "Global Foods," were a real company, it would be using the services of one of them, as do actual companies such as Kraft Foods, Coca-Cola and PepsiCo, Nestlé and Unilever. In the appendix to this chapter, we utilize actual historical data on the demand for wine that was obtained by IRI.²

There are numerous ways in which market researchers seek out the "truth" of consumer behavior. They might use the direct approach of a **consumer survey**, either face to face (e.g., stopping people in shopping centers) or by telephone. Another direct approach is the focus group. Typically, this method involves market researchers

²Go to www.acnielsen.com and www.symphonyiri.com.

observing behind a two-way mirror the answers and body language of selected groups of consumers who are asked specific questions about a company and its competitors. Perhaps some of you have been involved in this type of data gathering.³

Another way that ACNielsen measures the behavior of consumers is by having people participate in consumer panel surveys. Volunteers in this activity are provided with a scanner that enables them to record their store purchases by reading the bar codes of every item they bought. The survey participants then transmit the scanned data to the market research firm on a regular basis via a special modem attached to their home phones.

Technology is making it possible for retail stores to accumulate, store, and process large amounts of data about consumer behavior. The data are primarily collected with the aid of scanning technology such as bar code readers and point-of-sale terminals. RFID technology (radio frequency capability encoded into computer chips) has made it even easier for stores to track their inventory and sales. Wal-Mart, the largest company in the world, mandated that a number of its major suppliers attach these devices to their products. Already Wal-Mart has built a data warehouse of information about what every single customer purchases in every one of its stores every day of the year. The amount of data stored in its warehouse is allegedly second only to that stored by the U.S. government.

In recent years, one of the popular buzz words in the field of statistical analysis is *big data*. There seems to be different ways to define this term. One of the authors has heard it described as “data that is too much to handle by one computer.” Another variation of this definition is that it is “data that cannot be handled by Excel.” Regardless of how it is defined, Network World states that big data consists of “three V’s.”⁴ That is, volume (the massive amount of data now being generated by organizations), velocity (the speed at which this data is being generated), and variety (data from texts, audio, video, web logs, etc.). The extraction of this data for analysis is often referred to as data mining and the analysis of this data is usually called business analytics.

The analysis of data ranges from simple descriptive statistics on measures of central tendency (i.e., the mean, median, and mode) to very sophisticated inferential multivariate analyses such as **regression analysis**. The procedure most commonly used by economists is regression analysis. In the ensuing sections, we present a brief introduction to regression analysis and how it is applied in business. In this chapter, we focus on demand analysis. In the next two chapters, we show how regression analysis is used in the analysis of production and cost functions.

INTRODUCTION TO REGRESSION ANALYSIS

Specifying the Regression Equation and Obtaining the Data

In estimating the demand for a particular good or service, first determine all the factors that might influence this demand. Suppose we wanted to estimate the demand for pizza by college students in the United States. What variables would most likely affect their demand for pizza? We could start to answer this question by using price and all the nonprice determinants listed in Chapter 3 (i.e., tastes and preferences, income, prices of related goods, future expectations, number of buyers). But it is not always possible or appropriate to include all these variables in a particular demand analysis.

³One of the authors actually likes to participate in telephone consumer surveys because it gives him a good idea of what companies are looking for in their market research. He has also recently participated in a focus group session initiated by a major airline. The purpose of the session was to find out what the participants thought about one of the leading low-cost airlines and what the company could do to fight back against this type of competition. In exchange for his time, he was given 10,000 frequent flier points.

⁴<http://www.networkworld.com/news/2012/051012-big-data-259147.html>

As an example of this, in the demand for pizza, one would not think that “future expectations” would play an important role. In addition, there may be sudden shifts in tastes and preferences that may not be measured easily. For example, the “low carb” diet craze may well have an impact on the demand for pizza but it would be somewhat challenging to measure this change.⁵

Ideally, all variables that are believed to have an impact on demand should be included in the regression analysis. In reality, the variables used in regression analysis are based on the availability of data and the cost of generating new data. The two types of data used in regression analysis are **cross-sectional** and **time series**. Cross-sectional data provide information on variables for a given period of time. Time series data give information about the variables over a number of periods of time. For the purpose of illustration, let us assume we have obtained cross-sectional data on college students by conducting a survey of thirty randomly selected college campuses in the United States during a particular month.

Suppose we have gathered the following information for each campus from this survey: (1) average number of slices consumed per month by students, (2) average price of a slice of pizza in places selling pizza in and around the campus, (3) annual tuition cost, (4) average price of a soft drink sold in the pizza places, (5) location of the campus (urban versus suburban or rural), and (6) the type of college (residential versus nonresidential). The data from our hypothetical survey are shown in Table 5.3 and in the Multivariate Regression Excel App.



The reasons for selecting these variables are based on the economic theory of demand. Therefore, it should be clear why the price of pizza and the price of its complementary product, a soft drink, were selected for this study. But sometimes a researcher may have to use some creativity in coming up with variables that represent such factors as income and tastes and preferences. Because of the difficulty of finding out the average income of the students (or their families) who attend a particular college, tuition was used as a proxy variable. The location dummy variable is included to determine whether the demand for pizza is affected by the number of available substitutes for pizza. The assumption behind this is that colleges in urban areas may have more eating establishments from which to choose, and this might adversely affect the students’ demand for pizza. Finally, a dummy variable for residential colleges is included based on the assumption that students at these schools may have already purchased meal plans on campus that will moderate their food purchase behavior off-campus.

Using these data, we then express the regression equation to be estimated in the following linear, additive fashion:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5$$

where Y = Quantity of pizza demanded (average number of slices per capita per month)

a = Constant value or Y intercept

X_1 = Average price of a slice of pizza (in cents)

X_2 = Annual tuition (in thousands of dollars)

X_3 = Average price of a 12-ounce can of soft drink (in cents)

X_4 = Location of campus (1 if located in a concentrated urban area, 0 otherwise)

X_5 = Residential college (1 if college is residential, 0 otherwise)

b_1, b_2, b_3, b_4, b_5 = Coefficients of the X variables measuring the impact of the variables on the demand for pizza

⁵However, this may be just a short-run effect. The popularity of the low carb and other diets usually tends to have a limited life. But pizza has been a staple of campus life for a long time.

Table 5.3 Sample Data: The Demand for Pizza

College	Q_{slices}	P_{slice}	Tuition	$P_{\text{softdrink}}$	Urban	Residential
1	9.89	97.3	17.92	98.8	1	1
2	12.07	101.2	26.36	95.2	1	0
3	13.08	91.0	14.57	110.8	1	0
4	13.90	95.0	7.70	91.9	1	0
5	8.93	109.6	14.56	100.0	0	1
6	8.05	126.4	9.00	100.8	0	0
7	3.98	126.4	17.48	122.9	1	1
8	3.18	149.4	20.46	149.1	0	0
9	15.07	78.8	29.26	102.4	1	0
10	11.86	75.7	13.96	87.8	1	1
11	13.10	91.9	12.17	80.2	1	0
12	13.87	103.4	3.76	75.2	1	0
13	12.02	101.6	18.72	99.6	1	0
14	10.19	113.0	20.69	123.5	0	0
15	9.88	121.7	17.63	131.2	0	0
16	11.94	112.6	20.71	79.2	1	0
17	10.93	150.1	21.91	88.5	0	0
18	11.99	100.3	17.97	94.1	1	0
19	10.00	151.4	18.00	101.6	0	0
20	8.12	145.3	18.71	90.0	0	0
21	9.18	150.5	25.01	94.3	0	0
22	9.86	129.7	18.44	98.4	1	0
23	11.00	123.0	24.15	95.0	1	0
24	11.91	104.5	22.87	99.8	0	0
25	13.12	82.0	18.75	100.9	1	0
26	10.02	100.3	18.63	112.1	1	0
27	8.91	109.4	6.35	124.3	0	0
28	7.96	128.1	13.99	90.2	0	1
29	7.89	155.1	8.83	80.5	0	0
30	4.96	148.7	13.83	97.1	0	1

Y = Average quantity of slices per month, Q_{slices}
 X_1 = Average price per slice (in cents), P_{slice}
 X_2 = Tuition (in \$000)
 X_3 = Average price per soft drink (in cents), $P_{\text{softdrink}}$
 X_4 = Urban dummy variable
 X_5 = Residential dummy variable

Y , or the quantity demanded, is called the **dependent variable**. The X variables are referred to as the **independent** or **explanatory variables**. It is important to note the unit of measurement used for each variable. The researcher may choose how to record the data for use in regression analysis. Here we are measuring the prices of pizza and soft drinks in cents and tuition in thousands of dollars. Notice, too, that the unit of measurement for the location and residential variables are quite different from the others. “Urban” takes the value of 1 if the campus is located in an urban area and 0 otherwise. “Residential” takes the value of 1 if the college is a residential college and 0 otherwise. Both variables are **binary** or **dummy variables**. Given this particular setup of the regression equation and measurement scheme for the variables, we can now

estimate the values of the slope coefficients of the independent variables, b_i , as well as the intercept term, a , by using any one of the many available software packages containing regression analysis.

Estimating and Interpreting the Regression Coefficients

Among the software packages used by economists to conduct a regression analysis of the demand for a good or service are SPSS, SAS, and R (an open-source software). To estimate the demand for pizza, we employed the regression function contained in Excel. Although it only contains the basic elements of regression (e.g., it does not provide a Durbin-Watson statistic), we believe it is perfectly suitable for many types of regression analysis that would be conducted in business research. Besides, Excel is more available in both businesses and colleges and universities than are statistical software packages.

Using Excel to perform a regression analysis of the data in Table 5.3, we obtain the following output (see Excel Sheet 5.1).

Based on this output, we can express the following regression equation.

$$Q_{\text{slice}} = 29.16 - 0.093P_{\text{slice}} + 0.060Tuition - 0.082P_{\text{softdrink}} - 0.747Urban - 3.049Residential$$

(2.201) (0.042) (0.035) (0.013) (0.575) (0.493)

$$R^2 = 0.885 \quad \bar{R}^2 = 0.860 \quad F = 36.8 \quad \text{Standard Error of Estimate, SEE} = 1.06$$

The standard error of each coefficient, SEC, is listed beneath each coefficient in parentheses.

Before interpreting these results, we should first think about what direction of impact changes in the explanatory variables are expected to have on the demand for pizza as evidenced by the signs of the estimated regression coefficients. To put it more formally, we can state the following hypotheses about the



SUMMARY OUTPUT		<i>This is an example of Excel's regression output.</i>					
		<i>It was used to create the demand equation in the text.</i>					
<i>Regression Statistics</i>		<i>The only thing done to this output was to highlight the cells reported in that regression equation.</i>					
Multiple R	0.9405016						
R Square	0.8845432	<i>This output undoubtedly looks complex on first blush. Do not try to understand all of it at first. All you need to see initially is where the reported numbers come from.</i>					
Adjusted R	0.8604897						
Standard Er	1.0641804						
Observation	30	<i>Reg1 on the companion website is based on the pizza demand data in Table 5.3. Open this file to learn more about Excel's regression output or to run this yourself.</i>					
ANOVA							
		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>	
Regression	5	208.22903	41.645806	36.773994	1.718E-10		
Residual	24	27.179516	1.1324798				
Total	29	235.40855					
		<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	29.161319	2.2009619	13.249352	1.566E-12	24.618757	33.703881	
Pslice	-0.092571	0.0118269	-7.827178	4.635E-08	-0.116981	-0.068162	
Tuition	0.0597557	0.034522	1.7309453	0.0963003	-0.011494	0.1310056	
Psoftdrink	-0.081852	0.0128993	-6.345497	1.466E-06	-0.108475	-0.055229	
Urban	-0.746844	0.5750193	-1.298815	0.2063484	-1.933625	0.439938	
Residential	-3.04912	0.4925814	-6.190083	2.139E-06	-4.065758	-2.032482	

Excel Sheet 5.1

Source: Multivariate Regression Excel App, Sheet 5.1 worksheet

anticipated relationship between each of the explanatory variables and the demand for pizza:

- Hypothesis 1:** The price of pizza (X_1) is an inverse determinant of the quantity of pizza demanded (i.e., the sign of the coefficient is expected to be negative).
- Hypothesis 2:** Assuming tuition to be a proxy for income, pizza could be either a “normal” or an “inferior.” Therefore, we hypothesize that tuition (X_2) is a determinant of the demand for pizza, but we cannot say beforehand whether it is an inverse or a direct determinant (i.e., the sign of the coefficient could be either positive or negative).
- Hypothesis 3:** The price of a soft drink (X_3) is an inverse determinant of the demand for pizza (i.e., the sign of the coefficient is expected to be negative).
- Hypothesis 4:** Location in an urban setting (X_4) is expected to be an inverse determinant of the demand for pizza.
- Hypothesis 5:** Pizza demand is expected to be lower at residential colleges (X_5) than nonresidential colleges.

Turning now to the regression results, we observe that the P_{Slice} coefficient has a negative sign, and this is exactly what we would expect because of the law of demand. As the price of pizza changes, the quantity demanded for pizza will change in the opposite direction. This is what a negative slope coefficient tells us. The positive sign of the tuition coefficient tells us that tuition costs and quantity of pizza demanded are directly related to each other. Higher tuition costs are associated with a greater demand for pizza, and vice versa. Thus, pizza appears to be a “normal” product. The negative sign of the soft drink price confirms the complementarity between soft drinks and pizza. As the price of a soft drink goes up, college students tend to buy less pizza. The opposite would hold true for a reduction in the price of a soft drink. The negative sign of the location dummy variable tells us that those students who attend colleges in urban areas will buy, on average, about three quarters of a slice less per month (i.e., 0.747) than their counterparts in the suburbs or rural areas. Finally, the residential dummy variable tells us that students at residential colleges buy, on average, about three fewer slices per month (i.e., 3.049) than students at nonresidential colleges.

An interpretation of the magnitudes of the estimated regression coefficients of continuous variables is a bit more involved. Each estimated coefficient tells us how much the demand for pizza will change relative to a unit change in each of the explanatory variables. For example, a b_1 of -0.093 indicates that a unit change in price will result in a change in the demand for pizza of -0.093 in the opposite direction. Price, as you will recall, was measured in cents. Therefore, according to our regression estimates, a 100-cent (or \$1.00) increase will result in a decrease in the quantity demanded for pizza of 9.3 (100×0.093). A tuition increase of one unit (in this case \$1,000) results in an increase in the quantity demanded for pizza of 0.060. Are these changes and those associated with changes in the price of soft drinks, residential status of the college, and the location of the college campus substantial or inconsequential?

Researchers who are constantly estimating the demand for a particular good or service will have a fairly accurate idea whether the magnitudes of the coefficients estimated in a particular study are high or low relative to their other work. But if there are no other studies available for comparison, then researchers can at least use the elasticities of demand to gauge the relative impact that the explanatory variables have on the quantity demanded.

From our discussion of elasticity in Chapter 4, you can see that regression analysis results are ideal for point-elasticity estimation. Recall that the formula for computing point elasticity is

$$\epsilon_x = \frac{\partial Q}{\partial X} \cdot \frac{X}{Q}$$

where Q = quantity demanded and X = any variable that affects Q (e.g., price or income). To estimate elasticity, we need an estimate of quantity demanded. Suppose we are interested in estimating demand at a nonresidential college campus in an urban setting where tuition is \$14,000 per year and the average price of a slice of pizza is \$1.00 and the average price of a soft drink is \$1.10. Estimated demand in this instance is:

$$\begin{aligned} Q_{\text{slice}} &= 29.16 - 0.093 \cdot (100) + 0.060 \cdot (14) - 0.082 \cdot (110) - 0.747 \cdot (1) - 3.049 \cdot (0) \\ &= 10.99 \text{ (or about 11 slices per month)} \end{aligned}$$

To compute the point elasticities for each variable assuming the preceding values, we simply plug in the appropriate numbers into the point-elasticity formula. The partial derivative of Y with respect to changes in each variable (i.e., $\partial Y/\partial X$) is simply the estimated coefficient of each variable.

$$\begin{aligned} \text{Price elasticity: } & -0.093 \times \frac{100}{10.99} = -0.842 \\ \text{Tuition elasticity: } & 0.138 \times \frac{14}{10.99} = 0.076 \\ \text{Cross-price elasticity: } & -0.082 \times \frac{110}{10.99} = -0.819 \end{aligned}$$

With these estimates, we can say that the demand for pizza is somewhat price inelastic and that there is some degree of cross-price elasticity between soft drinks and pizza. Judging from the rather low elasticity coefficient of 0.076, tuition does not appear to have that great an impact on the demand for pizza.

Statistical Evaluation of the Regression Results

Our regression results are based on a *sample* survey of colleges across the country. How confident are we that these results are truly reflective of the *population* of college campuses? The basic test of the statistical significance of each estimated regression coefficient is called the ***t*-test**. Essentially, this test is conducted by computing a ***t*-value** or ***t*-statistic** for each estimated coefficient. This is done by dividing the estimated coefficient by its **standard error of coefficient, SEC**.⁶ That is:

$$t = \frac{\hat{b}}{\text{SEC of } \hat{b}}$$

As is the common practice in presentations of regression results, the standard errors in our pizza regression are presented in parentheses under the estimated coefficients. (Other researchers provide t statistics in parentheses. In this event, SEC is calculated as $\text{SEC} = \hat{b}/t$.) To interpret the value of t , we use the ***t*-table**. The convention in economic research is to select the .05 level of significance. This means you can be 95 percent confident that the results obtained from the sample are representative of the population. We also need to know the number of **degrees of freedom (df)** involved in the estimate. The term *degrees of freedom* is defined as $n - k - 1$, where “ n ” represents the sample size and “ k ” the number of independent variables. The “1” represents the constant or intercept term. Therefore, in our pizza example, we have $30 - 5 - 1$, or 24 degrees of freedom.

Turning to the t -table shown in Table A.4 in Appendix A, we see that the **critical *t*-value** at the .05 level of significance is 1.711 using a **one-tail test** and 2.064 using a

⁶In the following equation, the little “hat” (circumflex) over the b is a commonly used notation in statistical analysis to denote a value estimated from sample data. Similarly, \hat{Y} with a “hat”, \hat{Y} , means the estimated value of Y .

two-tail test.⁷ If the t -value computed for a particular estimated coefficient is greater than 1.711, we can say that the estimate is “significant at the .05 level” using a one-tail test. If it is greater than 2.064, then the same can be said, but with a two-tail test. A simple and useful way to handle the critical level is to use the **rule of 2**. This means that if the absolute value of t is greater than 2, we can conclude that the estimated coefficient is significant at the .05 level.

It is evident from the preceding regression equation that X_1 (price of pizza), X_3 (price of soft drinks), and X_5 (residential) are statistically significant because the absolute values of their t -statistics are 7.8, 6.3, and 6.2, respectively. The other two variables, X_2 (tuition) and X_4 (urban location), are not statistically significant because the absolute values of their t -statistics are less than 2.

If the estimated coefficient of a variable passes the t -test, we can be confident that the variable truly has an impact on demand. If it does not pass the t -test, then in all likelihood, the variable does not truly have an impact for the whole population of college campuses. In other words, the regression coefficients are nonzero numbers simply because of a fluke in the sample of campuses that we took from the population.

In statistical analysis, the best we can hope for is to be confident that our sample results are truly reflective of the population that they represent. However, we can never be absolutely sure. Therefore, statistical analysts set up degrees of uncertainty. As explained in greater detail later in this chapter, using the rule of 2 generally implies a 5 percent level of significance. In other words, by declaring a coefficient that passes the rule of 2 version of the t -test to be statistically significant, we leave ourselves open to a 5 percent chance that we may be mistaken.

Another important statistical indicator used to evaluate the regression results is the **coefficient of determination** or R^2 . This measure shows the percentage of the variation in a dependent variable accounted for by the variation in all the explanatory variables in the regression equation. (See the Geometric Introduction to Univariate Regression Excel App, How Regression Works (optional) worksheet for details.) This measure can be as low as 0 (indicating that the variations in the dependent variable are not accounted for by the variation in the explanatory variables) and as high as 1.0 (indicating that all the variation in the dependent variable can be accounted for by the explanatory variables). For statistical analysts, the closer R^2 is to 1.0, the greater the explanatory power of the regression equation.

In our pizza regression, $R^2 = 0.885$. This means that about 88 percent of the variation in the demand for pizza by college students can be accounted for by the variation in the price of pizza, the cost of tuition, the price of a soft drink, and the location of the college. As explained later in this chapter, R^2 increases as more independent variables are added to a regression equation. (See the Multivariate Regression Excel App, Adjusted R^2 worksheet for details.) Therefore, most analysts prefer to use a measure that adjusts for the number of independent variables used so equations with different numbers of independent variables can be more fairly compared. This alternative measure is called the **adjusted R^2** , or \bar{R}^2 . As noted earlier, $\bar{R}^2 = 0.86$ for the pizza demand equation.

Another test, called the **F -test**, is often used in conjunction with R^2 . This test measures the statistical significance of the entire regression equation rather than of each individual coefficient (as the t -test is designed to do). In effect, the F -test is a measure of the statistical significance of R^2 . The procedure for conducting the F -test is similar to the t -test. A critical value for F is first established, depending on the degree of statistical significance that the researcher wants to set (typically at the .05 or .01 level). The **critical F -values** corresponding to these acceptance levels are shown in Table A.3 in Appendix A. As can be seen, there are two **degrees of freedom** values that must be



⁷Please refer to a statistics text for a complete explanation of the one-tail and two-tail tests. We provide a brief explanation of these terms in the “Important Concepts” section at the end of this chapter.

incorporated in the selection of the critical F -value. These values essentially relate to the sample size and number of independent variables in the equation, and the sample size minus the number of independent variables plus the intercept of the equation. Therefore, because the pizza example has a sample size of 30 and five independent variables, the degrees of freedom are 5 and 24 ($30 - 5 - 1$). Table A.3 shows that at the .05 level, the critical F -value with those degrees of freedom is 2.62. At the .01 level, the critical value is 3.90. Because the regression results for the demand for pizza indicates an F -value of 36.8, we can conclude that our entire equation is **statistically significant** at the .01 level.

Review of Key Steps for Analyzing Regression Results

We now review all the key steps discussed so far in the regression analysis of a demand equation using the following stylized regression equation:

$$Q = 70 - 10P + 4P_R + 50I \quad \begin{matrix} \text{coefficient} \\ \text{(SEC)} \end{matrix} \quad \begin{matrix} \bar{R}^2 = 0.47 \\ n = 56 \end{matrix}$$

$$(3) \quad (1.8) \quad (30)$$

- where Q = Quantity of the product demanded
- P = Price of the product (in cents)
- P_R = Price of a related product (in cents)
- I = Per capita income (in thousands of dollars)
- n = Sample size
- \bar{R}^2 = Adjusted multiple coefficient of determination

Step 1: Check Signs and Magnitudes

The negative sign for the P variable indicates an inverse relationship between price and the quantity demanded for the product. A unit *increase* in price (i.e., 1 cent) will cause the quantity to *decrease* by ten units. A unit *decrease* in price will cause the quantity to *increase* by ten units. So, for example, if the price was decreased by \$1.00, quantity would increase by 1,000 units.

The positive sign for the P_R variable indicates a direct relationship between the price of a related product and the quantity demanded. This indicates that the related product is a *substitute* for the product in question. For example, if the price of the related product changes by one unit (i.e., 1 cent), then the quantity demanded of the product in question will change by four units in the *same direction*.

The positive sign for the I variable indicates that the product is *normal* or perhaps *superior*, depending on the magnitude of the income elasticity coefficient. A unit change in per capita income (i.e., \$1,000) will cause the quantity to change by fifty units in the *same direction*.

Step 2: Compute Elasticity Coefficients

To compute elasticity coefficients, we need to assume certain levels of the independent variables P , P_R , and I . Let us say they are as follows:

$$P = 100 \text{ (remember, this is 100 cents or \$1.00)}$$

$$P_R = 120 \text{ (also in cents)}$$

$$I = 25 \text{ (this represents \$25,000)}$$

Inserting these values into the previous equation gives us

$$Q = 70 - 10(100) + 4(120) + 50(25)$$

$$Q = 800$$

We now use the formula for point elasticity to obtain the elasticity coefficients. Recall that

$$\epsilon_x = \frac{dQ}{dX} \cdot \frac{X}{Q}$$

Using this formula, we obtain

$$\begin{aligned}\epsilon_p &= -10 \cdot \frac{100}{800} \\ &= -1.25 \\ \epsilon_{P_R} &= 4 \cdot \frac{120}{800} \\ &= 0.6 \\ \epsilon_I &= 50 \cdot \frac{25}{800} \\ &= 1.56\end{aligned}$$

Step 3: Determine Statistical Significance

Using the “rule of 2” as an approximation for the .05 level of significance, we can say that P and P_R are statistically significant because their t -values are both greater than 2 (e.g., 3.3 and 2.2, respectively). I is not statistically significant at the .05 level because its t -value is only 1.67.

As an added consideration, we note that the \bar{R}^2 of 0.47 indicates that 47 percent of the variation in quantity can be accounted for by variations in the three independent variables P , P_R , and I . Although this is not actually an indication of statistical significance, it does show the explanatory power of the regression equation. For cross-section data, this R^2 level can be interpreted as being moderately high.

Implications of Regression Analysis for Management Decisions

In our experience, the “bottom line” in the business world of any statistical analysis, including regression analysis, is the extent to which the results can help managers make good decisions. In our pizza example, the results indicate that the price of pizza and the price of its complementary product, the soft drink, are key factors influencing the demand for pizza. Their elasticity coefficients are both less than 1 and both variables’ coefficients passed the t -test. What does this mean for those in the pizza business? First, it means that they can expect price decreases to lead to decreases in revenue, other factors remaining constant. Therefore, they would probably not want to try lowering price in an effort to increase sales. But they could try lowering the price of soft drinks, with the anticipation that the lower price of the soft drink will attract people to buy the pizza.

In statistical analysis, it is often as important to find out what does *not* pass the t -test as much as it is to find out what passes. In our example, we learned that tuition and location do not have statistically significant impacts on pizza demand. Moreover, the magnitudes of their coefficients were relatively small. For managers of national chains such as Pizza Hut or Domino’s, this would indicate that they would not have to be very concerned about the type of college (private or public) or its location (urban or rural) in deciding where to open pizza franchises. They will, however, face competition from dining halls at residential colleges.

It is hoped that this summary is sufficient for those instructors and readers who simply want a general idea of how regression analysis can be used in business analysis and decision making. For a more detailed discussion, continue with the rest of this chapter and refer to the regression Excel Apps.



APP

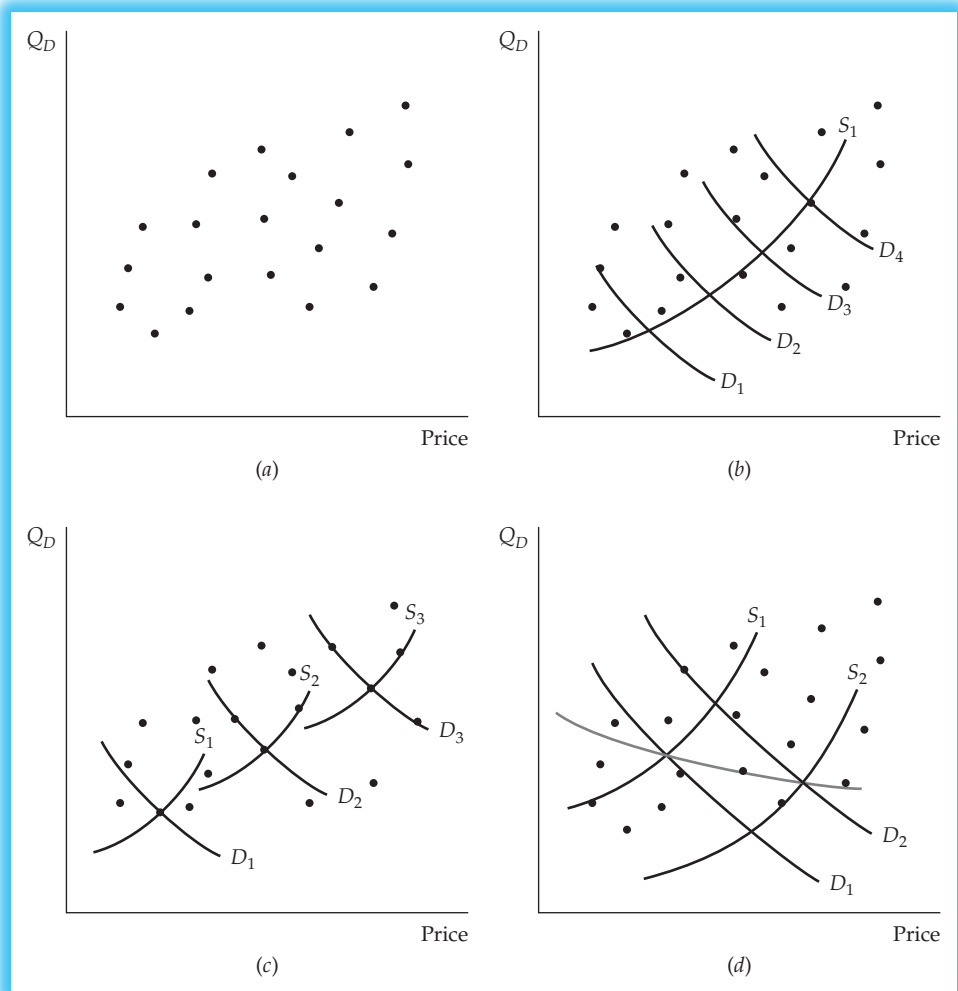
PROBLEMS IN THE USE OF REGRESSION ANALYSIS

Sometimes, even the best data and most powerful computer software cannot compensate for certain statistical problems arising from the efforts to measure the quantitative relationship between the amount demanded and the determinants of demand. We briefly address three of the most important problems.

The Identification Problem

The identification problem presents perhaps the greatest challenge to those using regression analysis to estimate the demand for a particular good or service. To explain this problem, let us return to our pizza example. Suppose we had time series data relating the per capita consumption of pizza with the price of pizza over a 20-year period. The scatter plot of this information is shown in Figure 5.1a. Notice that the scatter tends to slope upward and that a least squares regression estimate would reflect this pattern of relationship. Does this mean that the consumers of pizza behave irrationally and demand more pizza at higher prices? Common sense would balk at such a conclusion, but then why the positive coefficient of the price variable in the demand equation? The alert reader will state that what we have identified as a demand equation is probably some sort of supply equation or perhaps the result of the movement in *both* supply and demand for pizza over the past 20 years. As can be seen in Figure 5.1b, if the supply remained constant over the past 20 years while demand shifted upward (because of changes in such factors as income, number of buyers, and tastes and preferences over this time period), the regression equation would really be a reflection of the supply, supply curve S_1 . If the supply increased but demand increased more than the supply,

Figure 5.1 The Identification Problem



then the regression estimate would really be a reflection of the intersection of the various S and D curves in Figure 5.1 *c*. Figure 5.1 *d* shows still another possibility. In this case, supply shifts more than demand so the estimated regression line is downward sloping and more like what we would expect of a demand curve. Nonetheless, this estimated demand curve is flatter than the true demand curves, which have gradually shifted to the right over the years. Thus, the regression estimate of the relationship between price and quantity demanded would be biased in the sense that it would indicate a much greater price elasticity than actually exists in the population of pizza consumers.

There are advanced estimation techniques, such as the methods of *two-stage least squares* and *indirect least squares*, that can help the researcher deal with samples in which the simultaneous shifting of demand and supply takes place. Essentially, these techniques involve the simultaneous consideration of the supply and demand equations with the use of a single regression equation. A discussion of these techniques lies outside the scope of this text. But the principal point to remember is that if the identification problem is not recognized and dealt with by the researcher, the method of ordinary least squares will result in biased estimates of the regression coefficients.

Multicollinearity

One of the key assumptions made in the construction of the multiple regression equation is that the independent variables are not related to each other in any systematic way. If this assumption is incorrect, then each estimated coefficient may give a distorted view of the impact of the change in each independent variable. For example, suppose a regression model states that the demand for luxury foreign-made automobiles depends on price, income, and education. The latter variable is included because education is a proxy for tastes and preferences, and those with higher levels of education are hypothesized to have a greater preference for luxury foreign cars. But, as you would expect, education and income are closely associated. If their values tend to move up and down together, the least squares method may arbitrarily assign a high value to the coefficient of one variable and a low coefficient value to the other. In effect, if two variables are closely associated, it becomes difficult to separate out the effect that each has on the dependent variable. The existence of such a condition in regression analysis is called *multicollinearity*.

If the regression results pass the F -test (the measure of the overall statistical significance of the regression equation) but fail the t -test for each regression coefficient, it is usually a sign that multicollinearity is present in the sample data. Multicollinearity can also be detected by examining the correlation coefficient between two variables suspected of being closely related.⁸ As a rule of thumb, correlation coefficients of 0.7 or more provide a basis for researchers to suspect the existence of multicollinearity.

If multicollinearity is a serious problem in the regression analysis, it will tend to introduce an upward bias to the standard errors of the coefficients. This will tend to reduce the t -values (which, you will recall, are computed using the standard errors of the coefficients). This makes it harder to reject the null hypothesis and, of course, to identify statistically significant independent variables in the regression model.

It should be pointed out, however, that if the researcher simply wants to use the estimated regression coefficients as a basis for forecasting future values of the dependent variable, multicollinearity does not pose a serious problem. It is only when the researcher wants to understand more about the underlying structure of the demand function (i.e., what are the key determinants of demand) that this

⁸The correlation coefficient is a measure of the degree of association between two variables. This measure, denoted r , ranges from a value of -1 (perfect negative correlation) to 1 (perfect positive correlation).



APP

particular statistical problem should be resolved. Most software packages automatically produce a correlation coefficient matrix for the entire set of independent variables used in the regression equation. A standard remedy for multicollinearity is to drop one of the variables that is closely associated with another variable in the regression equation. For further discussion of multicollinearity, see the Multicollinearity Excel App.

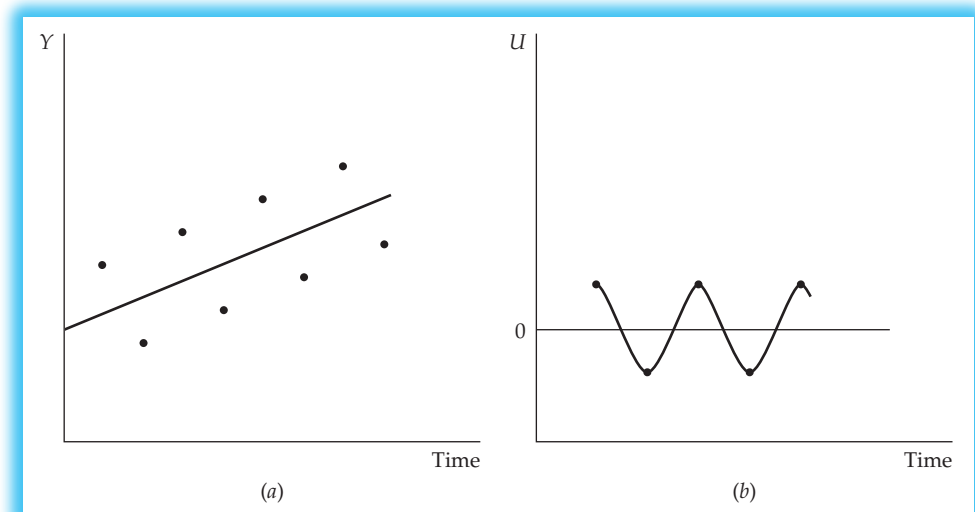


Autocorrelation

Autocorrelation is a problem that is usually encountered when time series data are used. For this reason, it is often referred to as *serial correlation*. Let us use the case of simple regression, involving only the dependent variable Y and one independent variable, time (X). Essentially, autocorrelation occurs when the Y variable relates to the X variable according to a certain pattern. For example, in Figure 5.2a, the scatter plot reveals that as time increases, the Y value deviates from the regression line in a very systematic way. In other words, the **residual term**, U (for unexplained), which is the difference between the observed value of Y and the estimated value of Y given X , \hat{Y} (i.e., $U = Y - \hat{Y}$), alternates between a positive and a negative value of approximately the same magnitude throughout the range of X values. In fact, if we were to plot these residuals on a separate graph, they would have the pattern shown in Figure 5.2b.

One possible cause of autocorrelation is that there are effects on Y not accounted for by the variables included in the regression equation. It might also be that the true relationship between Y and the independent variable(s) is nonlinear. See the Nonlinearities Excel App for a discussion of this topic. But regardless of the reason, if autocorrelation is present in the regression analysis, it creates a problem for the validity of the t -test. Simply stated, autocorrelation tends to increase the likelihood that the null hypothesis will be rejected. This is because autocorrelation gives a downward bias to the standard error of the estimated regression coefficient ($SE_{\hat{\beta}}$). Recalling that the t -value is defined as $\hat{\beta}/SE_{\hat{\beta}}$, we can see that a smaller $SE_{\hat{\beta}}$ would tend to increase the magnitude of the t -value, other factors held constant. Thus, in the presence of autocorrelation, researchers may well declare that certain independent variables have a statistically significant impact on the dependent variable when in fact they do not. From a policy standpoint, suppose the estimated coefficient of the advertising variable

Figure 5.2 Autocorrelation



in a regression model of demand passed the t -test when it really should not have. A firm might then be led to increase its advertising expenditures when in fact it should be looking at other ways to expand demand (e.g., through promotions, alternative channels of distribution, or price actions).

It may be difficult to identify autocorrelation simply by observing the pattern of the residuals of a regression equation. A standard test for identifying the presence of this problem is the *Durbin-Watson test*. The Durbin-Watson statistic (i.e., DW) is now routinely calculated in regression software packages and is presented automatically in the computer printout. As in the case of the t -test and the F -test, there is a Durbin-Watson table listing the critical values of this statistic for a given level of significance (usually the .05 level). We have included such a table in the appendix at the back of this text (see Table A.5 in Appendix A). As a rule of thumb, if the DW statistic is around 2, there is in all probability no autocorrelation present in the data. If the DW statistic indicates the presence of autocorrelation, there are certain things a researcher can do to correct the problem. These include transforming the data into a different order of magnitude or introducing leading or lagging data in the time series. The Serial Correlation and Durbin Watson Excel App provides further discussion of serial correlation and walks you through calculating and interpreting the Durbin-Watson statistic using Excel.

EXAMPLES OF REGRESSION ANALYSIS ACROSS THE DISCIPLINES

Regression analysis is perhaps the most widely used statistical tool for the analysis of empirical data in economics and business. But regression analysis is not restricted to economics and business. We will provide a brief glimpse of its use by academic researchers across the curriculum.

Probably the most practical application of regression analysis in economic research is in the estimation of consumer demand. Much of the data on consumer buying behavior, collected by such companies as ACNielsen and IRI, are proprietary. Sometimes these data are made available to researchers. Professor Stephen Erfle was able to obtain information on the consumer demand for wine collected by IRI and received permission to use this for his research. The results of his study using IRI data are presented in this chapter's appendix. We will also see regression modeling examples in the discussion of production in Chapter 6 and cost in Chapter 7.

Outside economics and business, regression analysis is widely used across academic disciplines. Due to the closeness of the other social sciences to economics, it is not surprising that regression is increasingly used in the other social science. This is especially the case for political science. Duchesne, Eagles and Erfle, for example, used regression models to examine the determinants of support for Quebec sovereignty in Canadian general elections.⁹

Outside the social sciences, its breadth is impossible to describe without spending a great number of pages. But here are two examples of the scope of its use: one in English and the other in medicine. A Google search of "regression analysis English literature" produced an article entitled "A regression model for the English benefactive

⁹Erich Duchesne, Munroe Eagles and Stephen Erfle, "Constituency Homogeneity, Economic Risk and Support for Quebec Sovereignty: A Research Note," *Canadian Journal of Political Science*, July/August 2003, 36:3, pp. 643-656.

alternation.”¹⁰ Researchers from Harvard Medical School used regression analysis to show that restricting the sale of sugar sweetened beverages from schools did reduce high school students’ consumption of such beverages. The abstract of that article is provided below.

“Effect of School District Policy Change on Consumption of Sugar-Sweetened Beverages Among High School Students, Boston Massachusetts, 2004–2006”¹¹

Abstract

Introduction Consumption of sugar-sweetened beverages has increased among youth in recent decades, accounting for approximately 13% of total calories consumed. The Boston Public Schools passed a policy restricting sale of sugar-sweetened beverages in Boston schools in June 2004. The objective of this study was to determine whether high school students’ consumption of sugar-sweetened beverages declined after this new policy was implemented.

Methods We conducted a quasi-experimental evaluation by using data on consumption of sugar-sweetened beverages by public high school students who participated in the Boston Youth Survey during February through April 2004 and February through April 2006 (N = 2,033). We compared the observed change with national trends by using data from the 2003–2004 and 2005–2006 National Health and Nutrition Examination Survey (NHANES). Regression methods were adjusted for student demographics.

Results On average, Boston’s public high school students reported daily consumption of 1.71 servings of sugar-sweetened beverages in 2004 and 1.38 servings in 2006. Regression analyses showed significant declines in consumption of soda (–0.16 servings), other sugar-sweetened beverages (–0.14 servings), and total sugar-sweetened beverages (–0.30 servings) between 2004 and 2006 ($P < .001$ for all). NHANES indicated no significant nationwide change in adolescents’ consumption of sugar-sweetened beverages between 2003–2004 and 2005–2006.

Discussion Data from Boston youth indicated significant reductions in consumption of sugar-sweetened beverages, which coincided with a policy change restricting sale of sugar-sweetened beverages in schools. Nationally, no evidence was found for change in consumption of sugar-sweetened beverages among same-aged youth, indicating that implementing policies that restrict the sale of sugar-sweetened beverages in schools may be a promising strategy to reduce adolescents’ intake of unnecessary calories.

GLOBAL APPLICATION: FOOD IN SPAIN, CIGARETTES IN TAIWAN

Example 1: Regression Analysis of Basic Foods in Spain

A time series study of the demand for six classes of food for the period 1964 to 1991 in Spain was performed. The six different classifications were

1. Bread and cereals
2. Meat
3. Fish
4. Milk, dairy products, and eggs
5. Fruits, vegetables, and potatoes
6. Fats and oils

¹⁰Daphne Theijssen, Hans van Halteren, Karin Fikkers, Frederike Groothoff, Lian van Hoof, Eva van de Sande, Jorieke Tiems, Véronique Verhagen and Patrick van der Zande, “A regression model for the English benefactive alternation,” *Proceedings of the 19th Meeting of the Computational Linguistics in the Netherlands*, 2009, Utrecht, The Netherlands, pp. 115–130.

¹¹Angie L. Cradock, Anne McHugh, Helen Mont-Ferguson, Linda Grant, Jessica Barrett, Claire Wang, and Steven Gortmaker. *Preventing Chronic Disease*, July 2011, 8:4, A74, pp. 1–9. Available at http://www.cdc.gov/pcd/issues/2011/jul/10_0149.htm

The results show that changes in food demand (in terms of calories consumed) do not depend only on current income and prices, but also on lagged values of these variables. Consumers of food in Spain do not immediately adjust their expenditures when incomes or prices change. Rather they appear to show a certain level of inertia. Of the above six categories, only meat is shown to be a “superior” good, with an income elasticity of 1.54. Among the other five groups, only fruits and vegetables have income elasticity close to unity (.9), while the elasticity of fats and oils is the lowest, at 0.35. In terms of own price elasticities, meat shows the highest coefficient of -0.8 . All the others are also negative; all have t -test results that are significantly different from zero. Most cross-elasticities are quite low and not statistically significant, showing that there is no great substitutability among these products. There are a few exceptions, the largest one being fish and meat, whose t -test result is significant. The R^2 for the equation is a relatively high 0.63.¹²

Example 2: Regression Analysis of the Demand for Cigarettes in Taiwan

A recent time series study investigated the demand for cigarettes in Taiwan, not only as a function of income and price, but also looked into the influence of antismoking campaigns and the impact of the opening of Taiwan’s market to imported cigarettes. Data for 30 years (1966–1995) were used. The basic estimating equation was written as follows:

$$C_t = a_0 + a_1P_t + a_2Y_t + a_3H_t + a_4X_t + e_t$$

where C_t = Annual cigarette consumption per capita,

P_t = Average retail price of cigarettes per pack,

Y_t = Disposable income per capita, in 1991 New Taiwan dollars,

H_t = Measure of smoke-related health information,

X_t = Other determinants,

e_t = Random error term, and

a_{0-4} = Regression coefficients

In addition, separate demand functions were estimated for domestic and imported cigarettes. The equations used in these calculations were similar to the previous one, except that one term was added to each equation. In the domestic cigarette demand equation, the price of imported cigarettes was included to measure cross-price elasticity. In the equation for imported cigarettes, the price of domestic cigarettes was included. Two measures of health information were used:

1. Market share of low-tar cigarettes
2. Strongly worded warning labels adopted in 1992. A dummy variable of 0 up to 1991 and 1 thereafter was employed.

Some of the other explanatory variables included lagged consumption, market share of imported cigarettes, and female labor force participation.

The results for the overall model were as follows:

Price elasticity -0.5 to -0.6

Income elasticity 0.14 to 0.22

Low-tar impact -0.04

¹²A. Gracia, J. M. Gil, and A. M. Angulo, “Spanish Food Demand: A Dynamic Approach,” *Applied Economics*, October 1998, 30(10), pp. 1399–1405.

All three of these had statistically significant coefficients at the 1 percent or 5 percent level. The effect of warning labels was also negative but not statistically significant, and the impact of imported cigarettes was positive but only significant at the 10 percent level in some of the models.

When domestic and imported cigarettes were considered separately, the results were similar. The cross-price elasticities in both equations were positive, indicating that domestic and imported cigarettes are substitutes. The coefficients of determination for the various models were quite satisfactory. They were approximately 0.91 for the overall equations and 0.71 to 0.88 for the equations in which domestic and imported cigarettes were studied separately.¹³

FORECASTING

INTRODUCTION

One of the authors remembers a poster he saw, many years ago, on the wall of the office of the director of market research of a large manufacturing corporation. It said: “Forecasting is very difficult, especially into the future.” One could add: “Accurate forecasting is even more difficult.” Certainly, there is a great deal of truth in this statement. But despite the difficulty of forecasting, and forecasting accurately, it is an integral part of our lives. Many of us eagerly watch television weather forecasts or pay heed to the predictions of a favorite stock market guru, knowing only too well how inaccurate they may be. In an even less formal sense we make forecasts when we buy a lottery ticket, bet on a horse, or decide whether to carry an umbrella when we leave the house.

In the worlds of business, government, or even nonprofit institutions, forecasting becomes even more important. In a world where organizations and their environments are continually becoming more complex and changes occur more rapidly, decision makers need help in weighing many factors and understanding constantly changing relationships to arrive at decisions whose results have ever-increasing impacts.¹⁴ Using existing resources and acquiring additional resources appropriately requires maximum information about the company’s future.

All organizations conduct their activities in an uncertain environment, and probably the major role of forecasting is to reduce this uncertainty. But no forecast, however extensive and expensive, can remove it completely. Managers who use forecasts in their work “need to develop realistic expectations as to what forecasting can and cannot do.”¹⁵ “Forecasting is not a substitute for management judgment in decision making; it is simply an aid to that process.”¹⁶

SUBJECTS OF FORECASTS

Business uses forecasting to obtain information about many subjects. In the final analysis, firms are interested in future sales and profits—the bottom line. But to get there, a large number of forecasts may be necessary. In this section of the chapter, we outline the various categories of forecasts—from the most macro forecasts to individual series.

¹³Chee-Ruey Hsieh, Teh-Wei Hu, and Chien-Fu Jeff Lin, “The Demand for Cigarettes in Taiwan: Domestic vs. Imported Cigarettes,” *Contemporary Economic Policy*, April 1999, pp. 223–234.

¹⁴The introductory sections of this chapter draw heavily on Steven C. Wheelwright and Spyros Makridakis, *Forecasting Methods for Management*, 5th ed., New York: John Wiley & Sons, 1989.

¹⁵*Ibid.*, p. 44.

¹⁶*Ibid.*, p. 30.

- Forecasts of gross domestic product—which describe the total production of goods and services in a country.
- Forecasts of the components of GDP—for example, consumption expenditure, producer durable equipment expenditure, and residential construction.
- Industry forecasts—for Global Foods, this would represent forecasts of sales of soft drinks, bottled water, and its other products.
- Forecasts of sales for a specific product—for instance, diet cola.

Demand Estimating and Demand Forecasting

In the first part of this chapter, regression analysis and demand estimation are discussed. There is a great deal of similarity between demand forecasting and demand estimating. The difference between the two lies largely in the ultimate purpose of the analysis.

The demand estimating technique will be used by a manager interested in probing the effect on the demand (or quantity demanded) of a change in one or more of the independent variables. Thus, a pricing manager may want to know the impact on the sales of the company's club soda as a result of a price change, of a competitor's price changes, or, for example, of a change in the company's advertising expenses.

In contrast, forecasting puts less emphasis on explaining the specific causes of demand changes and more on obtaining information regarding future levels of sales activity, given the most likely assumptions about the independent variables. Indeed, in some cases, to be discussed later in this chapter, forecasting is achieved without the introduction of any causal factors; future sales are predicted solely by projecting the past into the future.

PREREQUISITES OF A GOOD FORECAST

Certain conditions should be met by any good forecast:

- A forecast must be consistent with other parts of the business. For instance, a sales forecast of 10 percent growth must ensure there are sufficient manufacturing facilities and labor force to produce this increase.
- A forecast should be based on knowledge of the relevant past. However, when underlying conditions have changed significantly, past experience may be of no help in making a forecast. Moreover, sometimes there is no past on which to rely. This is the case when we are dealing with a new product or technology. Under these circumstances, analysts' judgments must be injected into the forecasting process. In some cases, "forecasts based purely on the opinion of 'experts' are used to formulate the forecast or scenario for the future."¹⁷
- A forecast must consider the economic and political environment, as well as any potential changes.
- A forecast must be timely. An accurate forecast that is too late to be acted on may be worthless.

FORECASTING TECHNIQUES

There are many different forecasting methods. One of the challenges facing a forecaster is choosing the right technique. The appropriate method depends on the subject matter to be forecast and on the forecaster, but we can discuss some of the factors that enter into consideration.¹⁸

¹⁷John E. Hanke, Dean W. Wichern, and Arthur G. Reitsch, *Business Forecasting*, 7th ed., Upper Saddle River, NJ: Prentice Hall, 2001, p. 421. This is another text we recommend for students who are interested in exploring the subject of forecasting in more detail.

¹⁸This section relies heavily on Wheelwright and Makridakis, *Forecasting Methods*, pp. 30–31.

1. The item to be forecast. Is one trying to predict the continuance of a historical pattern, the continuance of a basic relationship, or a turning point?
2. The interaction of the situation with the characteristics of available forecasting methods. The manager must judge the relation between value and cost. If a less expensive method can be used to achieve the desired results, it certainly should be.
3. The amount of historical data available.
4. The time allowed to prepare the forecast. Selection of a specific method may depend on the urgency of the situation.

One other point about forecasting cost and accuracy should be added here. Generally, when the requirements for forecast accuracy are high, more sophisticated and more complex methods may be used. Such methods are, as a rule, more costly. Thus, a manager will authorize greater expenditures when relatively high accuracy is warranted. However, “[e]mpirical studies have shown that simplicity in forecasting methods is not necessarily a negative characteristic or a detriment with regard to forecasting accuracy. Therefore, the authors would advise against discarding simple methods and moving too quickly to replace them with more complex ones.”¹⁹

Forecasting techniques can be categorized in many ways. We use the following six categories:

1. Expert opinion
2. Opinion polls and market research
3. Surveys of spending plans
4. Economic indicators
5. Projections
6. Econometric models

As we see in the following pages, some of the methods can be classified as qualitative, others as quantitative. **Qualitative forecasting** is based on judgments of individuals or groups. The results of qualitative forecasts may be in numerical form but generally are not based on a series of historical data.

Quantitative forecasting, in contrast, generally uses significant amounts of prior data as a basis for prediction. Quantitative techniques can be naive or causal (explanatory). **Naive forecasting** projects past data into the future without explaining future trends. **Causal or explanatory forecasting** attempts to explain the functional relationships between the variable to be estimated (the dependent variable) and the variable or variables that account for the changes (the independent variables).

Although judgmental (qualitative) forecasting is used frequently, the use of quantitative methods has been growing rapidly. A survey by the Institute of Business Forecasting reported that time series models that extrapolate past data into the future were used by 72 percent of the companies surveyed. Simple models (such as averages and simple trend) accounted for 60 percent of those using time series, and exponential smoothing (which we will discuss later in this chapter) accounted for 30 percent. “The reason why time series models are used most in businesses is because they are easier to understand and easier to use. Plus, they generally work well for short-term forecasting.” Causal forecasting methods were used by 17 percent and qualitative methods were used in 11 percent of the cases. Surveys were the most popular in the latter category. However, this survey also found that in many cases members of various

¹⁹Ibid., p. 309. This quote deals with the merits of relatively simple time series methods versus more complex explanatory techniques. “Thus, the evidence suggests that explanatory models do not provide significantly more accurate forecasts than time series methods, even though the former are much more complex and expensive than the latter” (p. 297).

functions (marketing, production, finance, and sales) of the companies met periodically to review the quantitative forecasts (often called baseline forecasts). “Where necessary, they collectively overlay judgment on the baseline forecast.” This was actually done by 83 percent of the companies surveyed. However, a somewhat earlier study of manufacturing companies concluded that judgmental methods are still prevalent in many business organizations.²⁰

Expert Opinion

Various types of techniques fit into the category of expert opinion. Only two are discussed here.

Jury of executive opinion: Forecasts are generated by a group of corporate executives assembled together. The members of the panel may represent different functions within the corporation (intraorganizational) or different corporations (interorganizational). This has been a rather successful technique. However, there is a danger that a panel member who is persuasive but not necessarily knowledgeable may have undue influence on the results. A similar method is to solicit the views of individual salespeople. However, “[o]ften sales people are either overly optimistic or pessimistic. At other times they are unaware of the broad economic patterns that may affect demand.”²¹

The Delphi Method: This method, developed by the Rand Corporation in the 1950s, is used primarily in predicting technological trends and changes. Delphi also uses a panel of experts. However, they do not meet. The process is carried out by a sequential series of written questions and answers. Although in the past such iterative procedure could have been quite time consuming, computers and e-mail have shortened the process significantly. The iterations are conducted until the range of answers is narrowed. Finally, a consensus (or convergence of opinions) is obtained. An early example of this method was a study that asked to forecast five subjects as far as 50 years into the future: scientific breakthroughs, population growth, automation, space programs, and future weapon systems.²²

Although the Delphi method has often been successful, it has drawbacks: “insufficient reliability, oversensitivity of results to ambiguity of questions, different results when different experts are used, difficulty in assessing the degree of expertise, and the impossibility of predicting the unexpected.”²³ Given the long-term nature of Delphi predictions, these criticisms do not appear to differ greatly from those leveled against forecasts in general.

Opinion Polls and Market Research

You are probably familiar with opinion polling because most of us have at one time or another been subjected to telephone calls or written questionnaires asking us to assess a product or sometimes a political issue. Rather than soliciting experts, opinion polls survey a population whose activity may determine future trends. **Opinion polls** can be very useful because they may identify changes in trends, which, as we see later

²⁰Chaman L. Jain, “Benchmarking Forecasting Practices in America,” *Journal of Business Forecasting*, Vol. 25, Issue 4, Winter 2006–07, pp. 9–13; Chaman L. Jain, “Benchmarking Forecasting Models,” *Journal of Business Forecasting*, Vol. 25, Issue 4, Winter 2006–07, pp. 14–17. Curiously, however, Hanke, Wichern, and Reitsch, *Business Forecasting*, p. 421, state that “research has shown that, when historical data are available, the judgment modification of the forecasts produced by analytical methods tends to reduce the accuracy of the forecasts.” Nada R. Sanders, “The Status of Forecasting in Manufacturing Firms,” *Production and Inventory Management Journal*, 38, 2, 1997, pp. 32–37.

²¹Wheelwright and Makridakis, *Forecasting Methods*, p. 242.

²²T. J. Gordon and O. Helmer, *Report on a Long-Range Forecasting Study*, Rand Corporation, P-2982, September 1964. (Described in Harold Sackman, *Delphi Critique*, Lexington, MA: Lexington Books, 1975, pp. 37–39, 104.)

²³Wheelwright and Makridakis, *Forecasting Methods*, p. 326

in this chapter, may escape detection when quantitative (both naive and explanatory) methods are used.

Choice of the sample population is of utmost importance because the use of an unrepresentative sample may give completely misleading results. Further, the questions must be stated simply and clearly. Often a question is repeated in a somewhat different form so the replies can be cross-checked.

Market research is closely related to opinion polling. Thorough descriptions of this method can be found in marketing texts. Market research will indicate “not only why the consumer is or is not buying (or is or is not likely to buy), but also who the consumer is, how he or she is using the product, and what characteristics the consumer thinks are most important in the purchasing decision.”²⁴ This information can then be used to estimate the market potential and possibly the market share.

Surveys of Spending Plans

The use of **surveys of spending plans** is quite similar to opinion polling and market research, and the methods of data collection are also quite alike. However, although opinion polling and market research usually deal with specific products and are often conducted by individual firms, the surveys discussed briefly here seek information about “macro-type” data relating to the economy.

1. **Consumer intentions.** Because consumer expenditure is the largest component of the gross domestic product, changes in consumer attitudes and their effect on subsequent spending are a crucial variable in the forecasts and plans made by businesses. Two well-known surveys are reviewed here.
 - a. *Survey of Consumers, Survey Research Center, University of Michigan.* This is probably the best known among the consumer surveys conducted. Initiated in 1946, it is conducted monthly. It contains questions regarding personal finances, general business conditions, and buying conditions. The answers to these questions are summarized into an overall index of consumer sentiment and a large number of indexes covering replies to the more detailed questions.
 - b. *Consumer Confidence Survey, The Conference Board.* A questionnaire is mailed monthly to a nationwide sample of 5,000 households. Each month a different panel of households is selected. The resulting indexes are based on responses to questions regarding business conditions (current and 6 months hence), employment conditions (current and 6 months hence), and expectations regarding family income 6 months hence. The replies are then summarized monthly into three indexes: the Consumer Confidence Index, the Present Situation Index, and the Expectations Index. This survey has been published since 1967.
2. **Inventories and sales expectations.** A monthly survey published by the National Association of Purchasing Agents is based on a large sample of purchasing executives.

Economic Indicators

The difficult task of predicting changes in the direction of activity is discussed previously. Some of the qualitative techniques discussed in the preceding sections are aimed at identifying such turns. The barometric technique of **economic indicators** is specifically designed to alert business to changes in economic conditions.

The success of the indicator approach to forecasting depends on the ability to identify one or more historical economic series whose direction not only correlates with, but also precedes that of the series to be predicted. Such indicators are used widely in forecasting general economic activity. Any one indicator series may not be

²⁴Ibid., p. 245.

very reliable; however, a composite of leading indicators can be used to predict. Such a series should exhibit a slowing (and an actual decrease) before overall economic activity turns down, and it should start to rise while the economy is still experiencing low activity.

Forecasting on the basis of indicators has been practiced in an informal fashion for many years. It is said that Andrew Carnegie used to assess the future of steel demand by counting the number of chimneys emitting smoke in Pittsburgh. Much of the work of establishing economic indicators was done at the National Bureau of Economic Research, a private organization. Today, economic indicator data are published monthly by The Conference Board in *Business Cycle Indicators*. These monthly data are reported in the press and are widely followed.²⁵

There are three major series: leading, coincident, and lagging indicators. As their names imply, the first tells us where we are going, the second where we are, and the third where we have been. Although the leading indicator series is probably the most important, the other two are also meaningful. The coincident indicators identify peaks and troughs, and the lagging series confirms upturns and downturns in economic activity.

Many individual series are tracked monthly in the *Business Cycle Indicators*, but only a limited number are used in the construction of the three major indexes. The leading indicator index contains ten series, and the coincident and lagging indicators are made up of four and seven components, respectively. All the series making up the indexes are listed in Table 5.4.

Table 5.4 Economic Indicators

Leading indicators

1. Average weekly hours, manufacturing
2. Average weekly claims for unemployment insurance
3. Manufacturers' new orders, consumer goods and materials industries
4. Index of supplier deliveries—vendor performance
5. Manufacturers' new orders, nondefense capital goods industries
6. New private housing units authorized by local building permits
7. Stock prices, 500 common stocks—S&P500
8. Real money supply (M2)
9. Interest rate spread, 10-year treasury bonds less federal funds rate
10. Consumer expectations, University of Michigan

Coincident indicators

1. Employees on nonagricultural payrolls
2. Personal income less transfer payments
3. Industrial production index
4. Manufacturing and trade sales

Lagging indicators

1. Average duration of unemployment
2. Ratio, manufacturing and trade inventories to sales
3. Change in index of labor cost per unit of manufacturing output
4. Average prime rate charged by banks
5. Commercial and industrial loans outstanding
6. Ratio, consumer installment credit outstanding to personal income
7. Change in consumer price index for services

²⁵The data can also be accessed at www.tcb-indicators.org.

It is rather evident why some of the indicators qualify as leading. They represent not present expenditures, but commitments indicating that economic activity will take place in the future. Among these are manufacturers' new orders and building permits. Others are not quite as obvious. But one would expect employers to increase the hours of their workforce as they increase production before committing themselves to new hiring. Stock market prices and money supply are usually believed to precede cycles.

The month-to-month changes in each component of the series are computed and standardized to avoid undue influence of the more volatile components. The individual series are then combined to create the index.²⁶ The Conference Board also computes a diffusion index that tends to indicate the breadth of the indicator index movement.

How good a forecaster are the leading indicators? To answer this question, we must establish some criteria. First, how many months of change in the direction of the index are necessary before a turn in economic activity is expected? A general rule of thumb is that if, after a period of increases, the leading indicator index sustains three consecutive declines, a recession (or at least a slowing) will follow. On this basis, the leading indicators have predicted each recession since 1948. Second, how much warning do the indicators give (i.e., by how many months do they lead) of the onset of a recession?

The following chart shows the lead times in number of months at both peaks and troughs of the business cycle since World War II.²⁷

Cycle	Peak	Trough
1948–49	–4	–4
1953–54	–6	–6
1957–58	–23	–2
1960–61	–10	–11
1969–70	–8	–7
1973–75	–9	–2
1980–80	–14	–2
1981–82	–8	–10
1990–91	–18	–2
2001–01	–11	–7
2007–09	–12	–3

It is obvious that the lead times vary considerably and that the lead times at peaks tend to be longer than those at troughs. So we can sum up by observing that leading indicators do warn us about changes in the direction of economic activity, but they really do not forecast lead times reliably. In addition, we must be aware of several other drawbacks:

1. In some instances, the leading indicator index has forecast a recession when none ensued.
2. A decline (or a rebound after a decline) in the index, even if it forecasts correctly, does not indicate the precise size of the decline (or rise) in economic activity.
3. The data are frequently subject to revision in the ensuing months. Thus, the final data may signal a different future from the one suggested by the originally published data.

The preceding criticisms of this forecasting method are certainly significant, but they indicate that this technique should be improved rather than discarded. The

²⁶An explanation of the methods to compute, update, and standardize the indexes is available on the Web site cited in footnote 25. An excellent source of information on the entire subject of economic indicators is *Business Cycle Indicators Handbook*, The Conference Board, 2001.

²⁷See any recent issue of *Business Cycle Indicators* for this information.

existing indicators are reevaluated periodically, and new ones are developed. As the structure of the economy changes, some of the indicators lose their relevance. For instance, in 1996, shortly after the Conference Board took over the publication of the indexes, two components of the leading indicator index were removed, while one series was added. The present index may still be overly weighted toward manufacturing activities and neglect service industries, which make up a continuously increasing portion of our economy. The growing importance of international trade and capital flows suggests the inclusion in the index of a series reflecting these activities. In the meantime, the use of leading indicators has spread to a large number of foreign countries.

In short, despite its drawbacks, the index of leading indicators (and the other two indexes as well) is a useful tool for businesspeople and will continue to be closely watched. As always, reliance on this method must be tempered by the knowledge of its imperfections.

Projections

We now turn to **trend projections**, which we previously identified as a naive form of forecasting. Several different methods are discussed, but they all have a common denominator: past data are projected into the future without taking into consideration reasons for the change. It is simply assumed that past trends will continue. Three projection techniques are examined here:

1. Compound growth rate
2. Visual time series projection
3. Time series projection using the least squares method

If annual data are to be forecast, any of these three methods can be used. However, more frequent data, such as monthly or quarterly, may be necessary. If there appear to be significant seasonal patterns in the data, a smoothing method must be applied. The moving-average method of smoothing is discussed, along with the least squares time series projection.



Module 5A

Constant Compound Growth Rate

The constant **compound growth rate** technique is extremely simple and is widely used in business situations. When quick estimates of the future are needed, this method has some merit. And, as we will find out, it can be quite appropriate when the variable to be predicted increases at a constant percentage (as opposed to constant absolute changes). But care must be exercised not to apply this technique when it is not warranted.



The most rudimentary application of this method is to take the first and last years of past data, and to calculate the constant growth rate necessary to go from the amount in the first period to the amount in the last. This problem is solved in the same way as if we were to calculate how much a specific sum deposited in an interest-bearing account grows in a certain number of years at a constant rate of interest that is compounded annually. This calculation can be made by using any handheld calculator or by using Table A.1 in Appendix A.

Frank Robinson will make the projection for Citronade using Table A.1 from Appendix A. The first and last period amounts are known, and the missing number is the growth rate. To solve for the growth rate, the formula is as follows:

$$(1 + i)^n = E/B$$

where E = Last year's amount
 B = First year's amount
 i = Growth rate
 n = Number of years

Applying this formula to the 2002 and 2012 sales of Citronade from Table 5.1,

$$6,954/3,892 = 1.7867$$

Because there are 10 years intervening between the two sales periods, the growth rate can be found by entering the table at ten periods and moving to the right until a number close to 1.7867 is found. At 6 percent, the factor is 1.7908. Thus, the growth rate of Citronade over the last 10 years is just under 6 percent. (One may also obtain this answer analytically by solving for i in the above equation: $i = (E/B)^{(1/n)} - 1 = 1.7867^{0.1} - 1 = 0.0598$.)

Therefore, can next year's sales be estimated at 7,371,000 cases, a 6 percent annual increase? Looking at the year-to-year increases in Table 5.1, Robinson noticed that the growth rate achieved in the early years (through 2006) has not been approached since. Whereas during the first 4 years of the available data, growth ranged between 6.5 and 8 percent, the second 4 years registered increases between 5 and 6 percent, and the rise in the last 2 years was less than 5 percent. Growth has been declining over the 10-year period.

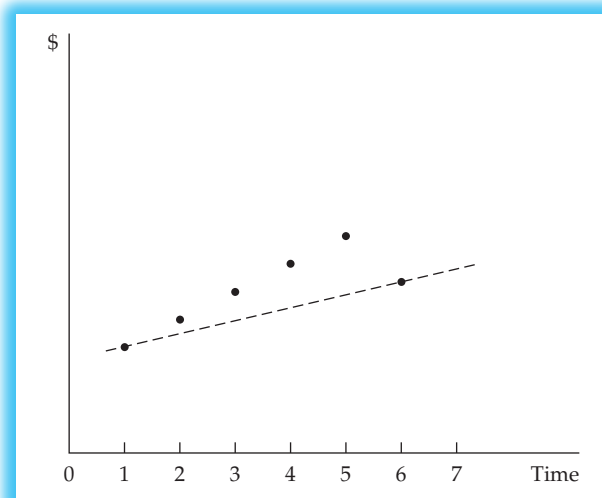
This illustrates a major problem that is frequently encountered when the compound growth rate method of projection is employed. The only two numbers considered in determining the growth rate were the first and the last; any trends or fluctuations between the original and terminal dates are disregarded. Thus, when year-to-year percentage growth is not stable, any estimates based on this result can be misleading.

Figure 5.3 illustrates what may happen. The data plotted in the graph show a strong increase in each year except for the last value, which may have been caused by an exceptional occurrence such as a severe recession or a strike. If the compound growth rate approach is used, any forecast based only on the first and last observations may be quite misleading.

Visual Time Series Projections

A series of numbers is often difficult to interpret. Plotting the observations on a sheet of graph paper (or entering paired data into Excel) can be very helpful because the shape of a complicated series is more easily discerned from a picture.

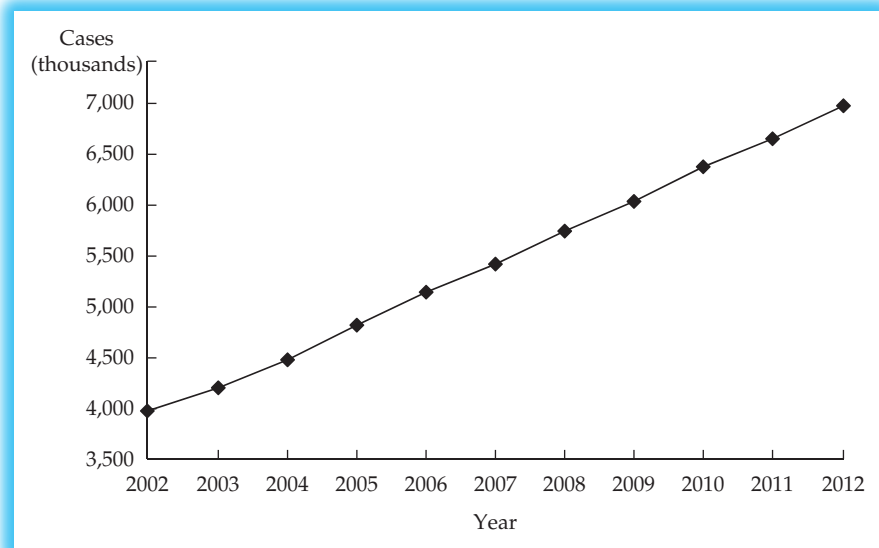
Figure 5.3 An Example in Which the Constant Compound Growth Rate Approach Would Be Misleading



Frank Robinson did just that. He had two types of graph paper available to him. One was the familiar kind, with arithmetic scales on both axes, which he used to plot the annual data of Citronade sales, as shown in Figure 5.4. The observations appear to form a relatively straight line. As a matter of fact, one could easily draw a straight line just by putting a ruler through the observations, so some points fall above and some below the line. Projecting the line to 2013 would give us a fairly good prediction—if the growth, in absolute number of cases, reflects past increases. The reason that a straight line is such a good fit is that the absolute annual increases in sales fluctuate around their mean of 306,000 cases without exhibiting any kind of trend.

There is another type of graph paper often used to observe trends—semilogarithmic. This graph has an arithmetic scale along the horizontal axis, but the vertical axis transforms numbers into a logarithmic format. Equal distances between numbers on the vertical scale represent a doubling. Thus, the distance between 1 and 2 is the same as that between 2 and 4, between 4 and 8, and so on.²⁸ In the semilogarithmic format, if observations that exhibit a constant growth rate are plotted, they will fall on a straight line.²⁹ The data for Citronade, as can be seen in Figure 5.5 where the

Figure 5.4 Annual Sales of Citronade (2002–2012)



²⁸Incidentally, the point of origin on the vertical axis is 1, not 0.

²⁹A simple example will explain why a constant percentage growth rate shows up as a straight line on a logarithmic scale. Assume the first period's quantity is 100 and it grows by 20 percent each year. Then translate each number into a common logarithm, and calculate the differences for each series:

Raw Data	Absolute Difference	Logarithm	Difference in Logarithms
100		2.0	
120	20	2.0792	0.0792
144	24	2.1584	0.0792
172.8	28.8	2.2375	0.0791
207.36	34.56	2.3167	0.0792

Although the absolute differences increase, the differences between logarithms remain constant. Thus, if the vertical scale of a graph is shown in terms of logarithms, constant percentage differences will plot as a straight line.

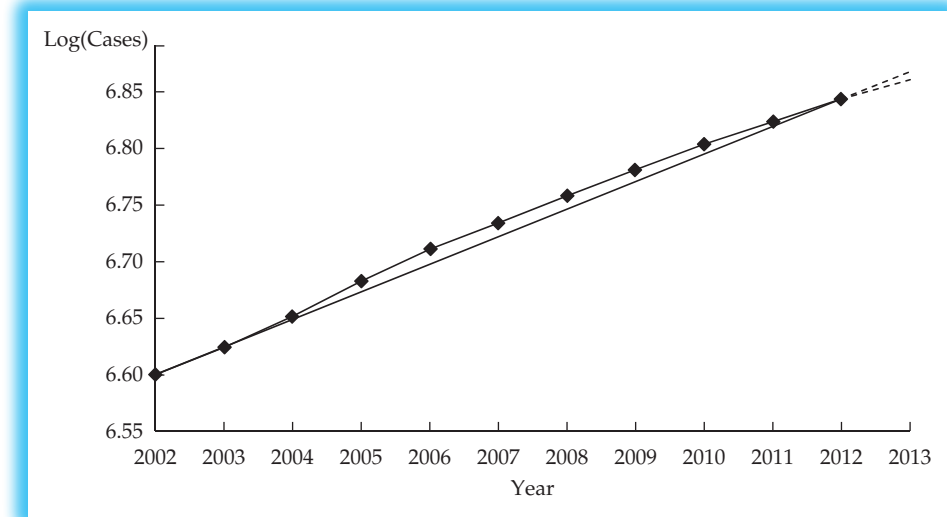


Figure 5.5 Semilogarithmic Graph of Annual Sales of Citronade (2002–2012)

annual observations are plotted, indicate growth at a decreasing rate. Thus, if Frank Robinson were to base his projections on this line, he would have to continue the curve on its decreasing slope.

Frank found that the compound growth rate for Citronade sales from 2002 to 2012 was 6 percent. A line representing an annual growth rate of 6 percent in Figure 5.5 is a straight line because constant growth is assumed. But a projection of the straight line would result in a higher sales estimate for 2013 than if the decreasing year-to-year percentage of growth were taken into consideration.

If we projected 2013 sales along the 6 percent growth line, our estimate would be 7,371,000 cases. Projecting sales by using the 10-year absolute average growth of 306,000 cases we would arrive at 7,260,000 cases, a 4.4 percent increase, a quantity that would closely correspond to the value that would be indicated by projecting the line showing actual observations. The latter would appear to be the more reasonable estimate.

Although under some circumstances this type of forecast may be sufficiently accurate, most forecasters would believe that a more detailed estimate—particularly given the seasonal nature of soft drink demand—is essential. Thus, Frank will move on to perform a time series analysis using a least squares equation, and also to identify the seasonal pattern of Citronade sales.

Time Series Analysis

In this section, we continue to explore **time series forecasting**. However, instead of visual estimation, a more precise statistical technique will be employed: the method of least squares. This method was introduced earlier in this chapter and was used to estimate demand. Whereas demand estimation requires the use of one or more independent variables, and the interactive relationship between these variables is of great importance, in the context of time series analysis there is only one independent variable: time. Thus, this system of forecasting is “naive” because it does not explain the reason for the changes; it merely says that the series of numbers to be projected changes as a function of time.

Despite the mechanical nature of this type of forecast, time series analysis has much to recommend it:

1. It is easy to calculate. A large number of software packages is available.
2. It does not require much judgment or analytical skill by the analyst.
3. It describes the line with the best possible fit, and provides information regarding statistical errors and statistical significance.
4. It is usually reasonably reliable in the short run, unless an absolute turn in the series occurs.

The fact that time series analysis does not take into consideration causative factors does not mean that an analyst using this method should not consider additional information about changes in the underlying forces. Any analyst using this naive method of prediction should try to fine-tune the conclusions based on information that could alter the results.³⁰

When data are collected over a number of periods in the past, they usually exhibit four different characteristics:

1. **Trend.** This is the direction of movement of the data over a relatively long period of time, either upward or downward.
2. **Cyclical fluctuations.** These are deviations from the trend due to general economic conditions. For instance, if one were to observe data for the GDP over time, a long-run upward trend would be evident. Also evident in this series would be movement around that trend as the economy rises more quickly or less quickly (or actually declines).
3. **Seasonal fluctuations.** A pattern that repeats annually is characteristic of many products. Toy sales tend to rise briefly before Christmas. Fashions have spring and fall seasons. In the soft drink industry, the expectation is for higher sales during the warmer periods of the year (i.e., June through September). Thus, time series in which data are collected more frequently than annually (monthly, quarterly) can exhibit seasonal variations.
4. **Irregular.** Departures from norm may be caused by special events or may just represent “noise” in the series. They occur randomly and thus cannot be predicted.

Thus, time series data can be represented by the following mathematical expression:

$$Y_t = f(T_t, C_t, S_t, R_t)$$

where Y_t = Actual value of the data in the time series at time t

T_t = Trend component at t

C_t = Cyclical component at t

S_t = Seasonal component at t

R_t = Random component at t

The specific form of this equation could be additive:

$$Y_t = T_t + C_t + S_t + R_t$$

Other forms could also exist. The specification most commonly used is the multiplicative form:

$$Y_t = (T_t)(C_t)(S_t)(R_t)$$

Thus, the changes in the actual values (Y_t) are determined by four factors. The task of the analyst is to “decompose” the time series of Y into its four components.

We can now return to Frank Robinson’s problem of forecasting Citronade sales not only for the entire year of 2013, but for each quarter. This entire procedure is summarized in Table 5.5. An alternative method of obtaining **seasonal**

³⁰It is essential that all changes and alterations are well documented by the analyst so a trail to her or his reasoning can be established.

Table 5.5 Citronade Quarterly Sales Analysis

(1) Quarter	(2) Actual	(3) Moving Average	(4) Centered Moving Average	(5) Ratio Actual/ CMA	(6) Adjusted Seasonal Factors	(7) Data Deseason- alized	(8) Trend	(9) Cycle and Irregular	(10) Cycle
1	842				0.889	946.7	940.1	100.70	
2	939				0.991	947.2	959.4	98.73	100.74
3	1,236	973.0	981.1	1.260	1.229	1,006.0	978.7	102.79	99.99
4	875	989.3	999.0	0.876	0.891	982.5	998.0	98.45	100.50
5	907	1,008.8	1,020.6	0.889	0.889	1,019.8	1,017.3	100.25	99.22
6	1,017	1,032.5	1,041.6	0.976	0.991	1,025.8	1,036.5	98.97	100.61
7	1,331	1,050.8	1,056.5	1.260	1.229	1,083.3	1,055.8	102.60	100.19
8	948	1,062.3	1,073.0	0.884	0.891	1,064.4	1,075.1	99.00	99.84
9	953	1,083.8	1,093.1	0.872	0.889	1,071.5	1,094.4	97.91	98.94
10	1,103	1,102.5	1,110.9	0.993	0.991	1,112.6	1,113.7	99.90	99.60
11	1,406	1,119.3	1,131.0	1.243	1.229	1,144.4	1,133.0	101.01	99.94
12	1,015	1,142.8	1,152.4	0.881	0.891	1,139.6	1,152.3	98.90	100.13
13	1,047	1,162.0	1,174.4	0.892	0.889	1,177.2	1,171.6	100.48	99.78
14	1,180	1,186.8	1,194.6	0.988	0.991	1,190.3	1,190.9	99.95	100.55
15	1,505	1,202.5	1,212.1	1.242	1.229	1,225.0	1,210.1	101.22	99.88
16	1,078	1,221.8	1,232.6	0.875	0.891	1,210.4	1,229.4	98.45	100.29
17	1,124	1,243.5	1,252.4	0.897	0.889	1,263.8	1,248.7	101.21	100.15
18	1,267	1,261.3	1,272.1	0.996	0.991	1,278.0	1,268.0	100.79	100.55
19	1,576	1,283.0	1,288.4	1.223	1.229	1,282.7	1,287.3	99.65	100.18
20	1,165	1,293.8	1,302.9	0.894	0.891	1,308.1	1,306.6	100.11	99.57
21	1,167	1,312.0	1,323.8	0.882	0.889	1,312.1	1,325.9	98.96	99.85
22	1,340	1,335.5	1,343.6	0.997	0.991	1,351.7	1,345.2	100.48	99.69
23	1,670	1,351.8	1,362.8	1.225	1.229	1,359.3	1,364.4	99.62	99.97
24	1,230	1,373.8	1,381.6	0.890	0.891	1,381.1	1,383.7	99.81	100.00
25	1,255	1,389.5	1,401.5	0.895	0.889	1,411.1	1,403.0	100.57	99.96
26	1,403	1,413.5	1,422.5	0.986	0.991	1,415.2	1,422.3	99.50	99.93
27	1,766	1,431.5	1,438.5	1.228	1.229	1,437.4	1,441.6	99.71	99.76
28	1,302	1,445.5	1,457.0	0.894	0.891	1,461.9	1,460.9	100.07	99.79
29	1,311	1,468.5	1,477.4	0.887	0.889	1,474.0	1,480.2	99.59	100.07
30	1,495	1,486.3	1,496.0	0.999	0.991	1,508.0	1,499.5	100.57	99.53
31	1,837	1,505.8	1,515.6	1.212	1.229	1,495.2	1,518.8	98.45	99.92
32	1,380	1,525.5	1,534.3	0.899	0.891	1,549.5	1,538.0	100.74	99.85
33	1,390	1,543.0	1,555.9	0.893	0.889	1,562.9	1,557.3	100.36	100.41
34	1,565	1,568.8	1,579.4	0.991	0.991	1,578.6	1,576.6	100.13	99.81
35	1,940	1,590.0	1,598.1	1.214	1.229	1,579.0	1,595.9	98.94	100.30
36	1,465	1,606.3	1,616.8	0.906	0.891	1,644.9	1,615.2	101.84	100.29
37	1,455	1,627.3	1,638.0	0.888	0.889	1,636.0	1,634.5	100.09	100.84
38	1,649	1,648.8	1,654.5	0.997	0.991	1,663.3	1,653.8	100.58	99.74
39	2,026	1,660.3	1,670.4	1.213	1.229	1,649.0	1,673.1	98.56	99.80
40	1,511	1,680.5	1,688.6	0.895	0.891	1,696.6	1,692.3	100.25	99.90
41	1,536	1,696.8	1,706.4	0.900	0.889	1,727.0	1,711.6	100.90	100.34
42	1,714	1,716.0	1,727.3	0.992	0.991	1,728.9	1,730.9	99.88	99.53
43	2,103	1,738.5			1.229	1,711.7	1,750.2	97.80	99.76
44	1,601				0.891	1,797.6	1,769.5	101.59	66.46



indices using regression analysis is examined in the Seasonality and Seasonal Indices Excel App.

Seasonality The process of decomposition starts with identifying and removing the seasonal factor from our series of numbers in order to calculate the trend. We use the method of **moving averages** to isolate seasonal fluctuations.

Column 1: This column simply represents all the quarters of the 11 years, 2002 through 2012, numbered sequentially.

Column 2: All the quarterly data, originally shown in Table 5.2, are listed here.

Column 3: The first step in the deseasonalizing process is to calculate a quarterly average for the first year (i.e., the first four quarters). The result is 973.0, so this number is placed next to quarter 3 in this column. The next number is obtained by moving down one quarter: The first quarter is dropped; the fifth quarter is added; and the average of quarters 2, 3, 4, and 5 is calculated to be 989.3. This number is placed next to quarter 4. The same procedure is then followed for the rest of the quarters. Because of the averaging, the data for the first two quarters and the last quarter are lost.

The first moving average was placed next to the third quarter. It could also have been placed next to the second quarter. Actually, the average of four quarters belongs between the second and third quarters. This computation is made in column 4.

Column 4: An average of two adjacent numbers in column 3 is calculated and placed at the third quarter. Because 973.0 in column 3 should have appeared between quarters 2 and 3, and 989.3 should have been between quarters 3 and 4, the average of these two, 981.1, is correctly placed at quarter 3. All centered moving averages are placed in column 4. With this procedure, the next to the last number is now lost.

Column 5: We now obtain the seasonal factors. They are the ratios obtained by dividing the actual numbers in column 2 by the centered moving averages of column 4. Observing the numbers in column 5, it can be seen that every fourth quarter, starting with quarter 3 (i.e., quarter 3, 7, 11, and so on), has a factor greater than 1. These quarters represent the summer months (July to September), when soda consumption is at its peak. It can also be seen that data for the second calendar quarters (i.e., quarters 6, 10, 14, and so on) are rather close to 1, whereas the quarters representing fall and winter are generally below 0.9.

Column 6: To obtain a quarterly seasonal index, we must average the ratios obtained by quarter in column 5. This is done in Table 5.6. The quarterly patterns are quite obvious.³¹

The four averages should add to 4; they do not because of rounding. A minor adjustment must now be made. The results are then transferred to column 6 of Table 5.5.

Column 7: When the actual data of column 2 are divided by the seasonal factors of column 6, the deseasonalized series of numbers in column 7 is obtained. Both the actual and the deseasonalized data have been plotted in Figure 5.6. The latter series, as should be expected, is much smoother than the former.

The first step of the decomposition procedure has now been completed:

$$(T \times C \times S \times R) / S = T \times C \times R$$

The new series has eliminated the seasonality. The next step is to calculate the trend.

³¹There appears to be minor downward trend in the third-quarter data. This is discussed later.

Table 5.6 Averaging of Seasonal Factors

Year	First Quarter	Second Quarter	Third Quarter	Fourth Quarter	Total
2002			1.260	0.876	
2003	0.889	0.976	1.260	0.884	
2004	0.872	0.993	1.243	0.881	
2005	0.892	0.988	1.242	0.875	
2006	0.897	0.996	1.223	0.894	
2007	0.882	0.997	1.225	0.890	
2008	0.895	0.986	1.228	0.894	
2009	0.887	0.999	1.212	0.899	
2010	0.893	0.991	1.214	0.906	
2011	0.888	0.997	1.213	0.895	
2012	0.900	0.992			
Average	0.890	0.992	1.229	0.891	4.002
Adjusted Average	0.889	0.991	1.229	0.891	4.000

The Trend Line Computation of the trend uses the least squares method. The dependent variable is the deseasonalized series in column 7 of Table 5.5. The independent variable is time. Each quarter is numbered consecutively; although the initial number usually does not matter, it is easiest and most logical to start with 1. Thus, the consecutive numbers in column 1 of Table 5.5 represent the independent variable.

The form of the equation depends on the shape of the deseasonalized series. A casual observation of the deseasonalized series appears to indicate that a straight line would be most appropriate, but we actually tried three different possibilities:

$$\text{Straight line: } Y = a + b(t)$$

$$\text{Exponential line:}^{32} Y = ab^t$$

$$\text{Quadratic line:}^{33} Y = a + b(t) + c(t)^2$$

Although all three calculations gave acceptable results, the simplest—the straight line—gave the best statistical answers. Its coefficient of determination (R^2) was 0.996, and the t -statistic for the independent variable was 103.6—highly significant.

The exponential and quadratic equations also had very high R^2 s, but their t -statistics were considerably lower (and the squared term in the latter was actually not statistically significant).

The straight-line trend equation was

$$Y = 920.8 + 19.2882t$$

Column 8: Using this equation, we arrive at the trend numbers in column 8.

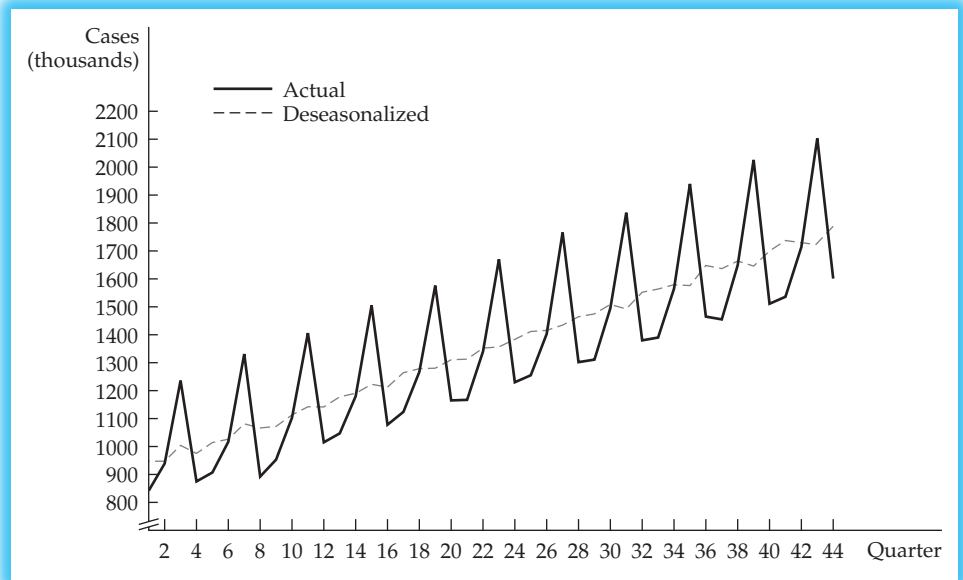
Column 9: Then, by dividing the data in column 7 by those in column 8, we eliminate the trend from the deseasonalized series. In terms of the fraction used previously,

$$(T \times C \times R) / T = C \times R$$



³²This is the equation that would best fit a straight line on a semilogarithmic chart, representing constant percentage growth from period to period. This is analyzed further in the Growth Models Excel App.

³³This equation would fit a curve on a graph with arithmetic scales.

Figure 5.6 Citronade Sales Deseasonalized

Thus, only the cycle and random elements remain in column 9, which is in percentage form.

Cycle and Random Elements At this point we could stop the analysis and do a forecast. The data in column 9 fluctuate rather irregularly within a range of about 5 percentage points. Part of this variation is due to random factors, which cannot be predicted and therefore should be ignored. However, there could be some longer business cycle wave in the data. To isolate the cycle, another smoothing operation can be performed with a moving average.

Column 10: The preferred length of the moving-average period cannot be determined, except in each individual case. In our illustration, we used three periods. If the moving average correctly eliminates the random variations remaining in column 9, then the swings that appear in column 10 should represent a cyclical index.

When the index rises above 100 percent, the indication is that economic activity is strong; the opposite holds for index numbers below 100 percent. The data in column 10 fluctuate very little—less than 2 percentage points. Therefore, no great adjustment is needed here. However, if the indication is that the economy is in a downturn, which may be expected to have some influence on soft drink sales, then a minor forecast adjustment can be made. Such an adjustment can be based on current general forecasts for the economy as a whole or on the latest published leading economic indicators.

Forecasting with Smoothing Techniques

Before we leave the section on projections, one other naive method should be mentioned. This method involves using an average of past observations to predict the future. If the forecaster believes that the future is a reflection of some average of past results, one of two forecasting methods can be applied: simple moving average or exponential smoothing.

The smoothing techniques, either moving average or exponential smoothing, work best when there is no strong trend in the series, when there are infrequent

changes in the direction of the series, and when fluctuations are random rather than seasonal or cyclical. Obviously, these are very limiting conditions. However, if a large number of forecasts is needed quickly and if the estimates involve only one period into the future, then one of these two techniques can be employed.

Moving Average The average of actual past results is used to forecast one period ahead. The equation for this construction is simply

$$E_{t+1} = (X_t + X_{t-1} + \dots + X_{t-N+1})/N$$

where E_{t+1} = Forecast for the next period ($t + 1$)

X_t, X_{t-1} = Actual values at their respective times

N = Number of observations included in the average

The forecasts for Citronade, with their strong trend and seasonal patterns, would certainly not lend themselves well to this simple projection method. Instead, a hypothetical example is used. Twelve observations are shown in the second column of Table 5.7.

The forecaster must decide how many observations to use in the moving average. The larger the number of observations in the average, the greater the smoothing effect. If the past data appear to contain significant randomness while the underlying pattern remains the same, then a larger number of observations is indicated. In Table 5.7, three moving averages have been computed—3, 4, and 5 months. The resulting estimates are shown in the forecast columns.

Table 5.7 Moving-Average Forecasting

Period	Actual	3-Month Moving Average			4-Month Moving Average			5-Month Moving Average		
		Forecast	Absolute Error	Squared Error	Forecast	Absolute Error	Squared Error	Forecast	Absolute Error	Squared Error
1	1,100									
2	800									
3	1,000									
4	1,050	967	83	6,944						
5	1,500	950	550	302,500	988	513	262,656			
6	750	1,183	433	187,778	1,088	338	113,906	1,090	340	115,600
7	700	1,100	400	160,000	1,075	375	140,625	1,020	320	102,400
8	650	983	333	111,111	1,000	350	122,500	1,000	350	122,500
9	1,400	700	700	490,000	900	500	250,000	930	470	220,900
10	1,200	917	283	80,278	875	325	105,625	1,000	200	40,000
11	900	1,083	183	33,611	988	88	7,656	940	40	1,600
12	1,000	1,167	167	27,778	1,038	38	1,406	970	30	900
13		1,033			1,125			1,030		
Total			3,133	1,400,000		2,525	1,004,375		1,750	603,900
Mean			348	155,556		316	125,547		250	86,271

Note: Some of the squared errors appear to be incorrect; this is so because decimals have been omitted.

For example, the forecast of 967 units in period 4 of the 3-month moving average was obtained as follows:

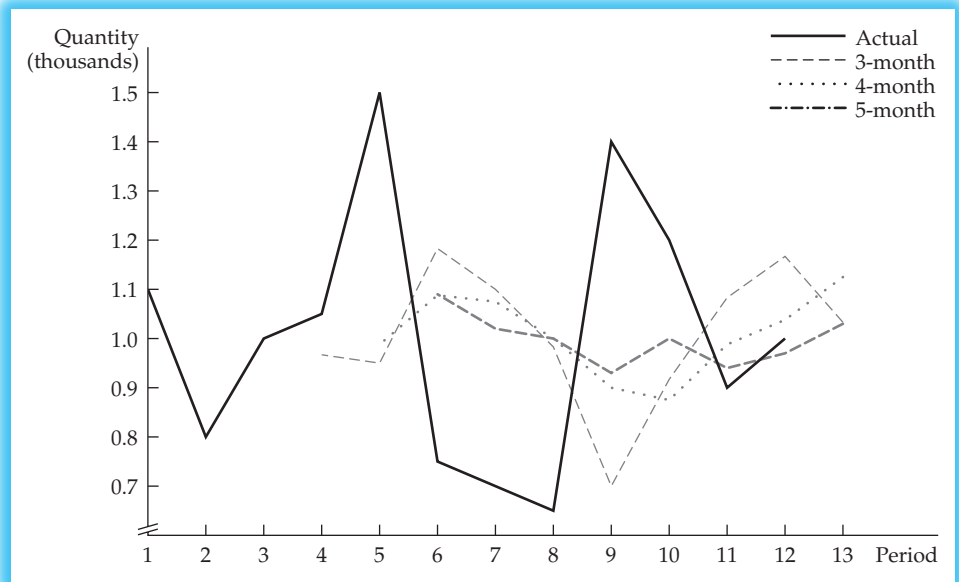
$$\begin{aligned} E_4 &= (X_1 + X_2 + X_3)/N \\ &= (1,100 + 800 + 1,000)/3 \\ &= 2,900/3 = 967 \end{aligned}$$

The three forecast columns diverge widely. These variations are shown graphically in Figure 5.7. As more observations are included in the moving average, the forecast line becomes smoother.

Which moving average should be used? One selection method is to calculate the mean error and the mean squared error of the differences between the actual data and the forecast.³⁴ The series with the smallest squared error would be preferred. In the example of Table 5.7, the 5-month moving average minimizes the deviations.

Exponential Smoothing The moving-average method awards equal importance to each of the observations included in the average, and gives no weight to observations preceding the oldest data included. However, the analyst may believe that the most recent observation is more relevant to the estimate of the next period than previous observations. In that case, it is more appropriate to employ the **exponential smoothing** method, which allows for the decreasing importance of information in the more distant past. This is accomplished by the mathematical technique of geometric progression.

Figure 5.7 Forecasting with Moving Averages



³⁴The formula for the mean squared error is

$$\frac{\sum_{n=1}^N (E_n - X_n)^2}{N}$$

Instead of the mean squared error, we could calculate the root mean squared error by taking the square root of the previous equation.

Older data are assigned increasingly smaller weights; the sum of the weights, if we approached an infinitely large number of observations, would equal 1. All the complex formulations of the geometric series can be simplified into the following expression:

$$E_{t+1} = wX_t + (1 - w)E_t$$

where w is the weight assigned to an actual observation at period t .

Thus, to make a forecast for one period into the future, all that is needed is the previous period's actual observation and the previous period's forecast. The analyst does not need the extensive historical data required for the moving-average method. The most crucial decision the analyst must make is the choice of the weighting factor. The larger the w (i.e., the closer to 1), the greater will be the importance of the most recent observation. Therefore, when the series is rather volatile and when w is large, the smoothing effect may be minimal. When w is small, the smoothing effect will be considerably more pronounced. This result can be seen in Table 5.8 and Figure 5.8 where w s of 0.2, 0.4, and 0.8 have been used.³⁵ Which weights should be used? We can again calculate the mean squared error, as was done in Table 5.8. Using $w = 0.2$ minimizes the error, so this is the best weight for this set of data.

Both of these naive forecasting techniques have their advantages and disadvantages. The simplicity of the methods is certainly an advantage. However, their usefulness is limited to cases where there is an underlying stability (i.e., no trend) with period-to-period random fluctuations. When a trend or a repeating pattern of fluctuations is present, the time series method with deseasonalization is a much better technique to use. In any case, the two methods just discussed should be employed only for extremely short-term estimates—preferably just one period into the future.

Table 5.8 Exponential Smoothing Forecasting

Period	Actual	Smoothing Factor = 0.2			Smoothing Factor = 0.4			Smoothing Factor = 0.8		
		Forecast	Absolute Error	Squared Error	Forecast	Absolute Error	Squared Error	Forecast	Absolute Error	Squared Error
1	1,100									
2	800	1,100	300	90,000	1,100	300	90,000	1,100	300	90,000
3	1,000	1,040	40	1,600	980	20	400	860	140	19,600
4	1,050	1,032	18	324	988	62	3,844	972	78	6,084
5	1,500	1,036	464	215,667	1,013	487	237,364	1,034	466	216,783
6	750	1,128	378	143,247	1,208	458	209,471	1,407	657	431,491
7	700	1,053	353	124,457	1,025	325	105,370	881	181	32,897
8	650	982	332	110,375	895	245	59,910	736	86	7,443
9	1,400	916	484	234,467	797	603	363,779	667	733	536,915
10	1,200	1,013	187	35,109	1,038	162	26,207	1,253	53	2,857
11	900	1,050	150	22,530	1,103	203	41,156	1,211	311	96,528
12	1,000	1,020	20	403	1,022	22	472	962	38	1,434
13		1,016			1,013			992		
Total			2,728	978,180		2,886	1,137,973		3,043	1,442,033
Mean			248	88,925		262	103,452		277	131,094

Note: Some of the squared errors appear to be incorrect; this is so because decimals have been omitted.

³⁵Note that the forecast for period 2 in each case is the actual observation of period 1. Because there is no forecast for period 1, a forecast must be made up. Any number could have been used, but the actual observation is probably the most logical choice.

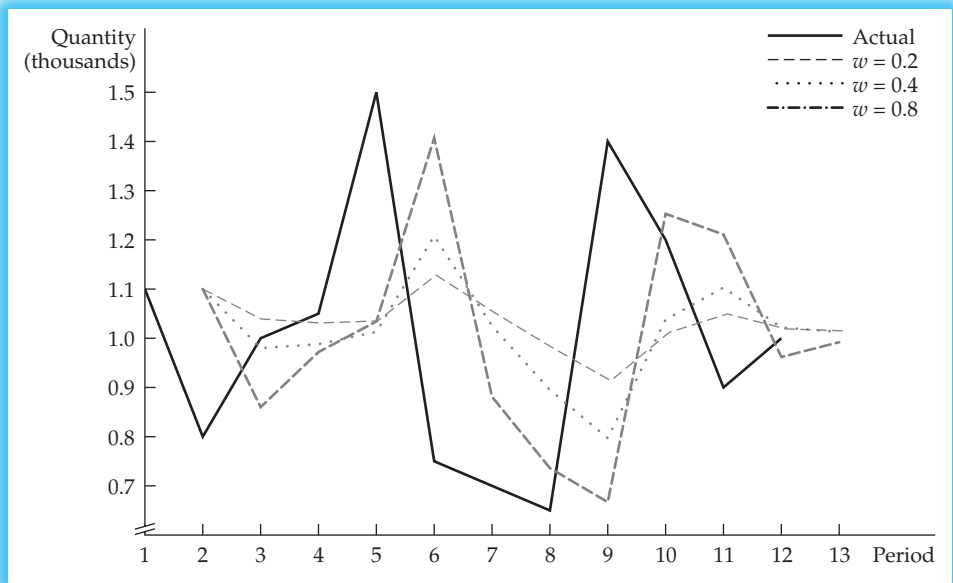


Figure 5.8 Exponential Smoothing Forecasting

Box-Jenkins Forecasting Method Only a brief mention of this sophisticated and fairly complex model is made here. Readers interested in greater detail should consult a business forecasting book.³⁶

The Box-Jenkins model, as do time series analysis and smoothing models, relies on prior observations rather than exogenous independent variables, to arrive at a forecast. However, it does not assume a specific pattern of prior data. Rather, it permits the choice among a large number of models. Through an iterative procedure, it arrives at the best result (i.e., the one that produces the smallest errors).

Recent Forecasting Methods and Results

A Sales Forecasting Survey³⁷ An article published in 2005 presents the results of a survey that was conducted to discover whether forecasting practices have changed over the previous 20 years and whether forecasting results have improved. It compares its results to two similar surveys that were conducted in 1995 and 1984.³⁸ Some of the major findings of this survey were as follows:

- The most familiar forecasting techniques were moving averages followed by exponential smoothing, regression and straight-line projection.
- The familiarity with qualitative methods (such as jury of executive opinion) decreased from the previous surveys.

³⁶For instance, Hanke, Wichern, and Reitsch, *Business Forecasting*, Chapter 9, or J. Holton Wilson and Barry Keating, *Business Forecasting*, 3d ed., Irwin McGraw-Hill, 1998, Chapter 7.

³⁷Teresa M. McCarthy, Donna F. Davis, Susan L. Golicic, and John T. Menzer, "The Evolution of Sales Forecasting Management: A 20-Year Longitudinal Study of Forecasting Practices," *Journal of Forecasting*, Vol. 25, Issue 5, August 2005, pp. 303–324.

³⁸John T. Menzer and James E. Cox, Jr., "Familiarity, Application, and Performance of Sales Forecasting Techniques," *Journal of Forecasting*, Vol. 3, Issue 5, January–March 1984, pp. 27–36; John T. Menzer and Kenneth B. Kahn, "Forecasting Technique Familiarity, Satisfaction, Usage and Application," *Journal of Forecasting*, Vol. 14, Issue 5, September 1995, pp. 465–476.

- Exponential smoothing appeared to be the most satisfactory method, followed by trend-line analysis.
- The satisfaction level with the various techniques appears to be declining from previous surveys.
- Despite the great improvement in tools such as software that are now available to forecasters, the overall degree of accuracy of the forecasts is substantially lower than that of the previous two studies.
- The authors conclude that the major reasons for the decline in accuracy are:
 - Lack of familiarity by users with techniques, and lack of satisfaction with forecasting systems.
 - Most forecasting personnel are not held accountable for performance.
 - Sales forecasting performance does not affect compensation of forecasting personnel.
- The authors state that “sales forecasting will not improve until companies commit the resources ... with personnel trained in the use of sales forecasting techniques ... properly measured and rewarded for performance.”

Forecast Errors A very recent article concerns itself with forecasting errors. The author concludes that despite advancements in computer technology and improvements in forecaster skills, data on U.S. industries combined as well as six individual industries show that errors have actually increased in many cases. The information is based on surveys performed by the Institute of Business Forecasting and Planning. The surveys examined the data at three levels (stock-keeping units, category, and aggregate) and looked at forecasts for one month ahead, two months ahead, one quarter ahead, and one year ahead. The data were classified into two periods of five years, 2000–2004 and 2005–2009, and compared the average forecasting error between these two groups. The errors were lower when the forecasts were made at an aggregate level and the errors were higher as the forecasts were further in the future. But when the two periods were compared, the results showed that, in the majority of cases, the errors increased. In some cases they remained the same, and in a few cases there was some improvement.³⁹

Forecasting the Demand for New Drugs A recent article reported on a method used by pharmaceutical companies to forecast the demand for newly introduced drugs. The firms use marketing research called “new drug acceptance and new drug adoption studies to estimate the percentage of the target market” that will use the new drug immediately and later over its life cycle. The method utilizes the Gompertz curve, which is an S-shaped curve that describes the new drug use. Please refer to Figure 5.9 to see a typical Gompertz curve. To develop this curve three variables must be entered: the intercept or minimum value, the asymptote or maximum value, and the slope or rate of change. To obtain these values the forecasters will estimate their values from “an analog or a previously launched drug of the same kind in the same market.” But because this is an innovative drug, such information may not always be available. Then researchers will have to look for information elsewhere. Thus, in the instance of a new drug for attention deficit hyperactivity disorder (ADHD), cited in the article, the marketing research was conducted among physicians two years before the launch of the product to ascertain the percentage growth of this drug over its life cycle. From this survey the forecasters were able to estimate the three variables to obtain the forecast over several years.⁴⁰

³⁹Chaman L. Jain, “Forecast Errors: How Much Have We Improved?” *Journal of Business Forecasting*, Summer 2011, pp. 27–30.

⁴⁰Michael Latta, “How to Forecast the Demand of a New Drug in the Pharmaceutical Industry,” *The Journal of Business Forecasting*, Vol. 25, Issue 3, Fall 2007, pp. 21–23, 28.

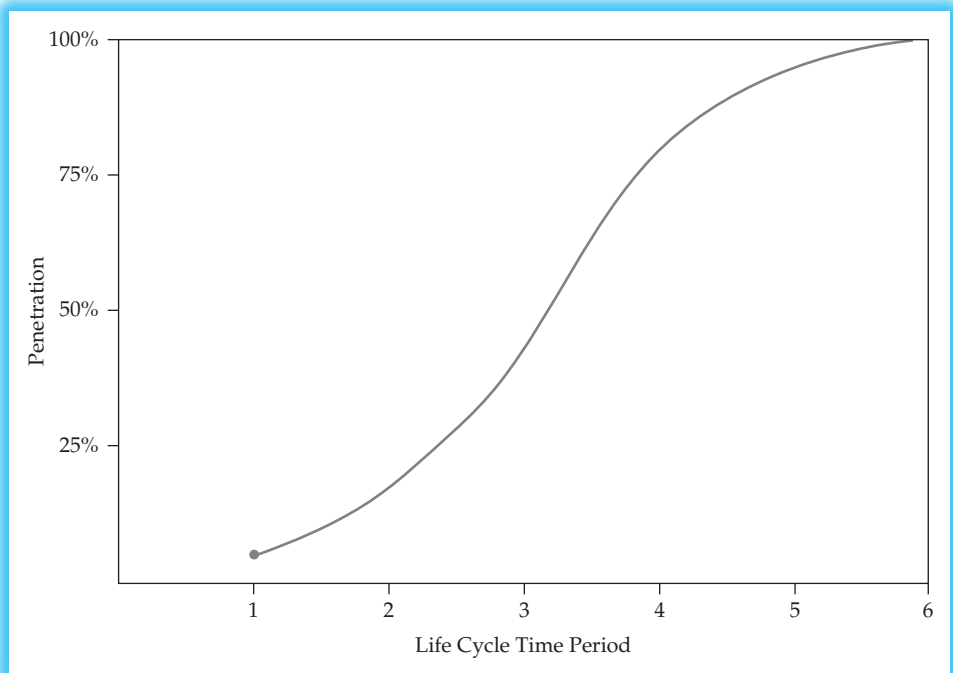


Figure 5.9 The Gompertz Curve

Econometric Models

Up to this point, all the quantitative forecasting techniques discussed can be classified as naive. In this section of the chapter, our brief discussion concentrates on models that are termed *causal* or *explanatory*. These **econometric forecasting models** were described in detail earlier in this chapter.

Regression analysis is an explanatory technique. Unlike the case of a naive projection, which relies on past patterns to predict the future, the analyst performing regression analysis must select those independent (or explanatory) variables that are influential in determining the dependent variable. Although simple projection models often give adequate results, the use of explanatory variables in the analysis can enhance the accuracy and the credibility of the estimates. Of course, no regression equation will explain the entire variation of the dependent variable because in most economic relationships there are numerous explanatory variables with complex interrelationships. As already explained, the person performing the analysis will have to settle for the inclusion of a limited number of variables; an equation that approximates the functional relationships; and results that explain a significant portion of, but not the entire variation in, the dependent variable.

The regression methods previously described employ a single equation to make estimates. Here we discuss a few studies of single-equation regression models of demand in which many of the variables are in the form of time series.

Among the many studies of automobile demand, the one written by Daniel Suits⁴¹ is quite interesting. He combined the demand equations for new cars and used cars into one, as follows:

$$\Delta R = a_0 + a_1\Delta Y + a_2\Delta P/M + a_3\Delta S + a_4\Delta X$$

where R = Retail sales, in millions of new cars

Y = Real disposable income

P = Real retail price of new cars

M = Average credit terms (number of months of the average installment contract)

S = Existing stock, in millions of cars

X = Categorical variable

All the variables are in terms of first differences; thus, the equation estimates year-to-year changes. The years 1942 through 1948 have been excluded from the time series because of the war years, when no automobiles were produced, and the immediate postwar period, when significant distortions existed in the automobile market.

The categorical variable X takes into account the fact that not all the remaining years in the series can be treated equally. The years 1941, 1952, and 1953 were considered exceptional. The first two were assigned a value of +1, the third a value of -1, and all the other years a value of 0.

Having specified the equation, Suits calculated the coefficients (the a s) and tested the results by predicting 1958 new auto sales. "The demand equation predicts a level of sales of slightly less than 6.0 million. This compares favorably with a preliminary estimate of 6.1 million actual sales for the year and stands in sharp contrast with sales forecasts of 6.5 million and over, which were common in the industry at the start of the year."⁴²

Another economist studied the demand for computers.⁴³ He specified the following equation in terms of logarithms:

$$\log(y_t/y_{t-1}) = a_0 - a_1\log p_t - a_2\log y_{t-1}$$

where y_t = Stock of computers in year t

y_{t-1} = Stock of computers in year $t - 1$

p = Real price of computers

The demand in this equation is measured as the percentage change in the stock of computers from one year to the next.

An additional example is given for coffee demand in the United States.⁴⁴ Coffee demand was estimated with the following equation:

$$Q = b_0 + b_1P_0 + b_2Y + b_3P + \sum_{i=1}^3 b_{4i}D_i + b_5T$$

where Q = Per capita (population over age 16) quantity of coffee

P_0 = Real retail price of coffee, per pound

Y = Real per capita disposable income

P = Real retail price of tea, per quarter pound

D_i = Binary (dummy) variable for quarters of the year

T = Time trend

⁴¹Daniel B. Suits, "The Demand for New Automobiles in the United States, 1929-1956," *Review of Economics and Statistics*, August 1958, pp. 273-80.

⁴²Ibid., p. 273.

⁴³Gregory Chow, "Technological Change and the Demand for Computers," *American Economic Review*, December 1967, pp. 1117-30.

⁴⁴Cliff J. Huang, John J. Siegfried, and Farangis Zardoshty, "The Demand for Coffee in the United States, 1963-1977," *Quarterly Review of Economics and Business*, 20, 2 (summer 1980), pp. 36-50.

The authors found that coffee consumption is sensitive to its own price, that there had been a long-term decline in coffee consumption, and that significantly less coffee was consumed during the spring and summer months. The other coefficients did not turn out to be statistically significant. Thus, the consumption of coffee could be estimated from the time trend, the quarterly changes, and an assumption made regarding coffee prices.

A last example is a more recent one, and represents a demand forecast prepared for a specific product for a specific company. Columbia Gas of Ohio forecasts daily demand by using both time series and regression analysis. Two years of historical data are used to obtain the regression. The time series serves to forecast growth. The forecast is then based on a combination of the regression analysis and the growth rate:

$$Q = (1 + G) \times (a + b_1T + b_2P + b_3W)$$

where Q = Daily demand in MDTH (thousand dekatherms)

G = Growth rate

T = Forecasted temperature

P = Prior day temperature

W = Forecasted wind speed

a = Intercept

b_1, b_2, b_3 = Coefficients

The R^2 of the regression equation is 0.983.

This forecast “is used mainly to ensure that scheduled supplies are in balance with forecasted demands over the next five-day period.” The company receives a 5-day weather forecast twice daily.⁴⁵

In each study, the numerical coefficients (the a s and the b s) were obtained by regression analysis. To forecast using these equations, it is necessary, of course, to make estimates of each independent variable for the period indicated.⁴⁶ Alternative forecasts can be created by assuming different values for each independent variable.

An important warning is in order regarding the use of regression equations for forecasting. All the parameters (coefficients) in a regression equation are estimated using past data (whether time series or cross-sectional analysis has been employed). The forecast based on such estimates will be valid only if the relationships between the dependent variable and the independent variables do not change from the past into the future. Literally, a regression equation is valid only within the limits of the data used in it. When we venture outside these limits (i.e., forecast with independent variables outside the range of the past), we are treading on dangerous ground. Still, forecasting with the least squares method is a much-used technique, often employed quite successfully. But its limitations must not be forgotten by the analyst.

Although many forecasting problems can be solved with the use of single-equation regression models, there are instances when one equation is not sufficient. In such cases, economists often turn to multiple-equation systems. Extremely large models used by economists to predict the GDP and its component parts are examples of such models, which may include hundreds of variables and equations.

A single-equation model can be used when the dependent variable can be estimated using independent variables that are determined by events not part of the equation. But what happens when the determining variables are determined by other variables within the model? Then a single equation is insufficient.

In a multiple-equation system, variables are referred to as *endogenous* and *exogenous*. Endogenous variables are comparable to the dependent variable of the single-equation

⁴⁵H. Alan Catron, “Daily Demand Forecasting at Columbia Gas,” *Journal of Business Forecasting*, Summer 2000, pp. 10–15.

⁴⁶In the case of lagged variables, actual data may be available. See, for instance, y_{t-1} , the stock of computers at the end of the previous year in the equation used in Chow, “Technological Change.”

model; they are determined by the model. However, they can also influence other endogenous variables, so they may appear as “independent” (i.e., on the right side of the equation) variables in one or more of the equations. Exogenous variables are from outside the system and are not determined within it; they are truly independent variables.

The following is an example of an extremely simple two-equation model of the private economy:

$$\begin{aligned}C &= a_0 + a_1Y \\ Y &= C + I\end{aligned}$$

where C = Consumption
 Y = National income
 I = Investment

Here C and Y are endogenous variables. I is assumed (rather unrealistically) to be exogenous, or determined by forces outside the system of equations.

Another simple multiequation model represents the interrelationships of activities within a firm:

$$\begin{aligned}\text{Sales} &= f(\text{GDP, Prices}) \\ \text{Cost} &= f(\text{Quantity of product, Factor prices}) \\ \text{Expenses} &= f(\text{Sales, Factor prices}) \\ \text{Product price} &= f(\text{Cost, Expenses, Profit}) \\ \text{Profit} &= \text{Sales} - \text{Cost} - \text{Expenses}\end{aligned}$$

Here there are five endogenous variables. The exogenous variables are GDP, Quantity of product, and Factor prices. These equations are given in functional form only. The specific forms of the equations would, of course, have to be specified.

Such systems of equations would then be solved to obtain all of the coefficients. The statistical methods used to solve for all the numerical values are beyond the scope of this textbook.

GLOBAL APPLICATION: FORECASTING EXCHANGE RATES

In Chapter 2, the additional challenges facing a multinational corporation are discussed. Such a corporation must forecast sales, expenses, and cash flows for its operations in different countries. The results from these transactions generated in the home country will depend on the exchange rate between the foreign and domestic currency. Often a company makes substantial investments in foreign operating facilities from which it expects to obtain cash flows. Thus, multinational corporations (MNCs) are vitally interested in forecasting exchange rates in both the short and long run.

A frequently used method to forecast exchange rates for relatively short periods into the future is through the forward rate. To understand this, two types of exchange rates must be defined:

1. Spot exchange rate—the price of one currency against another for transactions being completed immediately
2. Forward exchange rate—the price of one currency against another for a transaction that will occur at some point in the future

An extremely active exchange market exists, and financial officers can easily ascertain both spot and forward rates for the ten major currencies, which represent a predominant portion of all transactions. One view of the forward rate is that it represents the market’s consensus on what the future spot rate should be. Thus, if today’s spot rate is \$1.998 for 1 British pound sterling and the 90-day forward rate is \$1.989, then one can argue that the market consensus is that the pound will decrease in value by \$0.009 relative to the U.S. dollar 90 days hence.

Is the forward rate a good predictor of the spot exchange rate for major currencies? On the average, the forward rate will be equal to the future spot rate; in other words, negative and positive errors in the forecasts will be offset. Therefore, it can be said that the forward rate is an unbiased forecaster of the spot rate. It makes for a best-guess forecast. However, this does not make it an accurate forecast at any one time. Still, for the short run, it is probably as good an estimator as we have. In addition to its lack of accuracy, other shortcomings must be considered:

1. The present exchange rate system does not permit currencies to float freely. Governments interfere in the exchange rate markets when they consider it to be of benefit to their country's economy.
2. Although forward rates can be established for relatively long periods into the future (in some cases, they can go out as far as 10 years), by far the largest volume of forward contracts is for 180 days or less.
3. Reliable forward markets exist only for currencies of the leading industrial economies of the world.

Longer-term exchange rate forecasts often use econometric models. A major problem in constructing these multiple regression models is in finding appropriate reliable independent variables. In most cases, the independent variables are stated in terms of differentials between the domestic and foreign measures, such as:

1. Growth rates of GDP
2. Real interest rates
3. Nominal interest rates
4. Inflation rates
5. Balance of payments

A significant number of complex models is in use today. Here we use a simple hypothetical model to illustrate the estimate of a relationship between the domestic currency and a foreign currency:

$$E_t = a + bI_t + cR_t + dG_t$$

where E = Exchange rate of a foreign currency in terms of the domestic currency

I = Domestic inflation rate minus foreign inflation rate

R = Domestic nominal interest rate minus foreign nominal interest rate

G = Domestic growth rate of GDP minus the growth rate of foreign GDP

t = Time period

a, b, c, d = Regression coefficients

Such forecasts have all the advantages and disadvantages usually found in the application of regression analysis to economic problems:

1. Which variables should be included?
2. What form should the regression equation adopt?
3. Are accurate measurements of independent variables available? Data may be inadequate for other than the major industrial countries.
4. To forecast exchange rates, it may be necessary to forecast the independent variables. It is not surprising that it may be as difficult to forecast independent variables accurately as it is to forecast the exchange rates themselves.

By now the reader has probably concluded that forecasting exchange rates is not an easy task. But then, no forecasts are easy. The accuracy of exchange rate forecasts has been inconsistent at best. However, this should not deter businesspeople from attempting to make such forecasts. Even if inaccurate, they enhance a manager's

knowledge of her environment and provide a manager with insights into the economic variables that affect international decision making.

The discussion of this subject has been limited to its bare essentials. Students interested in studying this topic further can do so by reading one of several textbooks dealing with international financial management.⁴⁷

The Solution

Frank Robinson is now ready to make his forecast for 2013. He will do it in three steps:

1. He will project the trend for the four quarters of 2013.
2. He will apply the seasonal factors.
3. He may then make an adjustment for cyclical influences.

Because the trend equation is $Y = 920.8 + 19.2882(t)$, his trend forecasts will be as follows:

$$\text{1st qtr: } 920.8 + 19.2882(45) = 1,788.8$$

$$\text{2nd qtr: } 920.8 + 19.2882(46) = 1,808.1$$

$$\text{3rd qtr: } 920.8 + 19.2882(47) = 1,827.3$$

$$\text{4th qtr: } 920.8 + 19.2882(48) = 1,846.6$$

Next, each result must be multiplied by the seasonal factors:

$$\text{1st qtr: } 1,788.8 \times 0.889 = 1,590.2$$

$$\text{2nd qtr: } 1,808.1 \times 0.991 = 1,791.8$$

$$\text{3rd qtr: } 1,827.3 \times 1.231 = 2,245.8$$

$$\text{4th qtr: } 1,846.6 \times 0.891 = 1,645.4$$

Frank noticed, however, that the seasonal factors in Table 5.6 exhibit some trends over the 11 years. The third-quarter factor has been decreasing and the first- and fourth-quarter factors appear to be rising. To see the effect of these changes, Frank decides to use the average of the last three observations for each quarter (and adjust to a total of 4, as was done previously). When he applies his new factors to the trend numbers, he obtains the following results:

$$\text{1st qtr: } 1,788.8 \times 0.894 = 1,599.2$$

$$\text{2nd qtr: } 1,808.1 \times 0.993 = 1,795.3$$

$$\text{3rd qtr: } 1,827.3 \times 1.213 = 2,216.6$$

$$\text{4th qtr: } 1,846.6 \times 0.900 = 1,662.0$$

He decides that this last computation is more valid. The last step in his procedure is to evaluate the effect of cycles and random factors. He decides that the cycle's influence on soft drink sales will be neutral in the coming year. Although an estimate of random factors cannot be made, Frank knows that the company has plans to implement a strong advertising campaign on behalf of Citronade during the second quarter of 2013. At this point, however, he does not have sufficient information regarding the effect of advertising on sales. He makes a note to mention that in his report as a possible plus and to try to make some quantitative estimates later.

An additional point deserves attention. We have assumed the product whose sales Frank is forecasting has been in existence for at least 11 years. Had this been a new flavor, not previously produced, the forecast would have been much more difficult. If competing companies had been selling a similar product in the past, market information would be available in terms of very detailed statistics published by the beverage industry. Frank could then base his forecast on these data and assume some pattern of market penetration for the new brand. If, however, this is a completely new flavor not previously produced in the industry, it may be necessary to conduct market research to establish a base for a forecast. Another method would be to study the sales patterns of other soft drink products from point of introduction and to base the new product forecast on their histories.

⁴⁷For example, Dennis J. O'Connor and Alberto T. Bueso, *International Dimensions of Financial Management*, New York: Macmillan, 1990. The preceding discussion has drawn on material contained in this book.

SUMMARY

The aim of this chapter is to present an introduction to regression analysis and forecasting, with a primary emphasis on how the techniques of these disciplines are applied to management problems.

We focused the application of regression analysis on the estimation of demand. This particular application can be summarized as follows:

1. Specification of the regression model of demand
2. Collection of the relevant data
3. Estimation of the regression equation
4. Analysis and evaluation of the regression results (e.g., t -test, F -test, R^2), and adjustment or correction for any statistical problems (e.g., multicollinearity, autocorrelation, incorrect functional form)
5. Assessment of regression findings for use in making policy decisions

In a formal econometrics course, most of the emphasis is placed on steps 1, 3, and 4 (i.e., the more technical aspects of this type of statistical analysis). In business, the most important steps are 2 and 5. Powerful computers and sophisticated software packages are available to everyone today at such a reasonable cost that it has become elementary to apply regression analysis to the estimation of demand or any other aspect of business research. The real challenge is to obtain *good* data and to apply judiciously the results of the regression analysis to the managerial decision-making process. In these two areas of regression analysis, no textbook or course can take the place of actual hands-on experience.

Forecasting is an important activity in many organizations. In business, forecasting is a necessity.

This chapter summarizes and discusses a number of forecasting techniques. Six categories of forecasts are included:

1. *Expert opinion* is a qualitative technique of forecasting based on the judgment of knowledgeable people. The Delphi method is another type of expert opinion forecast that is generally applicable in forecasting technological advances.
2. *Opinion polls and market research* are conducted among survey populations, not experts, to establish future trends or consumer responses.
3. *Surveys of spending plans* are concerned with such important economic data as consumer sentiment. The forecasts are based on replies to questionnaires or interviews.
4. *Economic indicators* are indexes of a number of economic series intended to forecast the short-run movements of the economy, including changes in direction.
5. *Projections*, a quantitative method, employ historical data to project future trends. Usually, no causes for trends are identified.
6. *Econometric models* are explanatory or causal models in which independent variables that influence the statistic to be forecast are identified.

The chapter also examines the decomposition of least squares projections into trends, seasonal and cyclical fluctuations, and irregular movements.

One other naive forecasting method was mentioned: forecasting with smoothing techniques. Smoothing techniques fall into two major categories, moving averages and exponential smoothing, and are useful when there are no pronounced trends in the data and when fluctuations from period to period are random.

IMPORTANT CONCEPTS

Demand Estimation

Coefficient of determination (R^2). Measures the goodness of fit of a regression equation. The proportion of variation in the dependent variable explained by variation in the independent variable. Its value ranges from 0 to 1. Zero indicates that none of the

variation in the dependent variable is explained by the regression equation. One indicates that all of the variation in the dependent variable is explained by the regression equation. In multiple regression analysis, this measure is referred to as the *multiple coefficient of determination*. (p. 123)

Consumer survey. The attempt to obtain data about demand directly by asking consumers about their purchasing habits through such means as face-to-face interviews, focus groups, telephone surveys, and mailed questionnaires. (p. 116)

Critical t -value. A value derived from a t -table used to evaluate the significance of a variable. (p. 122)

Cross-sectional data. Data on a particular set of variables for a given point in time for a cross-section of individual entities (e.g., persons, households, cities, states, countries). (p. 118)

Degrees of freedom (df). An adjustment factor that is required in conducting the t -test. This number is found by subtracting the number of independent variables plus 1 from the number of observations in the sample; that is, $df = n - (k + 1) = n - k - 1$. (p. 122)

F -test. A test for the statistical significance of the R^2 value. If this test is passed, a researcher can be confident that all the estimated coefficients of a regression model together are not zero for the population under study. (p. 123 and Appendix, Table A.3)

One-tail test. This refers to the nature of the alternative hypothesis in the t -test. If the alternative hypothesis states that the population coefficient is positive, then the upper tail of the t -distribution is used. If the alternative hypothesis states that the population coefficient is negative, then the lower tail is used. In either case, only one tail is used. (p. 122)

Regression analysis. A statistical technique for finding the best relationship between a dependent variable and selected independent variables. If one independent variable is used, this technique is referred to as *simple regression*. If more than one independent variable is used, it is called *multiple regression*. (p. 117)

Rule of 2. A general rule of thumb employed by economists in conducting the t -test. Essentially, it states that any t -ratio of 2 or more indicates that the estimated coefficient is statistically significant at the 0.05 level. (p. 123)

Standard error of coefficient (SEC). A measure of the deviation of an estimated regression coefficient from the hypothesized value of the true (but unknown) population coefficient. In the t -test, the standard error of a particular estimated coefficient is divided into this coefficient, thereby indicating the t -value. (p. 122)

Standard error of estimate (SEE). Is a measure of the scatter of actual observations around the regression line. It is used for providing confidence intervals on estimated values of the dependent variable. (p. 120)

Time series data. Data for a particular set of variables that track their values over a particular period of

time at regular intervals (e.g., monthly, quarterly, annually). (p. 118)

t -statistic. A value calculated as $t = \text{coefficient}/\text{SEC}$. Used to evaluate the significance of a variable. Also called a t -value. (p. 122)

t -table. A table indicating the critical t -value based on the t -distribution whose mean value is zero. (p. 122 and Appendix A, Table A.4)

t -test. A test of the statistical significance of an estimated regression coefficient. If the *magnitude* of the t -statistic is greater than the critical t -value then the researcher can be confident that the value of the true population coefficient is not zero. (p. 122)

Two-tail test. A t -test in which the alternative hypothesis states that the population coefficient may be either positive or negative (i.e., it is not zero); that is, either the upper or the lower tail of the t -distribution may be used. (p. 123)

Forecasting

Causal (explanatory) forecasting. A quantitative forecasting method that attempts to uncover functional relationships between independent variables and the dependent variable. (p. 134)

Compound growth rate. Forecasting by projecting the average growth rate of the past into the future. (p. 139)

Delphi method. A form of expert opinion forecasting that uses a series of written questions and answers to obtain a consensual forecast, most commonly employed in forecasting technological trends. (p. 135)

Econometric forecasting model. A quantitative, causal method that uses a number of independent variables to explain the dependent variable to be forecast. Econometric forecasting employs both single- and multiple-equation models. (p. 153)

Economic indicators. A barometric method of forecasting in which economic data are formed into indexes to reflect the state of the economy. Indexes of leading, coincident, and lagging indicators are used to forecast changes in economic activity. (p. 136)

Exponential smoothing. A smoothing method of forecasting that assigns greater importance to more recent data than to those in the more distant past. (p. 149)

Jury of executive opinion. A forecast generated by experts (e.g., corporate executives) in meetings. A similar method is to ask the opinion of sales representatives who are exposed to the market on a daily basis. (p. 135)

Moving average method. A smoothing technique that compensates for seasonal fluctuations. (p. 145)

Naive forecasting. Quantitative forecasting that projects past data without explaining the reasons for future trends. (p. 134)

Opinion polls. A forecasting method in which sample populations are surveyed to determine consumption trends. (p. 135)

Qualitative forecasting. Forecasting based on the judgment of individuals or groups. Also called *judgmental forecasting*. (p. 134)

Quantitative forecasting. Forecasting that examines historical data as a basis for future trends. (p. 134)

Seasonal indices. A method for describing the seasonal variation in a time series. Seasonal indices may either be multiplicative, in which case the sum of indices should sum to the number of time periods per year (for example, Quarterly indices of 90, 80, 105, and 125), or additive, in which case the

sum of indices should sum to zero (for example, Quarterly indices of -10, -20, 5, and 25). (p. 143)

Surveys of spending plans. Examinations of economic trends such as consumer sentiment and inventory. (p. 136)

Time series forecasting. A method of forecasting from past data by using least squares statistical methods. A time series analysis usually examines trends, cyclical fluctuations, seasonal fluctuations, and irregular movements. (p. 142)

Trend projections. A form of naive forecasting that projects trends from past data. Trend projections usually employ compound growth rates, visual time series, or least squares time series methods. (p. 139)

QUESTIONS

Demand Estimation

1. Explain the difference between time series data and cross-sectional data. Provide examples of each type of data.
2. Would there be any differences in the set of variables used in a regression model of the demand for consumer durable goods (e.g., automobiles, appliances, furniture) and a regression model of the demand for “fast-moving consumer goods” (e.g., food, beverages, personal care products)? Explain.
3. Briefly explain the meaning of R^2 . A time series analysis of demand tends to result in a higher R^2 than one using cross-sectional data. Why do you think this is the case?
4. Summarize the steps involved in conducting the t -test. What is the basis for using the “rule of 2” as a convenient method of evaluating t -ratios?
5. Briefly explain the meaning of the F -test. Why do you think this test is considered to be more important in multiple regression analysis than it is in simple regression analysis?
6. What is *multicollinearity*? How can researchers detect this problem? What is the impact of this problem on the regression estimates? What steps can be taken to deal with this problem?
7. What is the identification problem? What effect will this problem have on the regression estimates of a demand function? Explain.

Forecasting

1. “The best forecasting method is the one that gives the highest proportion of correct predictions.” Comment.
2. Enumerate methods of qualitative and quantitative forecasting. What are the major differences between the two?
3. Discuss the benefits and drawbacks of the following methods of forecasting:
 - a. Jury of executive opinion
 - b. The Delphi method
 - c. Opinion polls
 Each method has its uses. What are they?
4.
 - a. Why are manufacturers’ new orders, nondefense capital goods, an appropriate leading indicator?
 - b. Why is the index of industrial production an appropriate coincident indicator?
 - c. Why is the average prime rate charged by banks an appropriate lagging indicator?

5. Discuss some of the important criticisms of the forecasting ability of the leading economic indicators.
6. Manhattan was allegedly purchased from Native Americans in 1626 for \$24. If the sellers had invested this sum at a 6 percent interest rate compounded semiannually, how much would it amount to today?
7. The compound growth rate is frequently used to forecast various quantities (sales, profits, and so on). Do you believe this is a good method? Should any cautions be exercised in making such projections?
8. Describe projections that use either moving averages or exponential smoothing. Under what conditions can these techniques be used? Which of the two appears to be the more useful?
9. How do econometric models differ from “naive” projection methods? Is it always advisable to use the former in forecasting?
10. You have been asked to produce a forecast for your company’s product, bottled water. Discuss the kind of information you would look for in order to make this forecast.
11. The following are the monthly changes in the index of leading economic indicators during 2001 and January 2002:

January	+1	August	0
February	+1	September	-.6
March	-.3	October	+1
April	-.1	November	+8
May	+6	December	+1.3
June	+2	January 2002	+6
July	+3		

What would be your prediction for the U.S. economy in 2002?

PROBLEMS

Demand Estimation

1. One of the most difficult tasks in regression analysis is to obtain the data suitable for quantitative studies of this kind. Suppose you are trying to estimate the demand for home furniture. Suggest the kinds of variables that could be used to represent the following factors, which are believed to affect the demand for any product. Be as specific as possible about how the variables are going to be measured. Do you anticipate any difficulty in securing such data? Explain.

Determinants of Demand for Furniture	Suggested Variables to Use in Regression Analysis
Price	
Tastes and preferences	
Price of related products	
Income	
Cost or availability of credit	
Number of buyers	
Future expectations	
Other possible factors	

2. You are the manager of a large automobile dealership who wants to learn more about the effectiveness of various discounts offered to customers over the past 14 months. Following are the average negotiated prices for each month and the quantities sold of a basic model (adjusted for various options) over this period of time.

- a. Graph this information on a scatter plot. Estimate the demand equation. What do the regression results indicate about the desirability of discounting the price? Explain.

Month	Price	Quantity
Jan.	12,500	15
Feb.	12,200	17
Mar.	11,900	16
Apr.	12,000	18
May	11,800	20
June	12,500	18
July	11,700	22
Aug.	12,100	15
Sept.	11,400	22
Oct.	11,400	25
Nov.	11,200	24
Dec.	11,000	30
Jan.	10,800	25
Feb.	10,000	28

- b. What other factors besides price might be included in this equation? Do you foresee any difficulty in obtaining these additional data or incorporating them in the regression analysis?
3. The maker of a leading brand of low-calorie microwavable food estimated the following demand equation for its product using data from 26 supermarkets around the country for the month of April:

$$Q = -5,200 - 42P + 20P_X + 5.2I + 0.20A + 0.25M$$

$$(2.002) \quad (17.5) \quad (6.2) \quad (2.5) \quad (0.09) \quad (0.21)$$

$$R^2 = 0.55 \quad n = 26 \quad F = 4.88$$

Assume the following values for the independent variables:

- Q = Quantity sold per month
 P (in cents) = Price of the product = 500
 P_X (in cents) = Price of leading competitor's product = 600
 I (in dollars) = Per capita income of the standard metropolitan statistical area (SMSA) in which the supermarket is located = 5,500
 A (in dollars) = Monthly advertising expenditure = 10,000
 M = Number of microwave ovens sold in the SMSA in which the supermarket is located = 5,000

Using this information, answer the following questions:

- Compute elasticities for each variable.
 - How concerned do you think this company would be about the impact of a recession on its sales? Explain.
 - Do you think that this firm should cut its price to increase its market share? Explain.
 - What proportion of the variation in sales is explained by the independent variables in the equations? How confident are you about this answer? Explain.
4. A manufacturer of computer workstations gathered average monthly sales figures from its 56 branch offices and dealerships across the country and estimated the following demand for its product:

$$Q = +15,000 - 2.80P + 150A + 0.3P_{pc} + 0.35P_m + 0.2P_c$$

$$(5,234) \quad (1.29) \quad (175) \quad (0.12) \quad (0.17) \quad (0.13)$$

$$R^2 = 0.68 \quad SEE = 786 \quad F = 21.25$$

The variables and their assumed values are

- Q = Quantity
 P = Price of basic model = 7,000
 A = Advertising expenditures (in thousands) = 52
 P_{pc} = Average price of a personal computer = 4,000
 P_m = Average price of a minicomputer = 15,000
 P_c = Average price of a leading competitor's workstation = 8,000

- Compute the elasticities for each variable. On this basis, discuss the relative impact that each variable has on the demand. What implications do these results have for the firm's marketing and pricing policies?
 - Conduct a t -test for the statistical significance of each variable. In each case, state whether a one-tail or two-tail test is required. What difference, if any, does it make to use a one-tail versus a two-tail test on the results? Discuss the results of the t -tests in light of the policy implications mentioned.
 - Suppose a manager evaluating these results suggests that interest rates and the performance of the computer (typically measured in millions of instructions per second, or MIPS) are important determinants of the demand for workstations and must therefore be included in the study. How would you respond to this suggestion? Elaborate.
5. You are given the following demand for European luxury automobiles:

$$Q = 1,000P^{-0.93}P_A^{0.75}P_J^{1.2}I^{1.6}$$

where P = Price of European luxury cars

P_A = Price of American luxury cars

P_J = Price of Japanese luxury cars

I = Annual income of car buyers

Assume that each of the coefficients is statistically significant (i.e., that they passed the t -test). On the basis of the information given, answer the following questions:

- Comment on the degree of substitutability between European and American luxury cars and between European and Japanese luxury cars. Explain some possible reasons for the results in the equation.
- Comment on the coefficient for the income variable. Is this result what you would expect? Explain.
- Comment on the coefficient of the European car price variable. Is that what you would expect? Explain.

Forecasting

- If the sales of your company have grown from \$500,000 five years ago to \$1,050,150 this year, what is the compound growth rate? If you expect your sales to grow at a rate of 10 percent for the next five years, what should they be five years from now?
- The sales data for the Lonestar Sports Apparel Company for the last 12 years are as follows:

2001	\$400,000	2007	\$617,000
2002	440,000	2008	654,000
2003	480,000	2009	700,000
2004	518,000	2010	756,000
2005	554,000	2011	824,000
2006	587,000	2012	906,000

- What is the 2001–2012 compound growth rate?
 - Using the result obtained in part *a*, what is your 2013 projection?
 - If you were to make your own projection, what would you forecast?
3. Based on past data, Mack's Pool Supply has constructed the following equation for the sales of its house brand of chlorine tablets:

$$Q = 1,000 + 100t$$

where Q is quantity and t is time (in years), with 2007 = 0.

- a. What is the sales projection for 2013?
- b. The tablet sales are seasonal, with the following quarterly indexes:

Quarter 1	80%
Quarter 2	100%
Quarter 3	125%
Quarter 4	95%

What is the quarterly sales projection for 2013?

4. The sales data over the last 10 years for the Acme Hardware Store are as follows:

2003	\$230,000	2008	\$526,000
2004	276,000	2009	605,000
2005	328,000	2010	690,000
2006	388,000	2011	779,000
2007	453,000	2012	873,000

- a. Calculate the compound growth rate for the period of 2003 to 2012.
 - b. Based on your answer to part *a*, forecast sales for both 2013 and 2014.
 - c. Now calculate the compound growth rate for the period of 2007 to 2012.
 - d. Based on your answer to part *c*, forecast sales for both 2013 and 2014.
 - e. What is the major reason for the differences in your answers to parts *b* and *d*? If you were to make your own projections, what would you forecast? (Drawing a graph is very helpful.)
5. The Miracle Corporation had the following sales during the past 10 years (in thousands of dollars):

2003	200	2008	302
2004	215	2009	320
2005	237	2010	345
2006	260	2011	360
2007	278	2012	382

- a. Calculate a trend line, and forecast sales for 2013. How confident are you of this forecast?
 - b. Use exponential smoothing with a smoothing factor $w = 0.7$. What is your 2013 forecast? How confident are you of this forecast?
6. You have the following data for the last 12 months' sales for the PRQ Corporation (in thousands of dollars):

January	500	July	610
February	520	August	620
March	520	September	580
April	510	October	550
May	530	November	510
June	580	December	480

- a. Calculate a 3-month centered moving average.
 - b. Use this moving average to forecast sales for January of next year.
 - c. If you were asked to forecast January and February sales for next year, would you be confident of your forecast using the preceding moving averages? Why or why not?
7. Office Enterprises (OE) produces a line of metal office file cabinets. The company's economist, having investigated a large number of past data, has established the following equation of demand for these cabinets:

$$Q = 10,000 + 60B - 100P + 50C$$

where Q = Annual number of cabinets sold

B = Index of nonresidential construction

P = Average price per cabinet charged by OE

C = Average price per cabinet charged by OE's closest competitor

It is expected that next year's nonresidential construction index will stand at 160, OE's average price will be \$40, and the competitor's average price will be \$35.

- Forecast next year's sales.
 - What will be the effect if the competitor lowers its price to \$32? If it raises its price to \$36?
 - What will happen if OE reacts to the decrease mentioned in part *b* by lowering its price to \$37?
 - If the index forecast was wrong, and it turns out to be only 140 next year, what will be the effect on OE's sales?
8. An economist has estimated the sales trend line for the Sun Belt Toy Company as follows:

$$S_t = 43.6 + 0.8t$$

S_t represents Sun Belt's monthly sales (in millions of dollars), and $t = 1$ in January 2008. The monthly seasonal indexes are as follows:

January	60	April	110	July	90	October	110
February	70	May	110	August	80	November	140
March	85	June	100	September	95	December	150

Forecast monthly sales for the year 2013.

9. The MNO Corporation is preparing for its stockholder meeting on May 15, 2013. It sent out proxies to its stockholders on March 15 and asked stockholders who plan to attend the meeting to respond. To plan for a sufficient number of information packages to be distributed at the meeting, as well as for refreshments to be served, the company has asked you to forecast the number of attending stockholders. By April 15, 378 stockholders have expressed their intention to attend. You have available the following data for the last 6 years for total attendance at the stockholder meeting and the number of positive responses as of April 15:

Year	Positive Responses	Attendance
2007	322	520
2008	301	550
2009	398	570
2010	421	600
2011	357	570
2012	452	650

- What is your attendance forecast for the 2013 stockholder meeting?
 - Are there any other factors that could affect attendance, and thus make your forecast inaccurate?
10. In the Columbia Gas of Ohio study that forecasted the demand for gas (see p. 155), the company developed the following coefficients for their equation:

Growth rate	.015
Intercept	1376.0
Forecasted temperature	-17.1
Prior day temperature	-3.7
Forecasted wind speed	4.2

Forecast the demand for gas for a day, when the expected average temperature will be 40 degrees, the prior day's average temperature was 37, and the average wind speed is predicted to be 8 miles per hour.



The Demand for White Zinfandel in Los Angeles

This appendix provides a demand analysis using actual data on consumer purchases for specific types of wine. The appendix has three parts. First, the analysis and results are discussed without specific reference to how those results were obtained. Next, we show how the analysis is performed in Excel. Finally, extensions to the analysis are suggested using the data provided.

An Analysis of the Demand for White Zinfandel

As noted in Chapter 5, the data obtained from companies like ACNielsen and Information Resources, Inc. (IRI) are highly proprietary in nature. Fortunately, the author of this section was given permission by IRI to use selected historical data on the wine industry extracted from its InfoScan Reviews. InfoScan data is provided using 4-week time periods (13 periods per year). For this analysis, we shall track sales in five wine categories (or types) over 52 time periods from 1996 to 1999 in the Los Angeles food market. Five statistics are given for each observation: dollar sales; volume sales; unit sales; average price per unit; and all commodity value (ACV) weighted distribution. Unit sales counts the number of bottles sold, regardless of bottle size. Volume sales is equalized on 9-liter cases. Readers can download this data from the Demand with Seasonality Excel App.

Ideally, we would like to calculate demand for a specific wine, for example, Mumm Cuvee Napa, Brut Prestige. Such an analysis would include information regarding Mumm's substitutes such as Domaine Chandon and Piper Sonoma. We could easily accomplish this task with the more detailed data that is available (for a fee) from IRI (indeed, the author has done exactly that analysis for Seagram). The data used in this appendix are more aggregated—information is provided on grape varieties, not specific brands of those varieties. Given aggregated data, the most effective demand analysis will occur for the variety that is most homogeneous. White zinfandel fits the bill. Simply put, there is not much price or perceived quality dispersion

between various brands of white zinfandel. Producers of white zinfandel may dispute this but greater quality and price dispersion exists for the other three grape varieties. White zinfandel is most like a commodity.

As can be seen in Figure 5A.1, unit sales of white zinfandel, Q , are flat across the four-year sample provided by IRI. Each year has a noticeable peak during the Thanksgiving and Christmas holidays (the twelfth and thirteenth time periods) each year. The four price series in Figure 5A.2 depict an interesting pattern. All four price series exhibit an upward trend but the prices for three of the varietals (chardonnay, cabernet sauvignon, and merlot) peak during the holidays. The same is *not* true of white zinfandel, where there is actually a noticeable dip in price during the holidays.

The scatterplot in Figure 5A.3 depicts an inverse relation between price and quantity for white zinfandel. The data “cloud” has a central core with peak observations to the right. As noted in Chapter 4, two functional forms are readily available: linear and constant elasticity. Both forms produce reasonable estimates, but the fit is better and coefficient interpretation is easier for the constant elasticity demand (CED) form. (This superior fit is not surprising if you consider the CED function depicted in Figure 4.1a of the previous chapter. For further discussion of CED estimation, see the Nonlinearities Excel App, CED worksheet. This curve fits the data depicted in Figure 5A.3 nicely.)

A line fit plot from a preliminary regression is provided in order to emphasize the point described in Figure 4.1. The CED *curve* in (Q,P) space in Figure 4.1a is *linear* in $(\ln(Q), \ln(P))$ space in Figure 4.1b. This linearity allows for estimation using regression analysis. This preliminary regression is:

$$\ln(Q_{wz}) = 14.038 - 0.819 \cdot \ln(P_{wz}) + 0.409 \cdot \text{Peak} \\ R^2 = 0.797, \text{SEE} = 0.084$$

Peak is a dummy variable that equals 1 during the twelfth and thirteenth time periods each

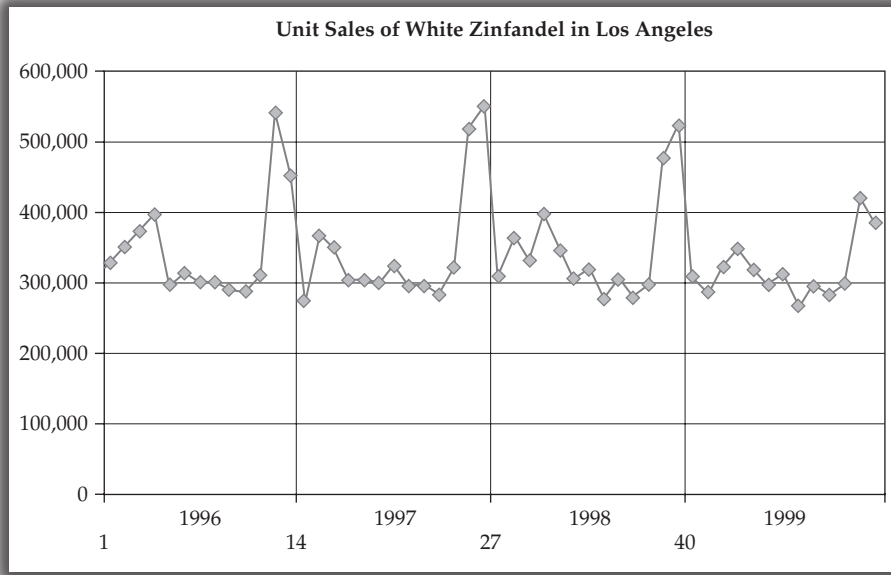


Figure 5A.1 Unit Sales of White Zinfandel in Los Angeles

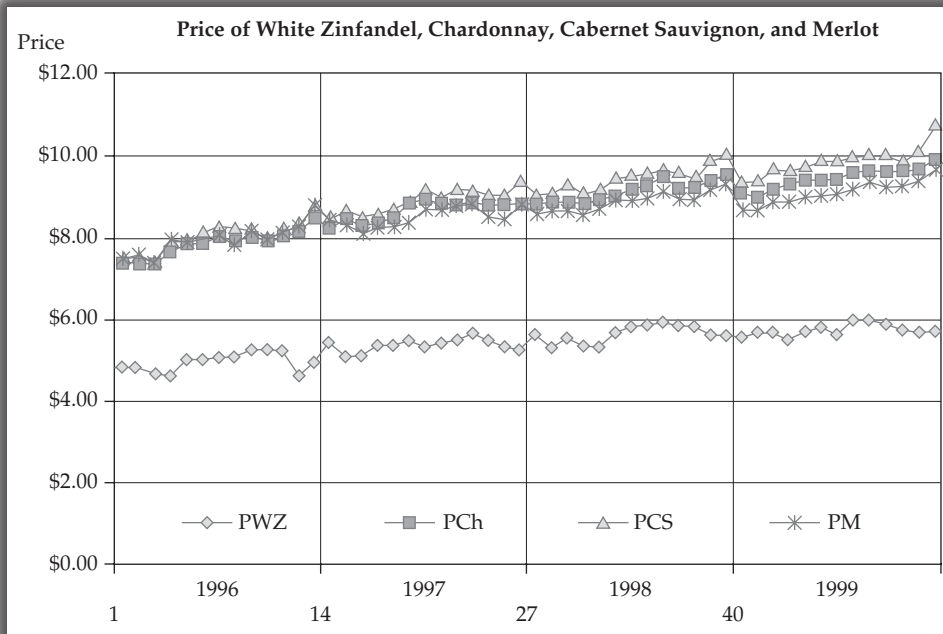
4-weekly observations, 1996–1999, 52 observations total. Time = 1 to 52.

year, otherwise it equals zero. Each coefficient is significantly different from zero, and this regression explains nearly 80 percent of the variance in white zinfandel sales over this period. The line fit plot from this regression is shown in Figure 5A.4.

This shows two lines of predicted $\text{Ln}(Q_{WZ})$ squares in $(\text{Ln}(Q), \text{Ln}(P))$ space—as expected, given that a dummy variable has been added to a univariate equation. Peak sales are approximately 50 percent higher than off-peak sales, all

Figure 5A.2 Price of White Zinfandel, Chardonnay, Cabernet Sauvignon, and Merlot

4-weekly observations, 1996–1999, 52 observations total. Time = 1 to 52.



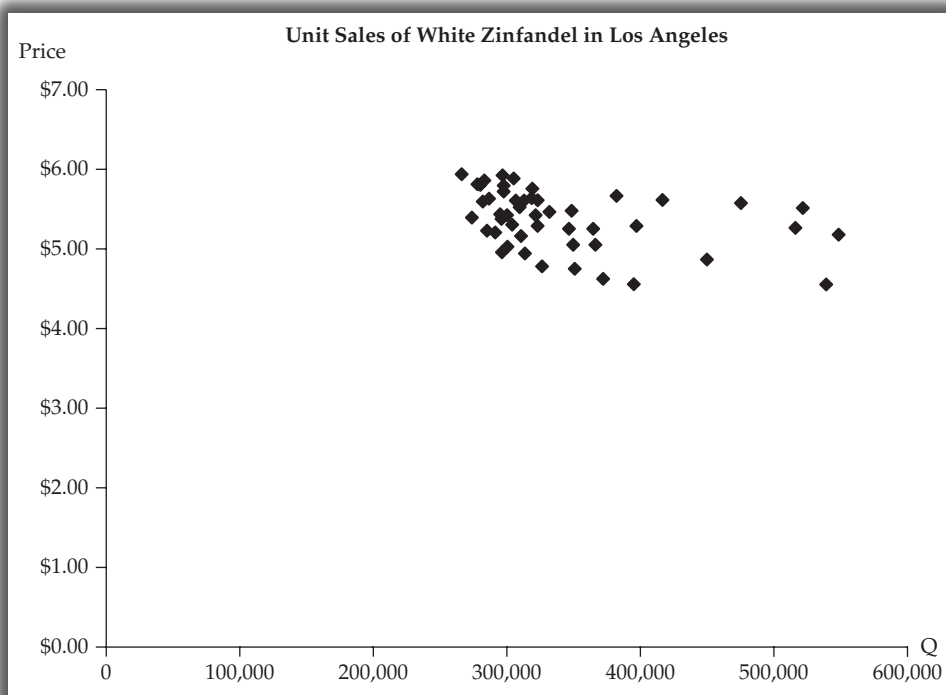


Figure 5A.3 Unit Sales of White Zinfandel in Los Angeles

else held constant.⁴⁸ The inner line is predicted off-peak sales and the outer line is predicted peak sales. This simplified model is *not* the preferred model but is provided to help explain the linearity inherent in CED functions in the log-log format. Line fit plots for the preferred models are also linear but this linearity is obscured in two-dimensional diagrams.

Various equations are estimated based on the presence of various cross-price series in the model. Each has the form:

$$\ln(Q_{WZ}) = a + b \cdot \ln(P_{WZ}) + c \cdot \ln(P_{\text{OtherWine}}) + d \cdot \text{Time} + e \cdot \text{Peak}$$

Time takes on values from 1 to 52. Each regression includes the price of white zinfandel because the good's own price must be included in any demand analysis. Coefficients and *t* statistics for each of the eight models are provided together with adjusted R^2 and standard error of estimates in

Table 5A.1. Each price coefficient in the log-log demand equation is a price elasticity of demand. The coefficient *b* is the own-price elasticity of demand and the *c* coefficients are cross-price elasticities. The regressions are grouped according to how many cross-price independent variables are included in the model. Before discussing the cross-price effects, it is worth noting the consistency of results provided by the independent variables common to all of the models.

Model A expands on the preliminary regression discussed above by including a time counter to test for trend. The coefficient of time, $d = 0.0051$, implies an annual growth rate of 6.8 percent per year, all else held constant.⁴⁹ The other models imply annual growth between 1.7 percent and 6.2 percent. The *t* statistic for the coefficient of time in each model suggests that sales of white zinfandel are flat—the trend that is apparent in Model A is due

⁴⁸This can be seen in Figure 5A.1, where off-peak sales are on the order of 300K units per period and peak sales are about 450K units per period. The factor of proportionality is given by $e^{\text{Peak coefficient}} - 1$. In this regression, that factor of proportionality is $50.6\% = e^{0.409} - 1$.

⁴⁹If *Y* is time series data with *n* time periods per year, you can find the annual growth rate of *Y* using a regression model with $\ln(Y)$ as the dependent variable and time as an independent variable. The annual growth rate in this instance is $g = (e^{\text{coefficient of time}})^n - 1$. In this instance, $6.8\% = (e^{0.0051})^{13} - 1$. The coefficient of time is often called the *periodic growth rate*.

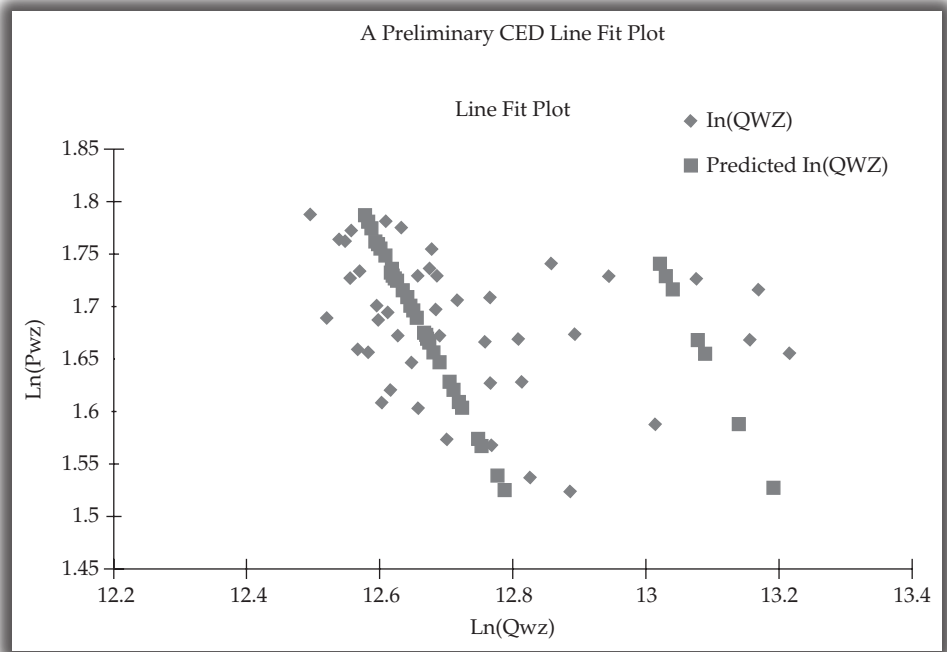


Figure 5A.4 Preliminary CED Line Fit Plot

to the lack of cross-price information in that model. In all specifications, demand for white zinfandel is elastic with own-price elasticity (coefficient b) ranging from -1.82 to -2.36 . The 95 percent confidence interval on a coefficient represents that range of values over which that individual coefficient is likely to occur. The 95 percent confidence interval on b represents the range of likely values of own-price elasticity. In each model, this 95 percent confidence interval does not include -1 , so we are confident in our belief that demand for white zinfandel is elastic.⁵⁰ Peak demand is approximately 40 percent higher than off-peak demand in each specification (coefficient e). The range of proportional increases is 38 percent to 43 percent (38 percent = $e^{0.32} - 1$ and 43 percent = $e^{0.36} - 1$).

Models B–D include a single cross-price independent variable. All three have positive coefficients for c , meaning that all are substitutes for white zinfandel. Model B is the best of the group because of the highest t statistic on the cross-price variable,

highest adjusted R^2 , and lowest standard error of estimate. This result is also reasonable from a market perspective—chardonnay is a white wine, and it is reasonable to assume that it would be a closer substitute for white zinfandel than either cabernet sauvignon or merlot, which are both red wines.

Models E–G include two cross-price independent variables. In this instance, one acts as a substitute and the other acts as a complement to white zinfandel, but only in model F are both statistically significant (merlot is significant at the 10 percent level). Model F is the best specification in Table 5A.1; it has the highest adjusted R^2 and lowest standard error of estimate of all regressions in the table. Chardonnay is a substitute and merlot is a complement to white zinfandel in this specification.

The final specification, model H, includes three cross-prices as well as the price of white zinfandel. Comparing models F and G with H provides a nice example of the multicollinearity problem. As is visually clear from Figure 5A.2, chardonnay, cabernet sauvignon, and merlot prices track one another quite closely. This is especially apparent during peak periods as noted above. This is confirmed by the correlations provided in Table 5A.2. Chardonnay, cabernet sauvignon, and merlot are all highly collinear with correlation coefficients between 0.975 and 0.988. These prices are also correlated

⁵⁰Were we to have brand-based demand for white zinfandel data (for example, the demand for Sutter Home white zinfandel as a function of both Sutter Home's price and Beringer's white zinfandel price), the resulting demand elasticity would be even more elastic. The 95 percent confidence interval on b is in cells **F18:G18** on each regression worksheet in Excel.

Table 5A.1 Constant Elasticity Demand Functions: Unit Sales of White Zinfandel in Los Angeles

i		Intercept	Ln(P_{WZ})	Ln(P_{Ch})	Ln(P_{CS})	Ln(P_M)	Time ⁱⁱ	Peak		
Model		a	b	c_{Ch}	c_{CS}	c_M	d	e	Adj. R_2	SEE
A	Coef.	15.60	-1.82				0.0051	0.36	0.823	0.077
	<i>t Stat</i>	28.13	-5.21				3.22	10.64		
B	Coef.	14.10	-2.36	1.16			0.0016	0.32	0.833	0.075
	<i>t Stat</i>	15.34	-5.44	2.00			0.71	8.52		
C	Coef.	14.78	-2.10		0.62		0.0029	0.33	0.824	0.077
	<i>t Stat</i>	16.58	-4.94		1.17		1.19	7.75		
D	Coef.	15.31	-1.89			0.20	0.0046	0.35	0.819	0.078
	<i>t Stat</i>	15.08	-4.57			0.34	2.30	8.33		
E	Coef.	14.19	-2.34	2.06	-0.95		0.0023	0.34	0.833	0.075
	<i>t Stat</i>	15.35	-5.40	1.87	-0.97		0.95	8.13		
F	Coef.	14.80	-2.36	2.42		-1.60	0.0013	0.35	0.842	0.073
	<i>t Stat</i>	15.29	-5.58	2.77		-1.89	0.58	8.84		
G	Coef.	15.22	-2.07		1.26	-0.88	0.0025	0.33	0.824	0.077
	<i>t Stat</i>	15.13	-4.85		1.46	-0.94	1.03	7.80		
H	Coef.	14.78	-2.35	2.60	-0.26	-1.51	0.0015	0.35	0.838	0.074
	<i>t Stat</i>	15.07	-5.51	2.29	-0.25	-1.61	0.63	8.42		

i. The dependent variable in each model is Ln(unit sales of white zinfandel in L.A.).

ii. 4-weekly observations, 1996–1999, 52 observations total. Time = 1 to 52. Peak is the 12th and 13th time period each year (i.e., Thanksgiving and Christmas).

with white zinfandel prices, but the correlation is not as high. Putting highly correlated independent variables in a regression model produces the multicollinearity problem discussed in Chapter 5. Estimated coefficients vary in magnitude and lose significance relative to models where fewer collinear variables are included. Notice, for instance, that cabernet sauvignon's coefficient changes sign between G and H due to the addition of pricing information on chardonnay. Put another way, the introduction of cabernet pricing in moving from F to H involves both chardonnay and merlot coefficients becoming less significant (reduced *t* statistics). Even though H uses more explanatory variables, it does a less accurate job of prediction as signified by the lower adjusted R^2 and higher SEE of H relative to F. In conclusion, F is the preferred model of demand for white zinfandel given the data available to us from IRI.

Using Excel to Perform the Regression Analysis

A number of available statistical software packages such as SPSS and SAS can be used to perform regression analysis. In this section, we demonstrate

how the most readily available generic software package, Excel, can also be used for this same task. The data used to perform this analysis are located in the **Data** sheet of the Demand with Seasonality Excel App. Five columns of data are provided for each of the five wine types—table wine, cabernet sauvignon (CS), chardonnay (Ch), merlot (M), and white zinfandel (WZ). As noted above, we focus on a comparison of unit sales and price. When performing regression analyses, it is convenient to create a separate worksheet that contains the regression data. This worksheet is called **DataReg**. (In this section of the appendix, **bold-faced** characters refer to Excel names or commands and *italic* characters refer to information (labels, numbers, or equations) you will need to type into cells. *Italics* are also used when *clicking* or *dragging* is required.)

If you want to try doing this analysis yourself, open your own blank document (by *clicking* the **New Blank Document** icon on the **Standard** toolbar in Excel 2003 or the **File** ribbon in Excel 2007), then try to create a file that looks like **Appendix5.xls**. Copy the entire **Data** worksheet to the new workbook by opening a new file, then going back to **Appendix5.xls** to the **Data** worksheet. Highlight the

Table 5A.2 Correlation Matrix of Wine Prices in Los Angeles

	P_{WZ}	P_{Ch}	P_{CS}	P_M
$P_{\text{white zinfandel}}$	1			
$P_{\text{chardonnay}}$	0.891	1		
$P_{\text{cabernet sauvignon}}$	0.862	0.988	1	
P_{merlot}	0.841	0.975	0.981	1

entire worksheet by *clicking* above the **1** and to the left of the **A** in the upper left corner of the spreadsheet. Once this is highlighted, *click* the **Copy** icon on the **Standard** toolbar. Move to the new worksheet, *click* on cell **A1**, and then *click Paste*.⁵¹ Copy individual series using the same strategy: Highlight individual columns on the **Data** worksheet by *clicking* the column label (i.e., the **A** or **B**), then paste to columns on the **DataReg** worksheet. The five IRI series used to produce the above analysis are copied from columns **L**, **R**, **X**, **AC**, and **AD** of the **Data** worksheet to columns **L:P** on **DataReg**. (Note: A group of rows, columns, or cells is denoted using a colon.) The order of columns (variables) has been reorganized so that quantity is first (unit sales of WZ) in column **L**, followed by the prices of WZ, Ch, CS, and M.

Figures are easy to create in Excel using the **Chart Wizard** found on the **Standard** toolbar in Excel 2003 or the **Insert** ribbon in Excel 2007. (Note that copies of Figures 5A.1 to 5A.3 are in **A2:AP67** on the **DataReg** worksheet.) For example, to create Figure 5A.1, highlight cells **L1:L53**, then *click Chart Wizard*. The **Chart Wizard** popup menu appears on the screen. Choose **Line Chart type** and maintain the **Chart sub-type** suggested (this is the **Line with markers displayed for each data value** subtype). *Click Press and Hold to View Sample* at the bottom of the popup menu to see a file that looks similar to Figure 5A.1. (When creating any figure in Excel, it is easy to check to see what you are getting by using this view sample feature.) *Click Finish* to obtain a copy that can be further formatted as required. (We will not focus on formatting

issues in this appendix.) A similar process produces Figure 5A.2 simply by highlighting **M1:P53** prior to using the **Chart Wizard**. Figure 5A.3 is produced by highlighting **L1:M53**, then using the **XY (Scatter) Chart type**. When creating scatterplots, the order of the series matters. Note that column **L** contains Q_{WZ} and **M** contains P_{WZ} , therefore the resulting scatterplot has the standard (Q, P) orientation used to describe demand.

The regressions that form the core of this analysis require a time counter as well as a dummy variable that is turned on during the twelfth and thirteenth periods each year (recall that IRI uses 4-week time periods rather than months for each observation—the exact timing of these observations is given in column **A** of the **Data** worksheet). The CED regressions also use logarithms of price and quantity rather than P and Q . Each of these variables is easily created in Excel.

To create a time counter (like the one in column **E**), type *Time* in **E1**, then type *1* in **E2** and *2* in **E3**. Then highlight **E2:E3** by *clicking* on **E2** and holding down the mouse and moving to **E3** (with the open white + sign showing). Lift the mouse and move to the lower right corner of cell **E3**. Once a black + sign without arrows appears, *click* and *drag* to **E53**. All *dragging* in Excel uses the black + sign. By contrast, if the + sign has arrows, then *dragging* will move the cell (or group of cells) rather than replicate the pattern.

The Peak dummy variable in column **F** is created by doing one year's worth of the series, then simply replicating what was done. Type *Peak* in **F1** and *0* in **F2**. *Drag* the zero down to cell **F12** (next to the 11 in **E12**). Type *1* in **F13** and **F14**. This is a full year of the dummy variable, Peak. From here, *click* on cell **F15**, type = *F2*, and press *Enter* to replicate the pattern. Notice that a 0 appears in cell **F15** even though it is based on an equation rather than a number entered from the keyboard. *Click* on cell **F15** and *drag* it to **F53**. Once you release the mouse, you will have a variable that takes on the value 0 in

⁵¹In some versions of Excel, you may see ##### in some cells after copying. This simply means that columns have not been formatted with sufficient width. To reformat column widths (or row heights) move the mouse to the right column boundary that you wish to change at the column icon (for example, if you wish to widen **A** move to the **A/B** boundary between the letters **A** and **B**) until you see a **double arrow with bar**, *click*, and move to adjust the width as necessary.

periods 1–11 and 1 in periods 12 and 13 for all four years.

The dependent variable, $\text{Ln}(Q_{WZ})$, is the natural logarithm of unit sales of white zinfandel. Type the label in **A1**, then in cell **A2** type the equation $=\text{ln}(L2)$ and press *Enter*. *Click* on **A2** and *drag* to **A53** to create this series. Do the same to create logarithmic versions of each of the price series.

You are now ready to run a regression. Excel requires that all independent variables be contiguous, therefore you will have to move series each time you wish to run a different regression. Each regression shown uses the same dependent variable, $\text{Ln}(Q_{WZ})$, but the independent variables vary from model to model. The **DataReg** worksheet is set up ready to run the preferred model, **F**. This model is shown on worksheet **CED F**. To create this regression yourself, start from the **DataReg** worksheet and *click* on the **Tools** dropdown menu in Excel 2003 or the **Data** ribbon in Excel 2007, then *click* **Data Analysis**.⁵² Once you *click* **Data Analysis**, the **Data Analysis popup menu** appears. Scroll down to **Regression** and *click* **Regression**, then *click* **OK**. The **Regression popup menu** appears. This menu allows you to describe what you want in the regression.

Regression requires you to identify the dependent and independent variables. The dependent variable is what you are trying to explain, and the independent variables are what you are using to explain the dependent variable. Unfortunately, Excel calls these concepts **Y range** and **X range** rather than dependent and independent variables. In “Excel speak,” **Y range** is the dependent variable and **X range** is the independent variable(s).⁵³ If you are wondering why this matters, it is because economists are exactly backward with respect to the most basic geometric models. Demand is quantity as a function of price—quantity is the dependent variable but it is on the X axis (and price is the independent variable despite being on the Y axis). With that said, *click* on the **Input Y range** area and highlight **A1:A53**. Next, *click* on the **Input X range** area and highlight **B1:F53**. *Click* **Labels**. (Always use **Labels**,

otherwise the regression will be difficult to decipher because it will call the independent variables X variable 1, X variable 2, etc.) I have also *clicked* **Residuals**, **Residual Plots**, and **Line Fit Plots** but that is not necessarily required.⁵⁴ *Click* **OK** and the regression is provided in a new worksheet.

All of the regressions in Table 5A.1 as well as the preliminary regression are obtained by simply adjusting the independent variables that are included in the **Input X range** of the regression model. Each time you adjust the independent variables to create a new model you must follow one rule—the independent variables must be contiguous. You cannot reference series one at a time that you wish to include in the set of independent variables. For example, suppose you want to run the model that includes all three “other” prices (model H) and you wish to have them in the order shown in Table 5A.1. In this instance, you must insert $\text{Ln}(P_{CS})$ between the prices of chardonnay and merlot (in columns **C** and **D**). To accomplish this you can simply move columns **D:F** over one column by highlighting all three columns (*click* on the **D** at the top of the spreadsheet, then hold the mouse down until it runs over **E** and **F** as well). Once all three columns are highlighted, they can be moved by moving the mouse to the right edge of column **G** until you see a large open diagonal white arrow pointing at a black + with arrows. *Click* and hold down the mouse, then *drag* all three columns one to the right (so that column **D** is now empty). Next *click* on column **J** and move it into column **D** using the same technique you just used on the other columns. When you run this regression, you need not redo the entire **Regression popup menu**—you need only change the **F** to a **G** in the **Input X range**, then *click* **OK**. This strategy streamlines running multiple regression models.

The nine regressions described above are shown on separate worksheets labeled **CEDPreliminary** and **CED A—CED H**. (To change a worksheet name—from **Sheet #**—simply *double click* on the tab and type the new name.) Figure 5A.4 is a **Line Fit Plot** from the **CEDPreliminary** worksheet, but

⁵²If **Data Analysis** is not available, *click* **Add-Ins**, then *click* **Analysis ToolPak** and **Analysis ToolPak—VBA**, then *click* **OK**. **Data Analysis** should now be available from the **Tools** dropdown menu. (If it is not, **Save** the file you have been working on Quit Excel, then open Excel. **Data Analysis** will now be available.)

⁵³Only one dependent variable is allowed in regression analysis (the **Input Y Range** must be univariate) but multiple independent variables are allowed (the **Input X Range** may be multivariate) in regression analysis.

⁵⁴**Residuals** is useful to *click* even if you do not plan to use residuals because it is one of the few places on Excel’s regression output where the dependent variable is listed. In this instance this is not much of an issue because there is only one dependent variable for all of the regressions, but often this will not be the case. To see where Excel “hides” this important information, see, for example, cell **B28** on worksheet **CED F**. The other place where this information is available is on the **Line Fit Plots**.

it is worth noting that the axes have been reversed (by changing some of the **Source Data** within the **Chart Wizard**) so that it has the standard orientation expected of a demand curve in (Q,P) space rather than (P,Q) space.⁵⁵ Of course, the Ps and Qs are actually Ln(P)s and Ln(Q)s. All of the **Line Fit Plots** on other worksheets have not been reversed—in each, the vertical axis is the dependent variable: Ln(Q_{WZ}).

Table 5A.1 is created by copying **Coefficient** and **t Statistic** information from each of the regression worksheets (starting in **B16** and **D16**) onto a new worksheet. This is accomplished in **B1:R8** of the **Table 5A.1** worksheet. Note that each regression has different cross-price information. As a result, take care to line up the data. For example, cells **C6:H6** are blank because the price of merlot is not included in models A–C. The regressions in **B1:R8** are oriented vertically (with two columns per regression) but Table 5A.1 depicts the regressions horizontally. To switch rows and columns,

⁵⁵To do this, **RIGHT** click on a chart you wish to change and a series of options emerge. The one used in this instance is **Source Data**. The X and Y values for both series must be switched on the **Series** tab.

use **Copy, Paste Special** (from the **Edit** pulldown menu), **Transpose**. From here the completed table is easily obtained by adding further labels and formatting and by copying adjusted R^2 and SEE information from **B6:B7** of each worksheet to **J13:K27**.

The coefficient of Time and the Peak coefficient are interpreted in the text based on the information provided in Table 5A.1. That interpretation uses calculations described in the text and performed in cells **L13:M27**. It is worth noting that only two equations need to be entered into Excel (in **L13:M13**). *Drag* these equations to **L27:M27**. In the process, the even-numbered rows are nonsense (since they are based on exponentiated t statistics), but these rows can be quickly deleted, leaving calculations of relevant statistics for each regression in odd-numbered rows.

The final part of the analysis is the correlation matrix provided as Table 5A.2. This is easily obtained from the **DataReg** worksheet. Once on this worksheet, *click* **Tools, Data Analysis, Correlation**. *Click* on the **Input Range** area. Highlight **M1:P53** and *click* **Labels in First Row** and **OK**. The correlation matrix will appear on a new worksheet. This is easily formatted to look like Table 5A.2.

Understanding Consumer Behavior Through Testing*

Test-and-learn is the new mantra in the tech world.⁵⁶ Be it new product features, new policy ideas, marketing messaging, or email content, test-and-learn is gaining momentum, graduating from simple A/B tests to complex multivariate tests. When it comes to fundamentally understanding consumer behavior and preference, however, this approach has not lived up to its promise in the industry. The primary reason is that, due to the very nature of this approach, one has to develop good hypotheses from the beginning. No matter how appropriate or powerful the statistical technique or how rich the data set, a researcher's work is only as useful as the hypotheses selected for testing. Therefore, developing the right hypotheses to test is the key to understanding consumer behavior. To understand how to develop creative, effective hypotheses, we must be clear on why we are testing in the first place.

Three Reasons for Testing

1. We Are Not Our Customers

A senior executive at a large payments company asked a group of about 100 employees a simple question: "If you were at a store buying something for less than \$50, how many of you would prefer to pay using cash?" About 10 percent raised their hands. Then he went on to share new market research suggesting that 80 percent of U.S. consumers (potential customers of this payments company) state that they prefer to pay using cash for face-to-face transactions of less than \$50.

This huge behavioral gap between employees of this company and its potential customers reinforces the point that we are not our customers; that is, we are not a representative sample of the base that includes our customers. Therefore, we must let go of

the belief that we can sit in a conference room, even with decades of collective experience among us, and dream up what customers would want, like, or prefer.

If we cannot trust ourselves to know what customers want, why not ask them?

2. Customers Are Unable to Articulate What They Want

Customers are particularly poor at articulating what they want, primarily because they don't know what is possible and what isn't. This is the reason why customer-focused companies such as Proctor and Gamble and Intuit like to passively observe customers using their products. It gives them valuable insights for improving their products and services.

A large producer of consumer goods released a new laundry detergent in a developing country in Latin America. Sales were not doing well despite a big advertising budget and objective evidence of superiority of their product over the competition. So company employees followed their customers home to observe how they used this particular product. These visits were nothing short of a revelation. This was a country where women had to fill buckets of water and bring it to the place where they washed the clothes. Based on this simple but illuminating observation, the company decided to create a high-efficiency detergent that required half the amount of water as before. This reduced the labor of these women by half, and the product was an instant success. The customers of this company would not have asked for a high-efficiency detergent because they did not know that manufacturing such a product was possible. A quote from Henry Ford also illustrates this point: "If I had asked people what they wanted, they would have said faster horses."

Customers are often unable to articulate what they want, primarily because they don't know the possibilities and they have not experienced anything similar. As an old saying goes something like this: "People don't know what they don't know."

*Appendix 5B was written by Mukul Patki, manager of analytics at PayPal.

⁵⁶For an excellent introduction to business experiments and analytics, see Eric T. Anderson and Duncan Simester, "A Step-by-Step Guide to Smart Business Experiments," Harvard Business Review, March 2011, <http://hbr.org/2011/03/a-step-by-step-guide-to-smart-business-experiments/ar/1>

If we cannot predict what customers want, and customers are unable to articulate what they want, what is the solution? Why not create a few options and ask them to choose? Herein lies the holy grail of experimentation: developing the right hypotheses to test and ensuring that one of the ideas we are testing is the magic answer. Development of creative and effective hypotheses relies on our understanding of consumer behavior: understanding that people's responses are often irrational, but that there can often be a "method to the madness." This brings us to the third reason why we need to test in the first place.

3. Customers May Respond to Our Offers and Propositions Irrationally

The field of behavioral economics sheds ample light on how irrational the responses of most people can be when they are faced with certain kinds of decisions. Awareness of this body of knowledge often allows for the development of hypotheses (and treatments addressing them) that are effective and have a high probability of producing a product or a service that resonates with consumers and hence creates a disproportionate impact on the business.

Traditional Economics Versus Behavioral Economics

Traditional economics assumes that all parties, given a choice, would choose the economically optimal option for themselves. However, decades of research in psychology and neuroscience have shown that most people make decisions based on emotions, fears, and irrational considerations and that they can be affected by factors such as advertisements, perceptions, and sometimes completely unrelated events. Such decisions are often economically suboptimal. This "irrational" or nonoptimal behavior is the premise of behavioral economics.

For researchers in consumer behavior, the good news is that the consumer decisions we make may often be irrational, but they are predictable. Hence, the apt title of the pioneering book in this field by Dan Ariely, *Predictably Irrational*.⁵⁷ In his book, Ariely

points out dozens of principles of behavioral economics, with scores of studies backing them up. All of these studies prove not only the irrationality of our behavior, but also the predictability of it. Below are five examples of the principles of behavioral economics.

1. Anchoring

Can your social security number (SSN) determine how much you pay for a commodity? Your first reaction may be that such an idea is preposterous. But is it? Consider the following experiment. The audience in one of my talks recently was asked to pretend that they were bidding on a particular consumer product (in this case, the overhead projector that I was using for my presentation). As part of the bidding process, participants were asked to write on a piece of paper the answer to the following three questions:

1. What are the last three digits of your SSN? (Let's say 1, 2, and 3.)
2. Would you be willing to pay this amount in dollars for this product (i.e., \$123)?
3. What is the maximum amount you would be willing to pay for this product?

When the results were analyzed, a huge correlation was found between the last three digits of a respondent's SSN and the maximum amount he or she was willing to pay. Specifically, participants with last three digits SSN in the highest 25th percentile (i.e., 750 to 999) were willing to pay five times more than the participants whose SSN ended between 000 and 250. How can this be?

This phenomenon is known as **anchoring**, which means that consumers have a very poor sense of what a commodity is truly worth. Thus, they look for ways to anchor on to something. The moment they were forced to think about their willingness to pay a certain price for my overhead projector (namely, the last three digits of their SSN in U.S. dollars), they started gravitating toward that number. The maximum amount they were eventually willing to pay ended up being closer to the last three digits of their SSN than it would have been otherwise.

The concept of anchoring can be applied in various situations. Suppose you are shopping for a television. Would it make a difference if the first TV sets you saw were \$3,000 versus \$300? This might help to explain why stores often display the most expensive television sets at the front.

⁵⁷Dan Ariely, *Predictably Irrational*, New York: Harper Collins, 2008. For other good books on this subject, see Jonah Lehrer, *How We Decide*, New York: Mariner Books, 2009; Robert Cialdini, *Influence: The Psychology of Persuasion*, New York: Harper Business, 2006; and Richard Thaler, *Nudge: Improving Decisions About Health, Wealth and Happiness*, New York: Penguin Books, 2009.

2. The Power of Free

Can reducing the price of two commodities by the exact same amount completely reverse consumer preference of one over the other? Traditional economics would predict that it cannot. But let us explore if our irrational minds prove otherwise.

A group of researchers offered participants of a study a choice between purchasing a Hershey's Kisses chocolate for 1 cent or a Lindt Lindor chocolate truffle for 15 cents. The participants recognized this as a good deal because the price difference in a supermarket is greater than 14 cents between the two options. Thus, most participants chose the Lindt truffle. In effect, what they were saying is that the incremental joy of consuming the exotic truffle over Hershey's Kisses is more than the pain of spending an incremental 14 cents. However, when the price of both was reduced by 1 cent, thus making Hershey's Kisses free and the Lindt Lindor price 14 cents, their preference completely reversed. An overwhelming majority chose the Hershey's Kisses.

What happened here? It's the same incremental joy for the same incremental amount spent, but consumer reaction is not the same. Why? Our response to price reduction becomes very nonlinear when the price reaches zero. We just love the word *free*. It evokes unreasonably positive feelings in the brain. Just the sight of the word *free* releases large quantities of dopamine in our brain to make us feel happy, and we respond irrationally.

So how does this occur in the real world? We are inundated with free offers every day and may believe that this does not affect us. But consider two economically identical deals—one offering “buy one get one free” and the other “50 percent off if you buy two.” Which one are you more likely to respond to?

3. Dominated Alternatives

If I ask you to choose between two legitimate options, can I increase the chances that you will buy the option I want you to choose by introducing you to a third option as a decoy?

Potential customers of *The Economist* were given the two subscription offers: an “online only” annual subscription for \$56, and “online + print” annual subscription for \$125. A large majority of people chose the first option (online only for \$56). The publishers wanted people to choose the second option (online + print for \$125). They introduced a third decoy option that they knew nobody would prefer: \$125 for print only. It was no surprise that

no one chose the third option. But something unexpected happened. An overwhelming majority now chose the second option (\$125 for online + print). The mere introduction of this third option, made the second option look very attractive. Who would choose print only for \$125, when she or he could get the online version (a perceived \$56 value) for free!

What happened here? This goes back to the idea that consumers have a very poor understanding of what a commodity is truly worth. They had no idea what a print or an online subscription of *The Economist* is truly worth in dollars. The first scenario with two options gave nothing to compare either option to, but the introduction of the third option made options 2 and 3 comparable. Option 2 was then significantly better because it appeared that consumers were getting an online version for free. Option 1 couldn't be compared to anything, so it was left out.

This principle has been demonstrated successfully in many different scenarios. The next time you are evaluating vacation packages, for example, pay attention to how different options are being positioned. Marketing professionals have figured this out through experience, even if they do not present the options with a clear explanation of their reasoning.

4. Irrational Value Assessment

Are you more likely to admire a \$5 bottle of wine if I lie to you and tell you that it costs \$45? Research says that you are. And here is the interesting part: It is not because you may be trying to show off. You will truly enjoy the \$45 wine more than the \$5 one, even though they are identical.

Consider this experiment. Members of the Stanford Wine Club were invited to taste five bottles of wine and rate them based on their preference. However, there were just three different wines in those bottles: Two wines had two bottles each. Each bottle was marked only with the price tag and nothing else. Some of the same wines were marked at significantly different prices. For example, the \$5 wine and the \$45 wine were actually the same, the true cost being \$5.

There was a clear correlation between the rating of the wine and the price tag: More expensive wines got systematically higher ratings. So the \$45 bottle of wine got a significantly higher rating than the \$5 bottle, although they were the exact same wine.

In another experiment, the same group was asked to rate the same wines again. This time, even the price tags were absent. The cheapest wine was ranked the highest in this case.

Before we start calling these wine experts snobs, consider this. Prozac was tested against a placebo. The placebo was sold at a higher price (\$2.50 per pill) than Prozac (\$2.00 per pill). The placebo outperformed Prozac!

Consider another experiment, where students were given a caffeine-and-sugar-rich drink that was supposed to improve their alertness and focus in the short term. Their task was to solve as many puzzles as they could. Half of the group was asked to pay the full price of the drink, and the other half was given a significant discount on the price. The group that got the drink at a discounted price solved 30 percent fewer puzzles. This result has been consistent in multiple such studies over time.

It turns out that we inherently expect cheaper products to be inferior. This feeling runs so deep, and its effect is so profound on our brain, that the cheaper product does perform more poorly. It becomes a self-fulfilling prophecy. So the people from the Stanford Wine Club were not being snobs when they rated the ostensibly more expensive wines as tasting better. They truly did enjoy the wines with higher price tags more. This was demonstrated by increased activity in the prefrontal cortex of the brain when the same experiment was done under an fMRI machine. Consumers of Prozac, deep within, expected a poorer performance compared to the more expensive placebo. This expectation and conviction were so strong that they did create inferior performance in the body.

5. Decision Paralysis

Can reducing the number of options available to consumers actually increase sales? Studies indicate that it can. For example, in one study, researchers sat down in a supermarket with bottles of jam on display. The expectation was that some users would stop by, fewer would taste, and yet fewer would purchase. One group sat with six varieties on display, and the other had twenty-four varieties on display. While more people stopped by in the case of the twenty-four-jar display, the percentage that bought was ten times less than the percentage in the six-jar scenario (3 percent versus 30 percent). Another example pertains to a company with a voluntary savings program. The company found that participation in the program fell by 2 percent for every ten mutual funds added to it.

The general belief is that consumers want more choices. But consumer behavior studies such as the one just mentioned show otherwise. When faced with too many options, however, consumers are unable to evaluate them all and they decide not to buy at all.

At the heart of these findings is our inability to process too much information. This concept is known as cognitive load, which incidentally does have a magic number: 7 ± 2 . This is the range of the number of items the human brain can simultaneously process. Consider the study where participants were falsely told that they were participating in a study on long-term memory. They were asked to memorize a number, walk down the hall, wait for some time, and repeat the number from memory to a different researcher in a different room. Half the group members were given a two-digit number, and the others were given a seven-digit number. But as the participants walked down the hallway, there were refreshments available with a choice of a decadent chocolate cake or a cup of fresh fruit. This was the real test—exercising self-control when the mind is occupied. The study found that a majority of participants in the seven-digit group chose the cake, while a majority in the two-digit group chose fruit.

There is a neurological explanation for this behavior. The part of the brain (the prefrontal cortex) that is occupied with memorizing irrelevant illogical information such as random digits is the same part that is charged with exercising self-control. Remembering seven digits is a tough task because the number of digits is approaching our cognitive limit. The brain is so preoccupied with trying to remember these numbers that it literally does not have the capacity to exercise self-control.

Utilizing Behavioral Economics in the Business Environment

The previous section examined a few principles of behavioral economics that shed some light on the irrationality of consumer behavior. How can businesses take advantage of this knowledge?

The subject area of behavioral economics fits into a broader framework for experimentation that involves systematic isolation and elimination of hypotheses in an iterative fashion that leads to generating valuable insights about behavior and preferences of customers. While experimentation in the modern industrial world is a subject of much detailed discussion, the following broad framework

should help readers understand the place of behavioral economics in it.

1. **Prelaunch Analysis.** Prelaunch analysis refers to a broad set of activities that must happen before a test is launched. The launch of a test refers to the date when this new offer (or functionality, or feature) is made available to a small group of customers. These activities include defining the goal for testing, developing hypotheses to test, and designing an experiment that enables clear unambiguous evaluation of these hypotheses postlaunch.
 - a. *Define the goal for testing.* This key here is to be specific. Clearly define the success metric and the success criteria. For example, we can say the following: *We expect revenue per customer to increase by 5 percent relative to the control group in the ninety days postlaunch, and we expect to detect such a difference with 95 percent confidence.* This goal is a good one because it is precise and specific.
 - b. *Develop hypotheses.* Hypotheses are ideas we believe can take us to the goal defined above. A good example of a hypothesis is the following: *providing a \$5 incentive to all customers who are likely to churn would produce the desired result.* Behavioral economics plays a key role in this step, which arguably is the most crucial step in the entire process.
 - c. *Design an experiment.* This step gets into the nuts and bolts of testing. Here are some questions that may need to be answered: How many variants of this test are we considering (e.g., \$5 cash back, \$5 coupon, \$10 cash back, \$10 coupon)? How many segments are we targeting (e.g., engagement level, type of account, gender, age of account)? How many customers/transactions do we want in each cell to achieve statistical significance within

the specified period? How long do we want to run the test? Is this a feasible design or do we need to negotiate for a smaller and simpler design?

2. **Postlaunch analysis.** This step assumes that the test was launched with no unexpected or unintended behavior. The goal is to analyze the data from the test group and compare it to the data from the control group. Most traditional analytics methodologies using time-series, confidence intervals, hypothesis testing, impact estimation, and so on, are applicable here.

Conclusion

A test is only as good as the hypothesis it tests. Developing creative, effective hypotheses is the key to a successful experimentation program to understand consumer behavior. When launching a test, the hope is that one of the hypotheses being tested is the “magic answer.” Behavioral economics provides a framework that helps businesses to systematically develop creative treatments, offers, products, or features that resonate with customers. Sometimes certain treatments will resonate more with specific segments of the customer base, and something entirely different might work best for the remainder of the customer base. There is much nuance to doing this the right way. The general overview in this appendix provides some information about how to find the answers to the questions that businesses are always asking: “What are our customers thinking? How can we provide goods and services that they really want or need?”

The Theory and Estimation of Production

Learning Objectives

Upon completion of this chapter, readers should be able to:

- Define the production function and explain the difference between a short-run and a long-run production function.
- Explain the “law of diminishing returns” and how it relates to the Three Stages of Production.
- Define the Three Stages of Production and explain why a rational firm always tries to operate in Stage II.
- Provide examples of types of inputs that might go into a production function for a manufacturing company and for a service company.
- Describe the various forms of a production function that are used in the statistical estimation of these functions.
- Briefly describe the Cobb-Douglas function and cite a few statistical studies that used this particular functional form in their analysis.
- (Appendix 6A) Discuss the key differences between productivity in the production of goods vs. productivity in the production of services, and list important factors that can help to improve productivity in services.
- (Appendix 6B) Describe isoquant and isocost curves and explain how the tangency of these curves help to determine the optimal combination of inputs for a firm to utilize.
- (Appendix 6C) Use the tools of calculus to determine a firm’s optimal combination inputs.

The Situation

The meeting of Global Foods’ top production managers must have been important because it was attended by CEO Bob Burns as well as by Jim Hartwell, vice president of manufacturing. Its purpose was to go over the plans for rolling out the company’s new bottled water product, “Waterpure.” Christopher Lim, manager of the company’s largest bottling plant, in St. Louis, Missouri, knew that this product was considered to be the hope of the sagging beverage division.

(continued)

(continued)

He also suspected that there was going to be more than the usual plans for the bottling operation because of the additional presence of Nicole Goodman, senior VP of marketing.

The meeting began with a few opening remarks by the CEO about the phenomenal growth of the bottled water industry and how and why Global Foods hoped to get into this market. Chris noticed that all but a few of the 15 people at the meeting had their own bottle of water in front of them and thought that this, more than any other statistic, made the point. Then came the surprise. The opening presentation was given by the senior VP of marketing, not manufacturing.

Nicole Goodman began to make a compelling case for a radical change in the packaging of this product. “You all have seen how Coca-Cola has spent millions of dollars in advertising and packaging in order to focus on the shape of its original glass bottle and how Pepsi-Cola recently changed the look of its label and logo. But packaging can be considered even more important in the market for bottled water, because after all what are we selling?

“What I propose is that we combine the tradition of Coca-Cola with the design innovation of Pepsi-Cola and a French company that makes another beverage called Orangina to create our own distinct packaging. We propose to sell our water in a green glass bottle shaped like a bottle of Champagne. To go along with this packaging, we intend to advertise our product as ‘the Champagne of bottled water.’”

“How original,” thought Chris. “I seem to recall that some time ago a beer company used a similar tag line in their advertising. But forget the advertising, doesn’t she realize how expensive it is to use glass rather than plastic? There has to be a reason why both Coke and Pepsi don’t use glass for most of their bottling, particularly here in the States.”

As if in anticipation of Chris’s negative thoughts, Nicole continued by saying, “I know some of you might be thinking that this Champagne idea is not very original. But our market research indicates that people between the ages of 15 and 35 are the key consumers of bottled water, and so most of our potential customers will be too young to remember that Miller beer commercial. We in marketing realize that it will be you in manufacturing that will actually implement our ‘creative’ ideas. We know that this will present some interesting production challenges in terms of the setup of the fill lines, the stacking of the cases for shipment and delivery, and so on. But I know you’re up to the challenge. We want to begin a pilot program in one of our bottling plants. In consulting with both Bob and Jim, they both recommended strongly that we begin with our bottling plant in St. Louis. Chris, we want you to put together a plan for rolling out our new product in the next quarter.”

THE PRODUCTION FUNCTION

Christopher Lim’s concerns are certainly justified. No matter how much revenue is generated by the marketing plan, if the cost of production cannot be contained, the company will not be able to earn an acceptable level of profit. In economics, the analysis of cost begins with the study of the production function. The **production function** is a statement of the relationship between a firm’s scarce resources (i.e., its **inputs**) and the output that results from the use of these resources. Economic cost analysis can then be seen as the application of a monetary unit such as dollars to measure the value of this input usage in the production process. Therefore, the discussion of costs in Chapter 7 requires an understanding of how inputs are put together to produce output as well as an understanding of how the cost minimizing method of producing a given level of output is determined. Chapter 6 examines both of these necessary prerequisites to understanding cost.

In mathematical terms, the production function can be expressed as:

$$Q = f(X_1, \dots, X_k)$$

where Q = Output

X_1, \dots, X_k = Inputs used in the production process

Note that we assume this relationship between inputs and output exists for a specific period of time. In other words, Q is *not* a measure of accumulated output over time.

There are two other key assumptions that you should be aware of. First, we are assuming some given “state of the art” in the production technology. Any innovation in production (e.g., the use of robotics in manufacturing or a more efficient software package for financial analysis) would cause the relationship between given inputs and their output to change. Second, we are assuming whatever input or input combinations are included in a particular function, the output resulting from their utilization is at the maximum level. With this in mind, we can offer a more complete definition of a production function:

A production function defines the relationship between inputs and the maximum amount that can be produced within a given period of time with a given level of technology.

For Christopher’s company, the *X*s could represent raw materials, such as carbonated water, sweeteners, and flavorings; labor, such as assembly line workers, support staff, and supervisory personnel; and fixed assets, such as plant and equipment.

For purposes of analysis, let us reduce the whole array of inputs in the production function to two, *X* and *Y*. Restating Equation (6.1) gives us

$$Q = f(X, Y)$$

where *Q* = Output
X = Labor
Y = Capital

Notice that, although we have designated one variable as labor and the other as capital, we have elected to keep the all-purpose symbols *X* and *Y* as a reminder that any two inputs could have been selected to represent the array.

As stated earlier, in economic analysis the distinction between the short run and the long run is not related to any particular measurement of time (e.g., days, months, or years). Instead it refers to the extent to which a firm can vary the amounts of the inputs in its production process. Thus, a **short-run production function** shows the maximum quantity of a good or service that can be produced by a set of inputs, assuming the amount of at least one of the inputs used remains unchanged. A **long-run production function** shows the maximum quantity of a good or service that can be produced by a set of inputs, assuming the firm is free to vary the amount of *all* the inputs being used.

A hypothetical production function with two inputs is displayed in Table 6.1. The numbers in the matrix indicate the amount of output that would result from various

Table 6.1 Representative Production Table

Units of Y Employed	Output Quantity (<i>Q</i>)							
8	37	60	83	96	107	117	127	128
7	42	64	78	90	101	110	119	120
6	37	52	64	73	82	90	97	104
5	31	47	58	67	75	82	89	95
4	24	39	52	60	67	73	79	85
3	17	29	41	52	58	64	69	73
2	7	17	29	40	48	52	54	52
1	4	8	14	20	27	24	21	17
	1	2	3	4	5	6	7	8
	Units of X Employed							

combinations of X and Y . For example, the use of 2 units of X and 2 units of Y yields 17 units of output. Adding one more unit of X while holding constant the amount of Y yields an additional 12 units of output ($Q = 29$). Increasing both X and Y by 1 unit yields 41 units of output. The additional 1 unit of X with Y unchanged is considered to be a “short-run” change. An increase in both inputs by 1 unit is an example of a “long-run” change.

A SHORT-RUN ANALYSIS OF TOTAL, AVERAGE, AND MARGINAL PRODUCT

Before we go on to a more detailed analysis of the production function, certain key terms employed throughout this chapter should be clarified. First, economists use a number of alternative terms in reference to inputs and output:

Inputs	Output
Factors	Quantity (Q)
Factors of production	Total product (TP)
Resources	Product

Second, in the short-run analysis of the production function, two other terms besides the quantity of output are important measures of the outcome. They are **marginal product (MP)** and **average product (AP)**. If we assume X to be the variable input, then

$$\text{Marginal product of } X = MP_X = \frac{\Delta Q}{\Delta X}, \text{ holding } Y \text{ constant}$$

$$\text{Average product of } X = AP_X = \frac{Q}{X}, \text{ holding } Y \text{ constant}$$

In other words, the marginal product can be defined as the change in output or **total product** resulting from a unit change in a variable input, and the average product can be defined as the total product (TP) per unit of input used.

The data in Table 6.2 shows a short-run production function. Here we see what happens to output when increasing amounts of input X are added to a fixed amount of input Y ($Y = 2$). Table 6.3 restates this information by focusing on the impact on

Table 6.2 Short-Run Changes in Production Showing Factor Productivity

Units of Y Employed	Output Quantity (Q)							
8	37	60	83	96	107	117	127	128
7	42	64	78	90	101	110	119	120
6	37	52	64	73	82	90	97	104
5	31	47	58	67	75	82	89	95
4	24	39	52	60	67	73	79	85
3	17	29	41	52	58	64	69	73
2	7	17	29	40	48	52	54	52
1	4	8	14	20	27	24	21	17
	1	2	3	4	5	6	7	8
	Units of X Employed							

Table 6.3 Short-Run Production Function with $Y = 2$

Units of Variable Input X	Total Product Q or TP	Marginal Product of X MP_X	Average Product of X AP_X
0	0		
1	7	7	7
2	17	10	8.5
3	29	12	9.67
4	40	11	10
5	48	8	9.6
6	52	4	8.67
7	54	-2	7.71
8	52		6.5

TP, MP, and AP resulting from increases in X , while holding Y constant at 2 units. The impact is also illustrated in the graphs in Figure 6.1. These tables and the figure indicate that the total product is 7 when 1 unit of X is used, it increases to a maximum of 54 when 7 units of X are used, and it decreases to 52 units when unit 8 of the X input

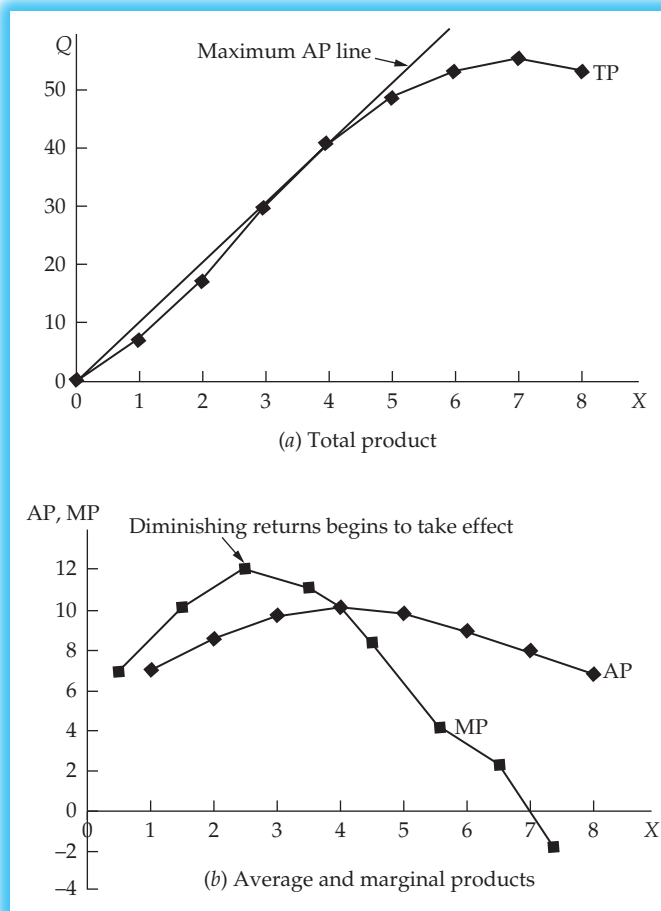


Figure 6.1 Short-Run Production with $Y = 2$

is added. Also notice in Table 6.3 that MP begins at 7 units, increases to a maximum of 12, and falls off to an ultimate value of -2 . Average product also begins at 7, increases to a maximum of 10, and then drops to 6.5 units when 8 units of X are combined with the fixed amount of Y . The pattern of these changes can be seen in Figure 6.1. The total product is plotted in Figure 6.1*a*, and the average and marginal products are plotted in Figure 6.1*b*.

We can observe that when Q , the quantity of the total product, reaches its maximum, $MP = 0$. We see also that initially (as more units of X are added to the production process), MP is greater than AP , and it then becomes less than AP . Furthermore, $MP = AP$ at AP 's highest point. The maximum AP line in Figure 6.1*a* allows you to see where this occurs on the TP curve. Because we are dealing with additional single unit changes in the input, it is difficult to see these points in Table 6.3, but they can be seen clearly in Figure 6.1. In the next section, we have more to say about the pattern of change in Q , AP , and MP and, more important, the reasons for the pattern of change.

The Law of Diminishing Returns

The key to understanding the pattern of change in Q , AP , and MP is the phenomenon known as the **law of diminishing returns**. This law states:

As additional units of a variable input are combined with a fixed input, at some point the additional output (i.e., marginal product) starts to diminish.

Diminishing returns are illustrated in both the numerical example in Table 6.3 and the graph of these same numbers in Figure 6.1. As you examine this information, think “change” as you see the word *marginal*. Therefore, the “marginal product” of an input such as labor is the change in output resulting from an additional unit of input. Notice in Table 6.3 that we have placed the marginal product *between* each interval of input and the resulting output. For example, the marginal product of the first unit of input, 7, is placed between 0 and 1 input. Continuing on, we see that marginal product reaches its maximum of 12 between the second and third units of input. It is precisely at this point, 2.5 units of input, that we can say the law of diminishing returns will begin to take effect.¹

In situations in which it would not be possible to consider intervals of less than one unit, it is necessary to approximate the point at which diminishing returns occurs. Suppose you were a manager who could actually measure and track your employees' marginal product. The first person's MP would be 7, the second one would be 10, and the third person's would be 12. It would be only on the addition of the fourth person that you would realize that the law of diminishing returns had occurred because this person's MP would be less than that of the previous worker. In summary we can say that, given the numbers in Table 6.3, the law of diminishing returns occurs in theory precisely at 2.5 units of input. In practice, a manager would have to add a fourth unit of labor in order to observe the law in action.

There are two key concerns of a practical nature that we advise readers to keep in mind when considering the impact of the law of diminishing returns in actual business situations. First, there is nothing in the law that states *when* diminishing returns will start to take effect. The law merely says that if additional units of a variable input are combined with a fixed input, *at some point*, the marginal product of the input will start to diminish. Therefore, it is reasonable to assume a manager will only discover the point of diminishing returns by experience and trial and error. Hindsight will be

¹The idea that diminishing returns will occur at this point is based on an analysis using calculus. See Appendix 6C at the end of this chapter for a complete explanation of this approach.

more valuable than foresight. Second, when economists first stated this law, they made some very restrictive assumptions about the nature of the variable inputs being used. Essentially, they assumed all inputs added to the production process were exactly the same in individual productivity. The only reason why a particular unit of an input's marginal product would be higher or lower than the others used was because of the order in which it was added to the production process.

Let us examine more closely the possible reasons for the occurrence of diminishing returns. Essentially, these reasons have to do with the physical limitations of the fixed input and the variable inputs that are added to this fixed input. Suppose we assume the numbers in Table 6.3 represent a simple manufacturing situation where the variable input is labor and the fixed input is the factory and its machinery. Suppose more labor is added to this fixed capital. Clearly when no workers are employed, TP or output is zero. The first worker produces 7 units. Thus, his MP is equal to 7 and his AP is also equal to 7. When two workers are used, their combined efforts yield a TP of 17. This implies that two people working together can produce more than the sum of their efforts working as separate individuals. (AP of one worker = 7 and AP of the two workers = 8.5.) Moreover, we see that the MP of the second worker is greater than that of the first (MP of second worker = 10 > MP of first worker = 7).

Because it is assumed in economic theory that each worker is equally productive, this must mean that the effect of teamwork and specialization enables additional workers to contribute more than those added previously to the production process, a phenomenon that we can refer to as “increasing returns.” But as still more workers are added, there are fewer and fewer opportunities for increasing returns through specialization and teamwork and at some point additional workers result in diminishing returns. Eventually, there may be so many workers relative to the fixed capacity that they may start to interfere with each other's activities. In this case, the additional workers lead to negative marginal returns, causing the total product to decrease. We see this occurring in Table 6.3 when the eighth person is added to the production process. In this case, the old adage “too many cooks spoil the soup” seems to have been realized.

Economists who first thought about the law of diminishing returns relied primarily on deductive reasoning rather than empirical verification to explain the law's existence. It was critical for them to establish this law because it then helped explain the “Three Stages of Production” and the phenomenon of increasing marginal cost (explained in detail in Chapter 7). Furthermore, when this law was established in the nineteenth century, the primary examples used to illustrate its impact on production involved agriculture, with land being the fixed factor of production, and farmers being the variable factors. At some point, it was reasoned, with a fixed amount of land, additional farm workers would result in diminishing amounts of additional harvested output.

To illustrate this law with more contemporary examples, we offer an example based on an actual situation involving a soft drink bottling facility. In addition, we provide two examples related to information technology.

The Sorting of Refillable Glass Bottles

In the early days of the soft drink industry, most drinks were packaged in returnable, refillable glass bottles. Now most drinks are packaged in plastic bottles or aluminum cans. But in some parts of the United States (particularly in Michigan) and in many parts of the rest of the world, returnable, refillable bottles are still used extensively.

There are three basic ways to sort and clean the returned bottles: (1) a totally automatic sorting system, (2) a totally manual sorting system, and (3) a hybrid system. The choice of system depends primarily on the anticipated volume of returned

bottles. A typical manual sorting area is about 30 feet long, and there is usually no room in a plant for expansion. Each sorter requires about 3 feet of space in which to work. A conveyor carrying the returned empties runs along a wall and sorters work from one side of this conveyor.

The standard productivity measure used in a plant is “cases sorted per person-hour.” If only one person is sorting bottles, he or she will not be able to keep up with the flow. The bottles tend to back up, and the system has to stop while the sorter tries to catch up.² In an interview with a bottling plant manager, the authors learned that typically two sorters per flavor are used.³ Thus, if five flavors need to be sorted, ten people work at the conveyor belt. These ten sorters take up the maximum length of the sorting area (30 feet) if the recommended amount of 3 feet per person is allocated. The plant manager explained that if more than ten sorters are used, productivity tends to decrease because the workers start to get in each other’s way. Although no quantitative detail was provided, we think that readers can well imagine that at this point, the law of diminishing returns would be starting to take effect, with each additional worker beyond ten sorting fewer returned bottles than the previously added worker.

Fixed input: Machinery and working area square footage

Variable input: People working as sorters

Development of Applications Software

Suppose you are the manager of a team of software engineers that is helping your company develop a new billing system. Suppose your project management plan calls for writing approximately 500 lines of program code every day to reach the objective. You have a team of five programmers each writing about 100 lines of code a day. To speed things up (as the deadline rapidly approaches), you decide to add more programmers to your team. You notice that the first additional person you add (the sixth member of the team) adds only 90 lines of code and the next person hired after that adds only 80 lines. Assuming the two additional programmers are equally skilled and work well as team members with the established group, you conclude that the law of diminishing returns must be at work.

Fixed input: The programming language and the hardware used to develop the applications program

Variable input: Software programmers

Response Time on a Data Network

You are a manager of a data communications network that is responsible for issuing approval codes for a credit card operation. After a credit card is swiped, the data are sent from the merchant’s bank to the bank that issued the cardholder the card via the data communications network. The approval code comes back to the merchant within an average of 3 seconds. The merchants complain that this is too long. Software programmers are put to work to make the adjustments in software that will reduce the average to 2 seconds. Ten programmers working 10 hours a week for 2 weeks reduce the response time down to 2 seconds. Merchants continue to complain and want the average reduced further to 1 second. As the network manager, you try again with another ten programmers working the same amount of time as the previous addition of

²Readers can imagine that this is very much what might happen in Stage I of the production process. (See next section for stages of production.)

³Interview with a former student of one of the authors, who used to work as a bottling plant manager for Pepsi-Cola.

programmers. You find that the same additional time and effort does not yield the desired 1-second reduction in response time. You conclude that the law of diminishing returns is at work.

Fixed input: The technological infrastructure (i.e., the hardware and bandwidth of the transmission facilities of the physical network and the limits and capabilities of the software)

Variable input: Software programmers

In summary, these examples illustrate how diminishing returns can be caused by the physical limitations of fixed capital such as machinery and work space, the technological limitations of computer hardware and software, and the personnel and management problems caused by increasing numbers of people working with a fixed capacity.

The Three Stages of Production in the Short Run

In competitive markets, the short-run production function can be divided into three distinct **stages of production**. To illustrate this phenomenon, let us return to the data in Table 6.3 and Figure 6.1. For your convenience, Figure 6.1 has been reproduced as Figure 6.2. As the figure indicates, Stage I runs from zero to four units of the variable

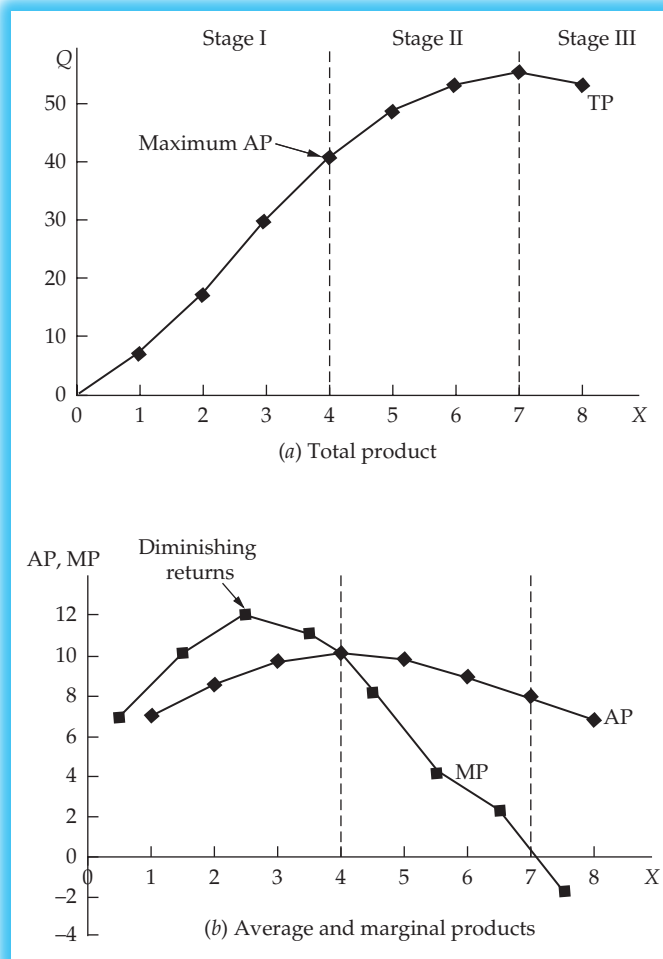
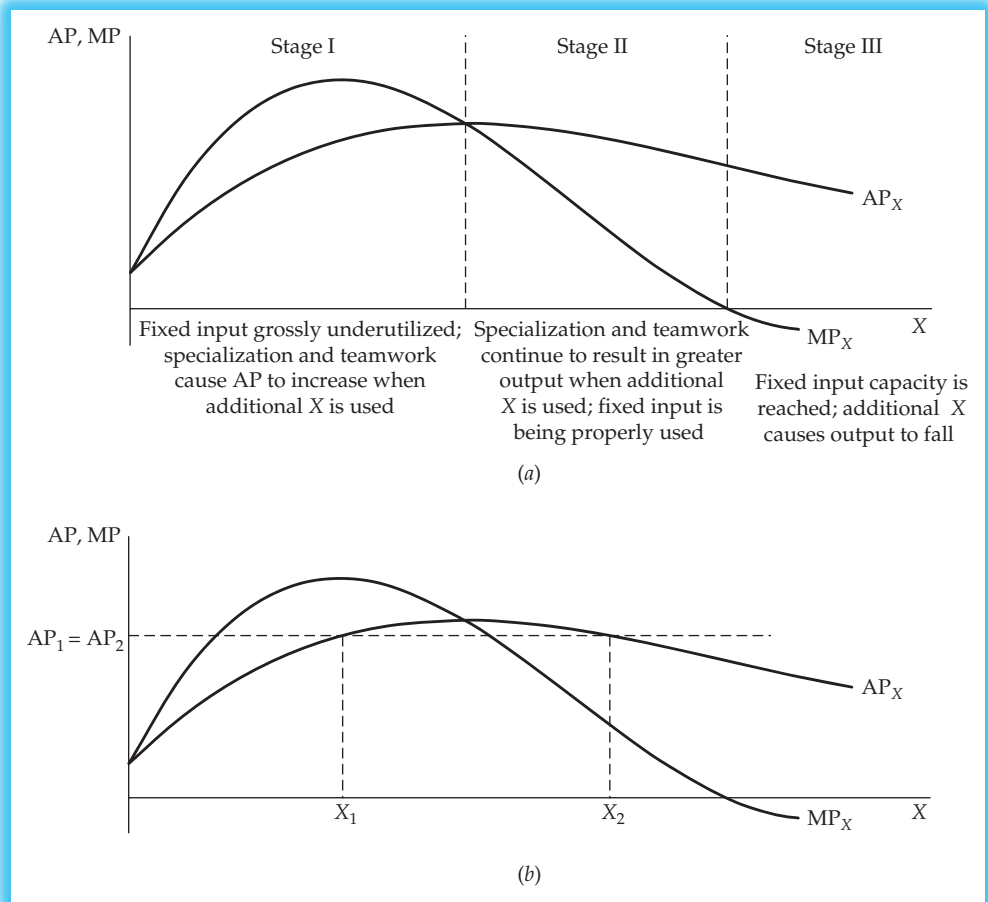


Figure 6.2 The Three Stages of Production

input X (i.e., to the point at which average product reaches its maximum). Stage II begins from this point and proceeds to seven units of input X (i.e., to the point at which total product is maximized). Stage III continues on from that point. According to economic theory, in the short run, “rational” firms should only be operating in Stage II. It is clear why Stage III is irrational: The firm would be using more of its variable input to produce less output! However, it may not be as apparent why Stage I is also considered irrational. The reason is that if a firm were operating in Stage I, it would be grossly underusing its fixed capacity. That is, it would have so much fixed capacity relative to its usage of variable inputs that it could increase the output per unit of variable input (i.e., average product) simply by adding more variable inputs to this capacity. Figure 6.3*a* summarizes the three stages of production and the reasons that the rational firm operates in Stage II of the short-run production function.

If you are still not clear about the irrational nature of Stage I, there is an alternative explanation. In Figure 6.3*b*, we have designated two levels of variable input usage: X_1 and X_2 . Here we see that the average product is the same whether X_1 or X_2 units of the variable input are used. If output per variable input is the same regardless of which input level is used, the firm should employ X_2 because the total product will be higher.

Figure 6.3 Explanation of Production Stages





Module 6A
Module 6B
Module 6C

Derived Demand and the Optimal Level of Variable Input Usage

Given that a firm’s short-run production has only one “rational” stage of production (i.e., Stage II), we must still determine the level of input usage within Stage II that is best for the rational, profit-maximizing firm. To demonstrate how we determine the optimal level of input usage within Stage II, we have created another numerical illustration, shown in Table 6.4. In this table, we observe that Stage II occurs from 60,000 to at least 80,000 units of total product. In terms of labor input, this Stage II ranges from four to at least eight units of labor. But which level of output (or input) usage should we consider to be optimal? In answering this question, note that this table resembles the setup of a spreadsheet program, because MP is placed at the *same* level of output that results from an additional unit of input rather than *between* the two levels of output as shown in Table 6.3. In fact, we have provided you with an Excel exercise so you can experiment with your own variations of the example shown in Table 6.4.

The answer to this question is based on a concept introduced in Chapter 4, derived demand. Recall that the demand for inputs is derived from the demand for their output. In other words, it would do no good for the firm to decide how many units of variable input to use without knowing how many units of output it could sell, the price of the product, and the monetary costs of employing various amounts of the X input.

To see exactly how this works, assume a firm is operating in a perfectly competitive market for its input and its output. That is, it can sell as many units of the product as it wants as long as it does so at the going market price given the assumption of competitive product markets, $MR = P$. Moreover, it can hire as many units of labor as

Table 6.4 Combining Marginal Revenue Product (MRP) with Marginal Labor Cost (MLC)

Labor Unit (X)	Total Product (Q or TP)	Average Product (AP)	Marginal Product (MP)	Total Revenue Product (TRP)	Marginal Revenue Product (MRP)	Total Labor Cost (TLC)	Marginal Labor Cost (MLC)	TRP – TLC	MRP – MLC
0	0		0	0		0		0	0
1	10,000	10,000	10,000	\$ 20,000	\$20,000	\$10,000	\$10,000	\$10,000	\$10,000
2	25,000	12,500	15,000	50,000	30,000	20,000	10,000	30,000	20,000
3	45,000	15,000	20,000	90,000	40,000	30,000	10,000	60,000	30,000
4	60,000	15,000	15,000	120,000	30,000	40,000	10,000	80,000	20,000
5	70,000	14,000	10,000	140,000	20,000	50,000	10,000	90,000	10,000
6	75,000	12,500	5,000	150,000	10,000	60,000	10,000	90,000	0
7	78,000	11,143	3,000	156,000	6,000	70,000	10,000	86,000	–4,000
8	80,000	10,000	2,000	160,000	4,000	80,000	10,000	80,000	–6,000

Note: Although not explicitly noted in each acronym, the univariate product functions (TP, AP, MP, TRP, and MRP) are functions of the variable factor, labor. Labor is specifically noted in MLC to distinguish it from marginal cost of production.

$P = \text{Product price} = \2

$W = \text{Cost per unit of labor} = \$10,000$

$MRP = MP \times P$

$TLC = X \times W$

$MLC = \Delta TLC / \Delta X$

it desires as long as it pays these inputs the going market price (i.e., the competitive wage rate). Note that four new measures are added to Table 6.4:

Total revenue product (TRP): The market value of the firm's output, computed by multiplying the total product by the market price ($Q \times P$).

Marginal revenue product (MRP): The change in the firm's total revenue product resulting from a unit change in variable input usage ($\Delta TRP / \Delta X$). It can also be computed by multiplying the marginal product of labor by the product price ($MP \times P$).

Total labor cost (TLC): The total cost of using the variable input, labor, computed by multiplying the wage rate (which we assume to be some given and constant dollar amount) by the number of variable inputs employed ($\text{Wage rate} \times X$).

Marginal labor cost (MLC): The change in total labor cost resulting from a unit change in the number of variable inputs used. Because the wage rate is assumed to be constant regardless of the number of inputs used, the MLC is the same as the wage rate.⁴

In deriving the figures for these measures in Table 6.4, we have assumed a product price of \$2 and a wage rate of \$10,000 per unit. Given these figures, you can see that a rational firm would want to hire 6 units of labor. Up to that point it pays for the firm to add more labor because the additional **marginal labor cost (MLC)** to the firm to do so is more than made up for by the additional **marginal revenue product (MRP)** brought in by the sale of the increased output. Beyond that point the firm would pay more in additional labor cost than it would receive in additional revenue.

Can you discern how the demand for the output is incorporated into this analysis, that is, how the demand for the input X is actually derived from the demand for the output? Suppose the market demand increased and drove the market price up to \$4. This would increase the market value of the labor input's efforts. In other words, the market value of each additional labor unit's contribution to the total product would double. This increase in the labor input's MRP would then justify the firm's use of a seventh unit of labor. Thus, an increase in the market demand for the output leads to an increase in the demand by the firm for labor input, all other factors held constant. The original figures from Table 6.4, as well as the case in which market price is assumed to increase to \$4, are illustrated in Figure 6.4.

We can summarize this relationship between the demand for the output and the demand for the input in terms of the following optimal decision rule:

A profit-maximizing firm operating in perfectly competitive output and input markets will be using the optimal amount of an input at the point at which the monetary value of the input's marginal product is equal to the additional cost of using that input—in other words, when **MRP = MFC**. In this instance, the profit-maximizing firm's optimal labor usage occurs where the marginal revenue product of labor equals marginal labor cost, **MRP = MLC**.

The Case of Multiple Inputs (Abridged Version)

This section presents a relatively nonquantitative explanation of optimal input decisions involving two or more inputs. For a complete treatment of multiple inputs using numerical tables, graphs, and algebraic equations, refer to Appendix 6B at the end of this chapter.

⁴This term is more generally referred to as *marginal resource cost (MRC)* or *marginal factor cost (MFC)*. In the present discussion, the variable factor is labor and hence marginal factor cost is marginal labor cost. By assuming the firm hires workers in a perfectly competitive labor market, we also assume it can hire as many workers as it chooses at some going market wage rate determined by the supply and demand for this particular type of worker. This would not be the case if the firm were operating in an imperfectly competitive labor market. We do not consider the imperfectly competitive case in this text. Interested readers should consult any intermediate microeconomics text for a full discussion on this type of labor market and its impact on the input decisions of the firm.

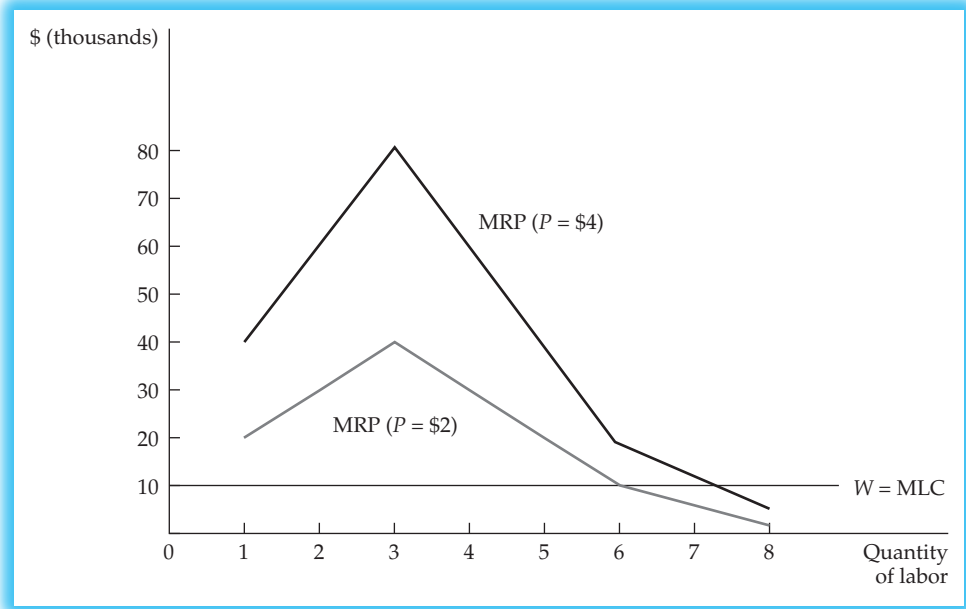


Figure 6.4 The Effect of Increased Market Price on the Demand for Labor

In the example illustrated in Table 6.4, we show how a firm could use the concept of equalizing at the margin to determine the optimal usage level of a single variable input. This same equalizing concept applies to a situation in which two or more inputs are being considered. In the multiple input case, we must consider the relationship between the *ratio* of the marginal product of one input and its cost to the *ratio* of the marginal product of the other input (inputs) and its (their) cost. Expressed mathematically for “*n*” inputs:

$$\frac{MP_1}{w_1} = \frac{MP_2}{w_2} = \frac{MP_n}{w_n}$$

Where w_i is the price of the i^{th} factor of production. It is important to note that i need not refer solely to labor inputs; therefore, economists often write this as $MP_i/P_i = MP_j/P_j$ across factors i and j . Economists describe this as having *equal bang for the buck across factors of production* because $1/P_i$ units of factor i can be purchased for a dollar.

Let us consider the case of a global manufacturing company that has manufacturing facilities in two different countries: a high-wage country and a low-wage country. We can analyze this problem by treating labor in the high-wage country as one input and labor in the other country as another input. At first glance, it might seem obvious that a country seeking to minimize its costs and hence maximize its profits should try to manufacture as much as possible in the low-wage country. This would mean using mostly the low-wage labor input. But production theory suggests that the firm should not only look at input costs, but also the marginal products of each input *relative* to their respective costs.

Suppose you are the production manager of a company that makes computer parts and peripherals in Malaysia and Costa Rica. At the current levels of production and input utilization in the two countries, you find that:

- Marginal product of labor in Malaysia (MP_{Mal}) = 18
- Marginal product of labor in Costa Rica (MP_{CR}) = 6
- Wage rate in Malaysia (w_{Mal}) = \$6/hr.
- Wage rate in Costa Rica (w_{CR}) = \$3/hr.

How much would you produce in each manufacturing facility? Because labor is cheaper in Costa Rica, you might be tempted to produce most of your output in that country. However, a closer look at the MP:wage ratios reveals the opposite conclusion. That is,

$$\frac{MP_{\text{Mal}}}{w_{\text{Mal}}} > \frac{MP_{\text{CR}}}{w_{\text{CR}}} \text{ or } \frac{18}{\$6} > \frac{6}{\$3}$$

This means that at the margin, the last dollar spent on a unit of labor in Costa Rica would yield 2 units of output (6/\$3), while in Malaysia the last dollar spent would result in 3 additional units of output (18/\$6). This inequality implies that the firm should begin to shift more of its production away from Costa Rica to Malaysia, until the two ratios are equalized. In theory, this optimal or equilibrium point would occur because, as more labor is used in Malaysia, the law of diminishing returns would start to reduce the MP of labor in this country. With less labor being used in Costa Rica, the law of diminishing returns would work in reverse, causing the marginal product of labor in this country to rise. Of course, this assumes a short-run condition whereby the complementary fixed inputs used along with labor remain constant.⁵

Once the implication of the basic model is understood, other factors can be brought in. If these factors outweigh the MP/input cost criteria, a company may well modify its decision. For example, despite Malaysia's higher MP/wage ratio, there may be political and economic risk factors to consider. (This was indeed the case when the Malaysian government imposed foreign exchange controls in 1998 by requiring foreign investors to keep their profits in Malaysia for at least 1 year before they could be repatriated.) In contrast, Costa Rica is a fairly stable economy with leaders who do not seem to want to impose any such trade and finance restrictions. Its proximity to U.S. markets would also reduce transportation costs. Furthermore, it has a democratic government (with no standing army) and a highly skilled workforce with strong English language skills.

THE LONG-RUN PRODUCTION FUNCTION

In the long run, a firm has time enough to change the amount of all its inputs. Thus, there is really no difference between fixed and variable inputs. Table 6.5 uses the data first presented in Table 6.1 and illustrates what happens to total output as both inputs X and Y increase 1 unit at a time. The resulting increase in the total output as the two inputs increase is called **returns to scale**.

Looking more closely at Table 6.5, we see for example that if the firm uses 1 unit of X and 1 unit of Y , it will produce 4 units of output. If it doubles its inputs (i.e., 2 units of X and 2 units of Y), it will produce 17 units of output. Thus, a doubling of inputs has produced more than a fourfold increase in output. Proceeding further, we notice that an additional doubling of inputs (i.e., 4 units of X and 4 units of Y) results in more than a threefold increase in output, from 17 to 60. What we are observing in this table is *increasing returns to scale*. The same results hold even when $L \neq K$ as seen in the boxed expansion path where $L = 2K$.

⁵Over time, there might be another factor causing the two ratios to begin to equalize. If there are enough companies that see this inequality, their combined increase in demand for labor in Malaysia may start to drive up the wage rate in this country, while a decrease in demand for labor in Costa Rica may slow the rate in increase in wage rates or even cause them to fall. Thus equalizing at the margin could take place because of changes in the values of the denominators of the ratios as well as the value of their numerators.

Table 6.5 Returns to Scale Along Two Expansion Paths

Units of Y Employed	Output Quantity (Q)							
	1	2	3	4	5	6	7	8
8	37	60	83	96	107	117	127	128
7	42	64	78	90	101	110	119	120
6	37	52	64	73	82	90	97	104
5	31	47	58	67	75	82	89	95
4	24	39	52	60	67	73	79	85
3	17	29	41	52	58	64	69	73
2	7	17	29	40	48	52	54	52
1	4	8	14	20	27	24	21	17
	1	2	3	4	5	6	7	8
	Units of X Employed							

According to economic theory, if an increase in a firm’s inputs by some proportion results in an increase in output by a greater proportion, the firm experiences *increasing returns to scale*. If output increases by the same proportion as the inputs increase, the firm experiences *constant returns to scale*. A less than proportional increase in output is called *decreasing returns to scale*.

You might simply assume firms generally experience constant returns to scale. For example, if a firm has a factory of a particular size, then doubling its size along with a doubling of workers and machinery should lead to a doubling of output. Why should it result in a greater than proportional or, for that matter, a smaller than proportional increase? For one thing, a larger scale of production might enable a firm to divide tasks into more specialized activities, thereby increasing labor productivity. Also, a larger scale of operation might enable a company to justify the purchase of more sophisticated (hence, more productive) machinery. These factors help explain why a firm can experience increasing returns to scale. In contrast, operating on a larger scale might create certain managerial inefficiencies (e.g., communications problems, bureaucratic red tape) and hence cause decreasing returns to scale. More will be said about the factors that can cause increasing or decreasing returns to scale in Chapter 7, when we discuss how returns to individual factors as well as returns to scale affect the shape of the various cost curves.

One way to measure returns to scale is to use a coefficient of output elasticity, E_Q :

$$E_Q = \frac{\text{Percentage change in } Q}{\text{Percentage change in all inputs}}$$

Thus,

If $E_Q > 1$, we have increasing returns to scale (IRTS).

If $E_Q = 1$, we have constant returns to scale (CRTS).

If $E_Q < 1$, we have decreasing returns to scale (DRTS).

Another way of looking at the concept of returns to scale is based on an equation that was first presented at the outset of this chapter:

$$Q = f(X, Y)$$

Recall in the original specification of this equation that it may include as many input variables as necessary to describe the production process (i.e., n variables). For ease of

discussion, we limit this number to two: X and Y . Now suppose we increase the amount of each input by some proportion $k > 1$. For example, if we increase the inputs by 10 percent, $k = 1.10$. If we double the inputs, $k = 2.0$. Of course, Q is expected to increase by some proportion as a result of the increase in the inputs. Let h represent the magnitude of this increase. Expressed in terms of Equation (6.3),

$$hQ = f(kX, kY)$$

Using this notation, we can summarize returns to scale in the following way:

If $h > k$, the firm experiences increasing returns to scale ($E_Q > 1$).

If $h = k$, the firm experiences constant returns to scale ($E_Q = 1$).

If $h < k$, the firm experiences decreasing returns to scale ($E_Q < 1$).

We illustrate returns to scale with a numerical example. Suppose we have the following production function of perfect substitutes:

$$Q_1 = 5X + 7Y$$

If we use 8 units of X and 12 units of Y , the output will be

$$\begin{aligned} Q_1 &= 5(8) + 7(12) \\ &= 40 + 84 = 124 \text{ units} \end{aligned}$$

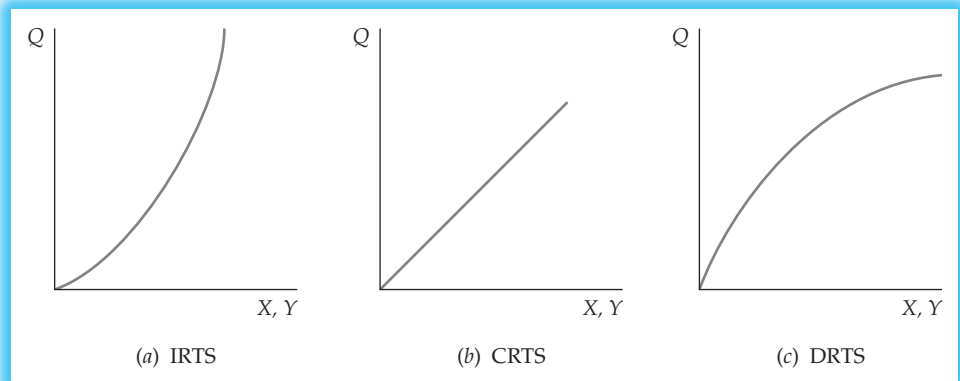
Now let us increase each input by 25 percent (i.e., $k = 1.25$). This will give us

$$\begin{aligned} Q_2 &= 5(10) + 7(15) \\ &= 50 + 105 = 155 \end{aligned}$$

The 25 percent increase in X and Y has led to a proportional increase in output (i.e., 155 is 25 percent more than 124).

We can also illustrate the concept of returns to scale graphically. Figure 6.5 shows the three possible types of returns to scale. In each case, we assume the inputs (X and Y) are increased by the same proportion; thus, they are both included on the horizontal axis. Obviously, these graphs are idealized representations of returns to scale. In reality, we would not expect the changes in output relative to the changes in inputs to behave in such a smooth and orderly fashion.

Figure 6.5 Graphic Representations of Returns to Scale



THE ESTIMATION OF PRODUCTION FUNCTIONS

Following our examination of demand theory, we turned to the subject of demand estimation. Now it is time to examine another of a managerial economist's important tasks, the estimation of production functions. This section deals with three major topics. First, we discuss the possible shapes of production functions. Second, we discuss the Cobb-Douglas production function, a form that has been commonly used by economists since its introduction in the 1920s. Last, we look at the data needed for estimation and some of the production function studies published by economists.

The Various Forms of a Production Function

Earlier in this chapter, the short-run production function was introduced. As you remember, the short run is characterized by the existence of a fixed factor to which we add a variable factor. Thus, the simple function, containing just one variable factor and one fixed factor, can be written as follows:

$$Q = f(L)_K$$

where output Q is determined by the quantity of the variable factor L (labor) with the fixed factor K (capital) given.⁶

The theoretical part of this chapter assumed the production function starts with increasing marginal returns followed by decreasing marginal returns. In other words, all three stages of production are present. This situation is represented by a cubic function:

$$Q = a + bL + cL^2 - dL^3$$

where a is the constant and b , c , and d are coefficients. Figure 6.6*a* shows the total product curve,⁷ and Figure 6.6*b* shows the average and marginal product curves. All three stages of production are present.

It is possible, however, that the data employed in the estimate will exhibit diminishing marginal returns but no Stage I. Such an estimate is represented by the quadratic function:

$$Q = a + bL - cL^2$$

Figures 6.7*a* and 6.7*b* depict total and per-unit product curves, respectively. Both graphs show diminishing marginal returns for all levels of L and thus Stage I is absent.⁸

In performing empirical research, it may be possible to identify a linear production function, $Q = a + bL$. This function exhibits no diminishing returns; the total product will be a straight line with slope b , and both the MP and AP lines will be horizontal and equal to b .⁹ Of course, a straight-line production function may hold in some

⁶In the prior discussion, the letters X and Y designated inputs. This emphasized the production function's generality. Here we use L (for labor) and K (for capital) because these are the letters that are frequently employed to indicate variable and fixed factors. In long-run analysis, of course, both L and K are variable.

⁷Total product curves should, in theory, be drawn from the origin because it is most likely that no production would occur in the absence of any variable factor. However, when production functions are estimated using regression analysis, it is probable that the fitted line will have a positive or negative y -intercept. Thus, in the equations that follow we include the intercept a . It should be noted that this value probably has no economic significance. Economists also force the intercept to 0 in this instance in order to remove a from the short-run production function. The resulting average and marginal product functions are $AP(L) = b + cL - dL^2$ and $MP(L) = b + 2cL - 3dL^2$. See Appendix 6C for an extended discussion of cubic production.

⁸If the intercept is present, average product is equal to $a/L + b - cL$. However, if the intercept is omitted, the average product will be a straight line, as shown in Figure 6.7*b*. In this event, $AP(L) = b - cL$ and $MP(L) = b - 2cL$.

⁹Again, if the intercept is present, the average product will show a decline. If, however, the intercept is omitted, the average product will be horizontal and equal to marginal product, b .

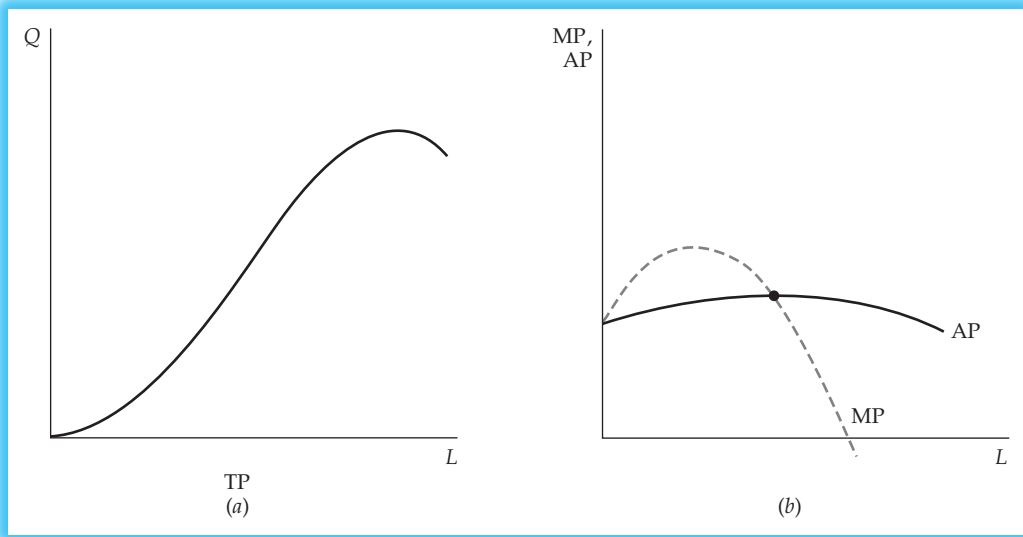


Figure 6.6 Cubic Production Function

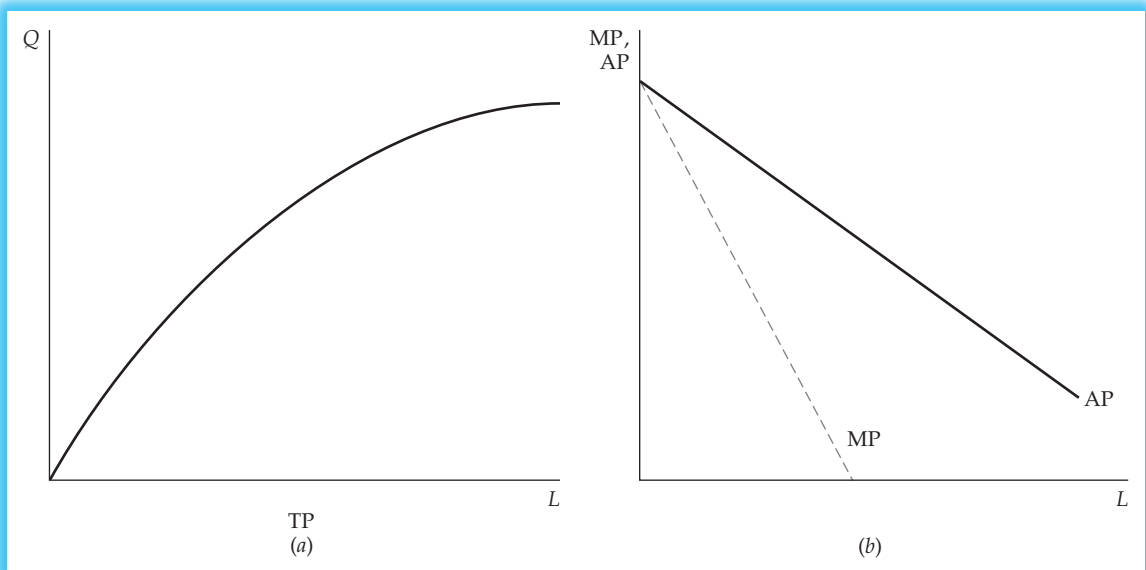
real situations, but given the existence of a fixed factor, constant marginal product should not be expected to prevail over a wide range of quantities produced.

Another form of the production function is the **power function**, which takes the following form:

$$Q = aL^b$$

The shape of this production function depends on the exponent b , as is shown in Figure 6.8.

Figure 6.7 Quadratic Production Function



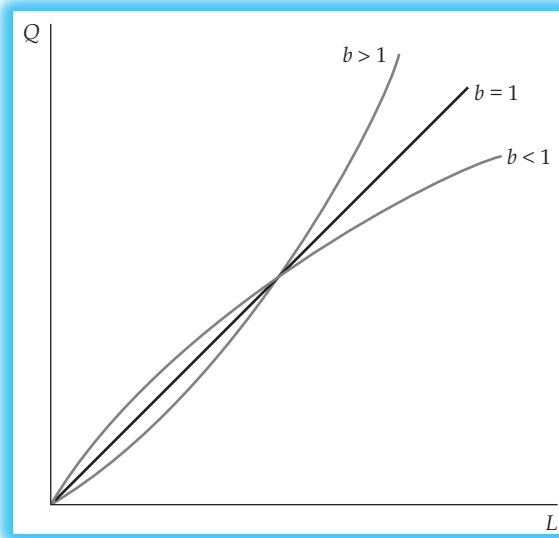


Figure 6.8 Power Production Function

A major advantage of the power function is that it can be transformed into a linear function when it is expressed in logarithmic terms, making it amenable to linear regression analysis:

$$\log Q = \log a + b \log L$$

The direction of marginal product depends on the size of the exponent b . If $b > 1$, marginal product is increasing; if $b = 1$, it is constant; and if $b < 1$, it is decreasing. It cannot, however, exhibit two directions for marginal product on the same function, as was possible with the cubic equation form. This function is used frequently in empirical work. One reason for its popularity is that it can be readily transformed into a function with two or more independent variables:

$$Q = aX_1^b \cdot X_2^c \cdot \dots \cdot X_n^m$$

Using more than one independent variable in a production function is certainly more realistic than limiting the analysis to only one, and when it is assumed all inputs are variable, we have then moved from short-run analysis to the long run. Indeed, this function can be employed in both analyses. In a simple two-variable model (e.g., labor and capital), the power function permits the estimation of marginal product (e.g., when labor changes and capital remains the same) and of returns to scale (when both variables change).

The Cobb-Douglas Production Function

The **Cobb-Douglas production function** was introduced in 1928,¹⁰ and it is still a common functional form in economic studies today. It has been used extensively to estimate both individual firm and aggregate production functions. It has undergone significant criticism but has endured. “It is now customary practice in economics to

¹⁰C. W. Cobb and P. H. Douglas, “A Theory of Production,” *American Economic Review*, 8, 1 (March 1928, suppl.), pp. 139–65. Charles W. Cobb was a mathematician; Paul H. Douglas was an economist, and later a U.S. senator from Illinois.

deny its validity and then to use it as an excellent approximation.”¹¹ It was originally constructed for all the manufacturing output (Q) in the United States for the years 1899 to 1922. The two inputs used by the authors were number of manual workers (L) and fixed capital (K). The formula for the production function, which was suggested by Cobb, was of the following form:

$$Q = aL^bK^{1-b}$$

What are the important properties of this function that have kept it popular for more than 80 years?

1. To make this equation useful, both inputs must exist for Q to be a positive number. This makes sense because total product is a result of combining two or more factors.¹²
2. The function can exhibit increasing, decreasing, or constant returns. Originally, Cobb and Douglas assumed returns to scale are constant. The function was constructed in such a way that the exponents summed to $b + 1 - b = 1$. (A brief description of the mathematics of the Cobb-Douglas function can be found in Appendix 6C.) However, in later studies, they (as well as other researchers) relaxed this requirement and rewrote the equation as follows:

$$Q = aL^bK^c$$

Under this construction, if $b + c > 1$, returns to scale are increasing and if $b + c < 1$, returns are decreasing.¹³ Constant returns exist when $b + c = 1$.

3. The function permits us to investigate the marginal product for any factor while holding all others constant. Thus it is also useful in the analysis of short-run production functions. In the two-factor model discussed here, the marginal product of labor turns out to be $MP_L = bQ/L$, and the marginal product of capital is $MP_K = cQ/K$. (The mathematics of this result are shown in Appendix 6C.) Each coefficient will usually be less than 1, and this will mean that they each exhibit diminishing marginal returns. Thus, production takes place in Stage II, which is the relevant area of production.
4. The elasticity of production is an important concept as discussed in Appendix 6C. In the Cobb-Douglas function, the elasticities of the factors are equal to their exponents, in this case b and c . Thus, the elasticities of labor and capital are constants.
5. Because a power function can be converted to a linear function by using logarithms, it can be estimated by linear regression analysis, which makes for a relatively easy calculation with any software package

$$\log Q = \log a + b \log L + c \log K$$

6. Although we limit our discussion to just two input variables (L and K), Cobb-Douglas can accommodate any number of independent variables, as follows:

$$Q = aX_1^b \cdot X_2^c \cdot X_3^d \cdot \dots \cdot X_n^m$$

7. A theoretical production function assumes technology is constant. However, the data fitted by the researcher may span a period over which technology has progressed. One of the independent variables in the previous equation could represent technological change (a time series) and thus adjust the function to take technology into consideration.
8. Given a Cobb-Douglas production function, spending on each variable factor is cost-minimizing if the proportion of variable cost devoted to that factor is proportional to the elasticity of that factor. For example, given a three factor Cobb-Douglas production function, $Q = aL^b M^c K^d$, then the fraction $b/(b + c)$ of variable cost should be spent on

¹¹George J. Stigler, *The Theory of Price*, 4th ed., New York: Macmillan, 1987, p. 153.

¹²If an additive function (e.g., $Q = a + bL + cK$) were employed, output would still be positive even if one of the two inputs were zero.

¹³Of course, if the function contained more than two factors, the exponents would still have to equal more than one for increasing returns, less than one for decreasing, and one for constant.

labor and the fraction $c/(b + c)$ of variable cost should be spent on materials in the short run in order to be cost-minimizing. In the long run, the fraction $b/(b + c + d)$ of total cost should be spent on labor, $c/(b + c + d)$ on materials and $d/(b + c + d)$ should be spent on capital. These spending patterns equalize MP_i/P_i across variable factors i .

The Cobb-Douglas function has the following shortcomings:

1. The Cobb-Douglas cannot show the marginal product going through all three stages of production in one specification. (A cubic function would be necessary to achieve this.)
2. Similarly, it cannot show a firm or industry passing through increasing, constant, and decreasing returns to scale.
3. There are also important problems with specification of data to be used in empirical estimates. These problems are discussed next.

Statistical Estimation of Production Functions

Let us turn now to a description of the data that would be used in empirical estimates of production functions. Later, we also review some published studies. When the statistical estimate concerns a plant, or group of plants, the data will come from company records—accounting, employment, purchasing, manufacturing, and others.

If only one product is produced in a plant, Q is specified in physical units (e.g., number, tons, gallons). However, if a plant produces a number of different products, and it is not possible to segregate properly the inputs and outputs of the products, estimation becomes considerably more difficult. In such a case, the investigator probably must settle for some measure of value, assigning weights to products depending on the value (in terms of cost or selling price) produced. There are some obvious problems with this procedure. First, over time the data will have to be deflated to account for price or cost changes. Second, the price or cost of a product may not be an exact reflection of the inputs combined in the total value. However, until better measurement methods are found, such valuing methods will have to suffice.

Measuring inputs also can vary in the level of difficulty. Inputs should be measured as “flow” rather than “stock” variables, and this is not always possible.¹⁴

Usually, the most important input is labor. Hours of labor input are probably the best measure for our purposes. Data for direct labor hours are ordinarily available from company records. If they are not, then number of direct workers is the next best choice. However, it must be remembered that number of workers is a stock variable and does not necessarily represent the amount of labor expended in production. For materials, a physical measure is again best. In some cases, such data are readily available (weight of materials consumed, for instance). Of course, because we do not want to proliferate the number of independent variables, choosing only the most important raw material may be indicated. Alternatively, a combination (by weight or value) may prove to be a viable option. Utilities (electricity, gas, etc.) may also be included; in this case, physical quantities should be fairly easily obtained.

The most difficult input variable is the all-important capital input. How can one measure the use of plant, machinery, and equipment in production? Because different components of plant and equipment are of varying durability and different input intensity, usage per period is hard to establish. In some cases, periodic depreciation may be an indicator of capital use. However, depreciation as recorded in the company’s books is often based on accounting convention or legal requirement. Further, the projected depreciation life of a piece of equipment tends to depend on tax rules

¹⁴Flows measure the usage of services consumed in producing a product, whereas stocks represent the amounts of factors present and available for productive use.

because the firm wants to take advantage of the fastest write-off permitted for tax purposes. Some capital items, such as land, are not depreciated. Unless some rather consistent measurement of capital usage can be designed by the researcher (and this is certainly a formidable task), the common method of measuring capital by the use of a stock variable (e.g., fixed assets) is probably indicated. Obviously, this is not a perfect solution because the price of these assets depends on when the assets were acquired. Thus, the asset figure must be adjusted by a price index. Should gross fixed assets (i.e., the original cost of all plant and equipment) be used, or net fixed assets (gross assets minus accumulated depreciation)? Again, this is a difficult question. There is no specific answer. The method used must be determined by the investigator as that which is most reasonable (and available) for each specific case.

If the production function estimate is to be accomplished using regression analysis, then we must choose between time series and cross-sectional analysis. The former would be preferable if data have been collected over a period of time in a given plant. However, it would be necessary to make adjustments for inflation if the variables are in monetary (rather than physical) terms. Also, a time variable (or a dummy variable) may have to be employed to account for changes in technology.

Cross-sectional analysis is favored when the data collected cover a number of plants in a given time period. But here again, problems may arise. The various plants may not employ the same level of technology. If the data are in monetary terms, an adjustment for differential price or wage levels at different geographic locations would be necessary.

Although a theoretical production function assumes output is produced at the most efficient input combinations, in reality such an ideal situation is certainly not assured whichever estimating method is used. Ultimately, there is no perfect way to measure and analyze the data. The researcher must choose the most appropriate method.

A Numerical Example of a Cobb-Douglas Production Function

A cross-sectional sample of twenty soft drink bottling plants has been selected. The data are given for a specific month in 1998. Only two independent variables are used: (1) number of direct workers and (2) plant size. Plants range from 1 to 1.75, based on a size and capacity measure developed by engineers.

Production, the dependent variable, is stated in terms of gallons of product shipped during the period. Table 6.6 contains the data and the results of the analysis.

The Cobb-Douglas function of the form $Q = aL^b K^c$ was applied to the numbers. The regression output in the table shows the results.¹⁵ The regression equation is as follows:

$$Q = 15.14L^{0.664} K^{0.321}$$

R^2 (the coefficient of determination) is quite high, showing that 98 percent of the variation in Q is explained by the model. The two coefficients are significantly different from zero because they both pass the t -test.¹⁶ The sum of the two coefficients is 0.985. Because this result is so close to 1, it can be assumed the plants exhibit constant returns to scale. Each input indicates decreasing marginal returns.



¹⁵The results were calculated using the Excel regression analysis program. The raw data were transformed into logarithms, and a straight-line regression in logarithms was computed. The constant of 1.18 is a logarithm, and 15.14 is its antilog. The Cobb Douglas Production (Regression) Excel App provides instructions on how to estimate a constant returns to scale Cobb-Douglas production function using this data. The resulting function, $Q = 15.55L^{0.655} K^{0.345}$, is quite similar to the one produced by allowing returns to vary.

¹⁶Labor's t -test is significant at the 1 percent level, and capital's is significant at the 5 percent level.

Table 6.6 Production Function: Soft Drink Bottling

Total Product	Labor	Capital
97	15	1.00
98	17	1.00
104	20	1.00
120	22	1.00
136	22	1.25
129	25	1.25
145	30	1.25
170	32	1.25
181	35	1.25
166	30	1.50
175	35	1.50
190	38	1.50
212	42	1.50
220	44	1.50
207	45	1.50
228	44	1.75
226	47	1.75
240	52	1.75
270	55	1.75
280	58	1.75

Summary Output		
Regression Statistics		
R Square		0.980965846
Adjusted R Square		0.978726534
Standard Error		0.020818585
Observations		20

	Coefficients	Standard Error	t Stat
Intercept	1.1800154	0.096022924	12.288892
LogL	0.6643702	0.075371367	8.8146228
LogK	0.3214714	0.147006777	2.1867796

Three Studies of Individual Production Functions

In 1967, a study of the production function of the Pacific halibut industry was published. A Cobb-Douglas function appeared to give good results and showed that constant returns are probably the rule in this industry. A “good captain” variable was added—a confidential rating of the management abilities of the captains of the thirty-two boats in the sample. It was found that good captains made a difference.¹⁷

Another study included management in the production function equation. This study dealt with a sample of plants of a multinational consumer goods manufacturer. Time series and cross-sectional data were combined to obtain 127 observations over 8 years (1975–1982). Management was measured as a performance ranking of each plant in terms of three criteria—output goal attainment, cost over- or underfulfillment, and quality level of output. The results showed that the management variable was statistically

¹⁷Salvatore Comitimi and David S. Huang, “A Study of Production and Factor Shares in the Halibut Fishing Industry,” *Journal of Political Economy*, August 1967, pp. 366–72.

significant. Another feature of this study was the conclusion that increasing returns existed up to a certain plant size and that decreasing returns resulted at larger sizes.¹⁸

A 2005 study used the Cobb-Douglas function as a base and then added several management policies to determine whether these policies contributed to increases in productivity. The study examined three relatively small companies over a period of four to five years. The three companies were in a sophisticated printing business, the manufacture of air filtration systems, and furniture production, respectively. The results suggest “that a substantial portion of productivity improvements are indeed under the control of management.” Of the six management policies examined, two had a positive and mostly statistically significant effect on productivity. These two were:

1. Dollars spent for asset maintenance and technology improvements
2. Benefits derived from job specialization and reorganization¹⁹

There are numerous statistical analyses of industries in other countries based on the Cobb-Douglas production. For example, there is a study that examined the relationship between human capital and labor productivity in the food industry of Iran from 1995 to 2006. This study looked at the effect that a change in the ratio of educated and skilled workers to total workers had on labor productivity. The results showed that a 1 percent increase in the ratio of educated workers to total workers increased labor productivity by 0.14 percent and a 1 percent increase in the ratio of skilled workers to total workers increased labor productivity by 0.42 percent.²⁰

Aggregate Production Functions

A large proportion of the studies performed using the Cobb-Douglas function did not deal with data for individual firms, but rather with aggregations of industries or even the economy as a whole. Although much of this work has proved to be quite fruitful in describing production functions, the interpretation of the results may not be quite as meaningful as for individual production functions. When data for the economy as a whole are used, the model must accommodate different technologies and different processes, and thus does not represent a specific technological process of a given firm. When the aggregation is done at the level of an industry rather than the overall economy, the assumption of similar technology is more appropriate, but even in such a case many dissimilarities may occur.

Gathering data for such aggregate functions can be difficult. For the economy as a whole, gross national or domestic product—in real terms—could be used to measure output. For specific industries, data from the Census of Manufactures or the production index published by the Federal Reserve Board can be employed. Data for investment and depreciation by industry are also available for the construction of appropriate indexes for the capital variable. The Bureau of Labor Statistics publishes a great deal of data on employment and work hours.

Cobb and Douglas performed their earliest study on U.S. manufacturing in the form of a time series regression for the years 1899 to 1922. Using the original version of their formula, they obtained the following result:

$$Q = 1.01L^{0.75}K^{0.25}$$

Other studies used the same technique with similar results.

¹⁸Robert N. Mefford, “Introducing Management into the Production Function,” *Review of Economics and Statistics*, February 1986, pp. 96–104.

¹⁹Craig S. Galbraith and Emmanuel Nkwenti-Zamcho, “The Effect of Management Policies on Plant-Level Productivity: A Longitudinal Study of Three U.S. and Mexican Small Businesses,” *Journal of Small Business Management*, Vol. 43, Issue 4, October 2005, pp. 418–31.

²⁰Ahmad Afrooz, Khalid B Abdul Rahim, Zaleha Bt Mohd Noor, and Lee Chin, “Human Capital and Labor Productivity in Food Industries of Iran,” *International Journal of Economics and Finance*, Vol. 2, No. 4, November 2010, pp. 47–51.

In 1937, David Durand suggested that the equation should not necessarily limit the results to constant returns to scale.²¹ After Durand's article, it was accepted that the exponents in the equation no longer had to equal 1. Douglas corrected the original study and found that the coefficient of labor was reduced to about two-thirds, whereas capital's exponent rose to about one-third. But the sum of the two exponents still summed to about 1, and thus constant returns to scale appeared to prevail.²²

At about the same time, cross-sectional analysis, rather than time series, came into use, and most of the studies done since then have used this technique. The observations were now individual industries in a particular year. Many other studies were conducted on data from the United States, Australia, Canada, and New Zealand. In a majority of these investigations, the sum of the exponents turned out to be very close to 1.

In another study, the author performed a cross-sectional study of eighteen industries using data from the Census of Manufactures.²³ The observation units were individual states. Three independent variables were used:

1. Production worker hours
2. Nonproduction worker years
3. Gross book value of depreciable and depletable assets

The dependent variable, Q , was represented by the value added in each industry.

The results showed that the sums of the three exponents in the eighteen industries ranged from about 0.95 to 1.11, indicating a span from decreasing to increasing returns. Tests of significance showed that in only five of the eighteen industries was the sum of the coefficients significantly different from 1. Thus, again, in the majority of cases, constant returns to scale appeared to dominate.

A more recent time series study estimated China's economic growth during the period 1952 to 1998 and attempted to predict its growth through 2010. It uses the traditional Cobb-Douglas function with GDP as the dependent variable. Labor and capital are the independent variables, and a third variable is also added. This is a trend variable that represents technical progress. The computed coefficients are 0.6136 for capital and 0.4118 for labor, and the trend coefficient equals 0.0263. Actually, the trend variable is fixed at zero through 1978 and shows the above increase after 1979. The results show returns that are very close to constant.²⁴

Another recent study investigated the impact of computer-related input on firm productivity for a sample of 399 enterprises in Shanghai, China. It found that, in the heavy industry sector, the number of computer-using workers had a much more positive effect on output than the number of workers not using computers. This effect was much less for state-operated enterprises than non-state-operated enterprises. Also interesting was the conclusion that the contribution of labor to output was less than capital's contribution.²⁵

A rather interesting application of the cross-sectional Cobb-Douglas function is represented by a study of the importance of various baseball statistics to the success of a baseball team. The dependent variable was a team's win/loss percentage relative

²¹David Durand, "Some Thoughts on Marginal Productivity with Special Reference to Professor Douglas," *Journal of Political Economy*, 45, 6 (December 1937), pp. 740–58.

²²The history of the Cobb-Douglas function and the major studies were summarized by Douglas himself in his article, "The Cobb-Douglas Production Function Once Again: Its History, Its Testing, and Some New Empirical Values," *Journal of Political Economy*, 84, 5 (October 1976), pp. 903–15.

²³John R. Moroney, "Cobb-Douglas Production Functions and Returns to Scale in U.S. Manufacturing," *Western Economic Journal*, 6, 1 (December 1967), pp. 39–51.

²⁴Gregory C. Chow and Kui-Wai Li, "China's Economic Growth: 1952–2010," *Economic Development and Cultural Change*, 51, 1 (October 2002), pp. 247–56.

²⁵Y. C. Ng and M. K. Chang, "Impact of Computerization on Firm Performance, a Case of Shanghai Manufacturing Enterprises," *Journal of the Operational Research Society*, 54, 10 (October 2003), pp. 1029–37.

to the league average. The independent variables were slugging percentage, on-base percentage, number of stolen bases, unearned runs allowed, and earned run average. Surprisingly, the offensive statistics proved to be more important than the defensive statistics. All coefficients were statistically significant, and the equation exhibited increasing returns to scale.²⁶

THE IMPORTANCE OF PRODUCTION FUNCTIONS IN MANAGERIAL DECISION MAKING

As stated in the introductory section of this chapter, the production function is an important part of the economic analysis of the firm because it serves as the foundation for the analysis of cost examined in Chapter 7. But for managers, an understanding of the basic concepts discussed in this chapter also provides a solid conceptual framework for decisions involving the allocation of a firm's resources both in the short run and in the long run. Another key management principle illustrated in the economic theory of production is planning, discussed in the following subsection.

Careful Planning Can Help a Firm to Use Its Resources in a Rational Manner

In our discussion of the short run, we state that a firm is expected to have three stages of production. Stage I represents the underutilization of a firm's fixed inputs relative to its variable ones. Stage III represents an overutilization of its fixed inputs relative to its variable ones. Indeed, firms operating in this stage would find their total output *decreasing* as they increased their variable input. The only stage for a rational firm to be in is Stage II. Assuming this information is well known to managers, why would a firm find itself in Stage I or III? The answer is, of course, that production levels do not depend on how much a company wants to produce, but on *how much its customers want to buy*.

Suppose we consider the short-run production function shown in Figure 6.9. As we can see, Stage II applies to production levels between $Q_1 = 200$ and $Q_2 = 275$. If people want to buy less than 200 units or more than 275 units, for example, then in the short run the firm would be forced to operate in either Stage I if demand is less than 200 units or fail to meet demand of more than 275 units even if X employment is expanded into Stage III.

The information in Figure 6.9 implies that for a firm to avoid having to operate in either Stage I or Stage III, there must be careful planning regarding the amount of fixed inputs that will be used along with the variable ones. In business, this is often referred to as *capacity planning*. For example, if the firm anticipated that the demand for its product would be in the range of 200 to 275, then the capacity implied in Figure 6.9 is perfect for its needs. However, if a firm forecasts the demand to be greater than 275, it would have to consider increasing its capacity so Stage II would include the higher level of output. By the same token, if the firm forecasts a demand less than 200, it would have to consider decreasing its capacity. These alternative capacity levels are illustrated in Figure 6.10.

Good capacity planning requires two basic elements: (1) accurate forecasts of demand, and (2) effective communication between the production and marketing functions (particularly in large organizations where these functions are often handled by separate work groups). The first element is rather obvious but, as you have seen in Chapter 5, not easy to achieve. The second element may not be so obvious, especially

²⁶Mark D. Woolway, "Using an Empirically Estimated Production Function for Major League Baseball to Examine Worker Disincentives Associated with Multi-Year Contracts," *The American Economist*, 41, 2 (Fall 1997), pp. 77–83.

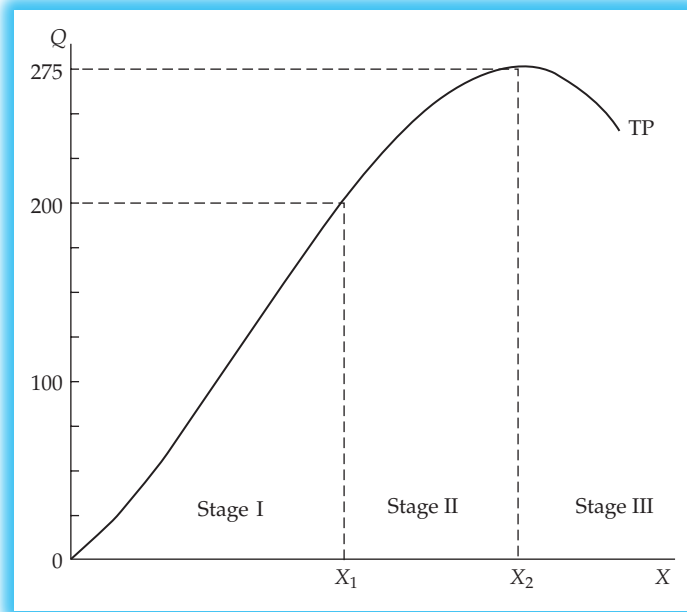
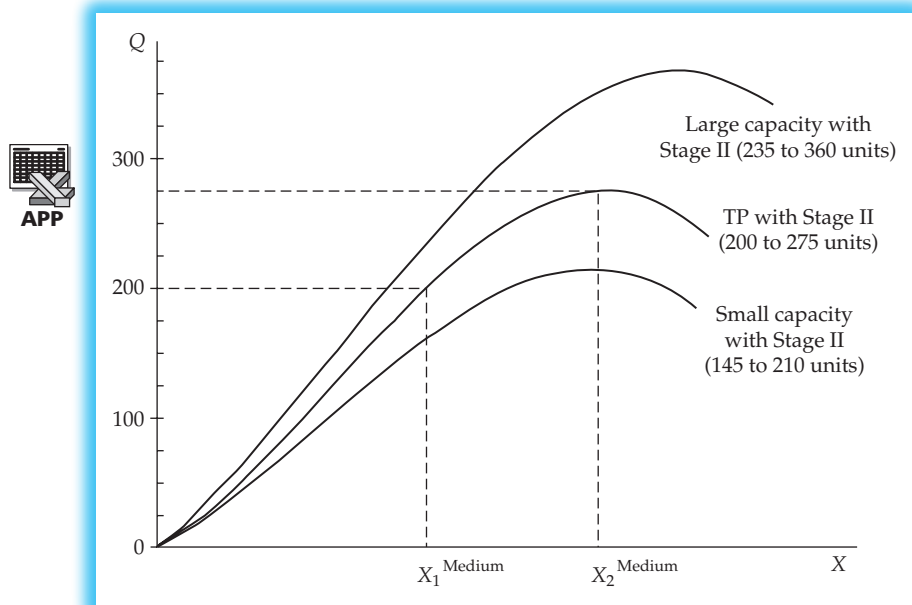


Figure 6.9 Production Stages and Capacity Planning

for those who have not had work experience in large organizations. It is not uncommon for manufacturing people to proceed merrily with their production plans on a purely technical basis (i.e., from a strictly engineering point of view) without fully incorporating the marketing plans of those whose main responsibility is to sell the

Figure 6.10 Adjusting Capacity Based on Demand



products. It is also quite possible for marketing people to try to sell as many units of the product as possible (as marketing people are supposed to do) without consulting the production people as to whether the firm has the capacity to meet the increase in demand. A full discussion of these problems of management and organization is beyond the scope of this chapter. However, they are mentioned to underscore the importance for managers of understanding production theory.

GOING “BEYOND THE CURVES”: CURRENT PRODUCTION ISSUES AND CHALLENGES FOR TODAY’S MANAGERS

The heart of the economic theory of production can be expressed as production curves showing the relationship between inputs and output. But the intensity of current global competition often requires managers to go beyond these curves. Being competitive in production today mandates that today’s managers also understand the importance of speed, flexibility, and what is commonly called “lean manufacturing.”

Speed of production is required because rapidly changing technology renders products on the market obsolete very quickly and because today’s consumers often change their tastes and preferences at the same fast pace. These “supply side” and “demand side” changes are quite evident in the world of consumer electronics, in particular the market for mobile devices such as smartphones and mobile phones. For example, in recent years, there has been a significant shift by consumers away from cell phones to smart phones. Companies such as Apple and Samsung are benefiting from this shift, while companies such as Nokia have been falling behind. Nokia was once the dominant seller of mobile devices because of the success of its cell phones. Now it has just about lost its top position to Samsung (in revenue) and to Apple (in profit) because it failed to keep up with the shift to the market for smart phones. Other companies that once led the field for mobile devices have also fallen in stature because of their failure to cater to changing demand. Blackberry, once the leader in sales of smart phones, has dropped behind Apple and Samsung. Motorola, once the leader in cell phone sales in the early days of the mobile device market, sold its mobile phone division to Google because it was simply unable to compete any longer in the rapidly changing market.

One of the best examples of the importance of speed and flexibility in production is the phenomenal growth of the Spanish retail store Zara. Numerous case studies have been written about the amazing rise of this trendy clothing chain. In Spain, Zara’s home base, the company has its own manufacturing facilities located relatively close to its stores. Zara does contract out some of its production to offshore operations in Asia, but reserves these facilities for the production of less trendy, classic fashions. Its company-owned local factories are in constant communication with its stores in order to get feedback on the latest consumer whims and fancies. Its factories are able to respond to suggestions from its stores within 2 to 4 weeks. Competitors’ normal response time could be between 4 to 12 months. Zara produces each fashion item in small quantities, thereby creating a type of scarcity. Shoppers in a Zara store are often told by salespeople that they should buy the item when they want it because they might not find it when they return to the store.²⁷ Zara also produces a lot more styles, thereby

²⁷Salespeople in Costco, the successful warehouse store chain based in the United States, are known to tell this to their customers as well.

giving its customers more choices while at the same time increasing its chances of finding “hit” fashion items.²⁸

An important tool for helping companies to increase both their speed to market and their productivity is an approach originally called lean manufacturing.²⁹ Because this approach is now being applied to services as well as manufacturing, it is often simply referred to as lean, or the lean approach. The lean approach was developed and perfected by Toyota in the early 1980s. At that time, the senior managers of Toyota decided that the mass production, assembly line approach developed years ago by the Ford Motor Company would not work for them. As implied in its name, mass production works best when large numbers of the same type of output are produced. Only then can a manufacturer begin to reap the benefits of economies of scale.

In 1981, Japan voluntarily agreed to restrict their export of cars to the United States (see Chapter 4). But at the same time, Japan’s consumers were demanding a wide variety of cars. Additional factors affecting the labor force also played a role in Toyota’s move to the lean manufacturing approach. The U.S. occupation of Japan after World War II introduced labor unions into the country, which in turn led to no-layoff clauses and higher wages in labor contracts. Japan’s strict immigration laws made it nearly impossible to use temporary as well as permanent foreign workers (as was the case in the United States and Germany). To adjust for an expensive, inflexible labor force; limited production runs of one type of automobile; and an overall lower level of total production, Toyota had to adopt an entirely new approach to making cars.³⁰

Some important aspects of the lean approach are (1) *kaizan*, or continuous improvement events; (2) the elimination of *muda* (Japanese for “waste”); and (3) the 5 S’s. Examples of *muda* are:³¹

1. Overproduction
2. Waiting (underutilization of assets affects downstream production)
3. Unnecessary movement of materials (wastes time and increases chances of handling damage)
4. Extra processing (extra operations such as rework, preprocessing, handling, or storage)
5. Inventory
6. Motion (unnecessary motion)
7. Defects

The 5 S’s are terms in Japanese that all begin with the letter S. Translations of these terms may vary. For example, one version of the 5 S’s is sort, set in order, shine, standardize, and sustain.

Over the years, Toyota has certainly set the standard in lean manufacturing, and its success in this production approach has undoubtedly helped the company to achieve its status as the best-selling car maker in the world. However, the game of business never ends. In the first quarter of 2012, Toyota continued to lead its closest rivals—GM and Volkswagen—in number of cars sold. But GM has come back strong

²⁸Devangshu Dutta, “Retail@the Speed of Fashion,” reference article taken from www.3isite.com. See also K. Ferdows, M. Lewis, and J. Macluca, “Rapid Fire Fulfillment,” *Harvard Business Review*, Vol. 11, no. 11, November 2004. On a recent trip to Madrid, one of the authors was told by a salesperson at one of Zara’s busiest stores near the Plaza Mayor that it was not unusual for new items to arrive several times in the same day.

²⁹For an excellent brief overview of the lean approach, see George Koenigsaecker, *Leading the Lean Enterprise Transformation*, New York: CRC Press, 2009.

³⁰John Gattorna, *Dynamic Supply Chains*, 2nd ed, Upper Saddle River, NJ: Prentice Hall, 2006, pp. 198–99.

³¹See www.leaninnovations.ca/seven_types.html

from bankruptcy, and Volkswagen is determined to achieve the number 1 spot by 2016. In 2010, Toyota sales suffered when it had to recall hundreds of thousands of its hybrid cars due to a faulty brake problem, leaving many to wonder about the infallibility of its lean production approach.³²

CALL CENTERS: APPLYING THE PRODUCTION FUNCTION TO A SERVICE

Most production function examples involve manufacturing or agriculture. To illustrate how the concepts in this chapter can be used for a service activity, let us consider the example of a call center represented by the following production function:

$$Q = f(X, Y)$$

where Q = number of calls

X = variable input (this includes call center representatives and the complementary hardware such as PCs, desks, and software if site licenses are sold on a per-user basis)

Y = fixed input (this includes the call center building, hardware such as servers and telecommunications equipment, and software site licenses up to the designated maximum number of users)

Using this basic production function, it is possible to consider a number of applications of concepts. The following are offered as examples.

Three Stages of Production: Stage I could be a situation in which there is so much fixed capacity relative to number of variable inputs that many representatives sit around idle, waiting for calls to come in. Stage II could be a situation in which representatives are constantly occupied and callers are connected to representatives immediately after the call is answered or are kept waiting no more than a certain amount of time (e.g., 3 minutes). If callers are kept waiting for longer than 3 minutes, the call center manager might consider adding more call center representatives. Stage III could be a situation in which callers begin to experience a busy signal on a more frequent basis or call representatives may begin to experience a slower computer response or more frequent computer “down times.” These are all technical manifestations of overloaded computer processing power or transmission capacity (i.e., insufficient bandwidth).

Input Combinations: A number of input combinations could be considered in the operation of a call center. To begin with, there is the location trade-off in which the productivity of call center representatives in two or more locations might be considered relative to the cost of each representative (including wages, training costs, costs associated with turnover rates, etc.).

In the early days of call center operations, numerous cities in the United States established themselves as attractive locations for call centers because of their large pools of qualified labor available at wage rates relatively lower than the national average. High on the list of these were Sioux Falls, South Dakota; Jacksonville, Florida; St. Louis, Missouri; and Salt Lake City, Utah. Early on, Salt Lake City was identified as being a desirable location for call centers because of its large concentration of people with foreign language skills.³³ Then, call center operations in the United States and the

³²See <http://www.nytimes.com/2010/02/04/business/global/04prius.html>

³³Alan Crane, “Telephone Contact with Customers Is Ripe for Outsourcing,” *Financial Times*, September 13, 1996.

United Kingdom started to shift to India, the Philippines, and Central America, where there are large pools of qualified English-speaking workers.³⁴

Returns to Scale: One can well imagine that a call center would be amenable to increasing returns to scale. This is evidenced by the number of smaller companies who outsource their operations to companies that specialize in call center operations. These “third-party” vendors are able to provide call center services at a lower cost to the company mainly because their size enables them to take advantage of increasing returns to scale.

GLOBAL APPLICATION: SHIFTING TRENDS IN GLOBAL OUTSOURCING

China is sometimes called the world’s factory and India is sometimes called the world’s back office. Economic data indeed supports these generalizations. China began its climb to global prominence by manufacturing products that required a high labor content and relatively less capital and technology. The apparel industry is a good example. By the same token, in the service sector, India began by gaining prominence in call center operations. In recent years, companies have been starting to shift their production of apparel and their call center operations to countries other than China and India, respectively. One would assume the wage rate is the primary factor driving their decisions to do so. But this is not quite the case.³⁵ True, wage rates in China have been rising significantly. However, there are other factors such as speed to market. In Western Europe, for example, more of the big-name designers have been shifting their production to countries such as Morocco, Tunisia, Macedonia, and Bulgaria, where the hourly wage for garment workers are about 20 to 30 percent higher than in China, but where the supply chain is shorter and less complex. The next time you shop for clothes at some of the major retailers such as the Gap, Banana Republic, Abercrombie and Fitch, or Club Monaco, check out the labels. You are as likely to see “Made in Vietnam” or “Made in Madagascar” as you are to see “Made in China.”

For years, India has led the world in global outsourcing revenue. But in recent years, call center operations have grown considerably in the Philippines, to the point where it appears to have passed India in the number of call center workers.³⁶ However, India still leads the world in the total revenue by its companies in this industry.³⁷ India’s government and business leaders believe that the country’s position as the leading outsourcing provider of higher value-added services in software and business consulting is not being threatened by the movement of call center operations to the Philippines and other countries. Going back to the point made about the apparel industry in China, we see that, in fact, the Chinese government is carrying out policies to focus the country more on higher value-added manufacturing in the high-tech sector and less on lower-value, labor-intensive sectors such as apparel and toys.

³⁴“U.S. Dials 1-800-Bangalore: Toll-Free Inquiries Are Increasingly Answered in India,” *Washington Post*, August 5, 2001. It is interesting to note that in France, call center operations have been set up in the French-speaking countries in North and Central Africa.

³⁵<http://www.reuters.com/article/2012/03/22/us-china-fashion-idUSBRE82L0SC20120322>

³⁶<http://call-center-outsourcing.tmcnet.com/topics/call-center-outsourcing/articles/239669-philippines-surpasses-india-call-center-workers.htm>

³⁷http://www.businessweek.com/magazine/content/10_50/b4207017538393.htm

The Solution

To: Robert Burns, CEO
Nicole Goodman, SVP of Marketing
Jim Hartwell, SVP of Manufacturing

From: Christopher Lim, Production Manager

Re: The Economic Feasibility of Packaging in Glass Bottles and Key Related Issues

Before beginning plans to bottle and distribute Waterpure, I feel that I must bring up certain production issues that I was not able to cover at our meeting last week. I have had considerable experience in glass packaging from my previous job as head of bottling for our Latin American beverage division. Glass is a more expensive material than either plastic or aluminum. With the latest slump in aluminum prices and technology developments in plastic, the price differential is widening even further. The reason why glass is so expensive is that it is much heavier and more fragile than these other materials. This causes higher production costs because (1) there are more back injuries to employees (thereby increasing our medical insurance premiums), (2) there is a need for special handling care (e.g., bottle carriers, protective cardboard), (3) breakage occurs, and (4) the greater weight increases the cost of fuel used in transportation.

We use a considerable number of glass bottles in emerging market countries because the total cost differential between glass and aluminum or plastic is not as great in these countries. For example, because of lower labor costs in South America, we can better afford to use the more labor-intensive process of handling the heavy glass bottles. The machinery to blow out the plastic pellets into 1- and 2-liter bottles is a lot more expensive in these countries than here in the United States. Furthermore, people in emerging market countries seem to like the idea of receiving their deposit back when returning glass bottles, however small the amount might seem to you.

However, a real issue that I think we should address even before devising a packaging, production, and distribution plan is the source of our bottled water. As you know, to be considered “spring water,” bottled water must actually come from a spring. The industry has grown so rapidly over the past few years that companies are having a harder time finding good sources of spring water. I have heard of bottled water companies hiring geologists to look for water just as oil companies hire them to look for new sources of oil. Therefore, we should consider the cost of “manufacturing” the product itself—water—before even considering the cost of packaging. I eagerly await your response to the issues raised in this memo.

SUMMARY

The topics in this chapter represent the foundation for the economic analysis of supply. After all, people may be willing to purchase a firm’s product at a certain price, but will the firm be willing to supply the product at this price? The answer to this question begins with the relationship between the firm’s inputs and the resulting output, that is, the firm’s production function. In the short run, where at least one of the firm’s inputs is fixed, we have learned that the firm is subject to the law of diminishing returns and the three stages of production. This means that as additional inputs are added to the fixed input, at some point the additional output (i.e., marginal product) resulting from the additional input will start to diminish. Once this level of production is exceeded, the output per unit of variable input (i.e., average product) will reach a maximum and then start to diminish. The point of maximum average product marks the end of Stage I and the beginning of Stage II, the stage in which the rational firm should be operating. The use of still more units of variable inputs will eventually cause the total output to decline (i.e., cause MP to assume negative values). By assigning monetary values to both the variable input and the output along with the use of marginal analysis, we were able to determine precisely where in Stage II the firm should be operating. A similar analysis can be used to derive the conditions for the optimal use of more than one input. (See Appendix 6C.)

The long-run function, in which a firm is able to vary all its inputs, was also considered in this chapter. When a firm is able to vary its entire scale of production, it may experience varying returns to scale. That is, the increase in output may be proportional, less than proportional, or greater than proportional to the increase in all its inputs.

Most studies of production functions have used an exponential expression that results in a monotonically increasing output as inputs are added. This model was introduced by Cobb and

Douglas in the 1920s. The original studies used the time series method of analysis, but researchers soon switched to cross-sectional regression, which they found more useful. The Cobb-Douglas function permits the investigation of both marginal product in short-run situations (with the presence of a fixed factor) and returns to scale in the long run. It is difficult to summarize the large number of studies conducted over the years, but the results generally indicate that constant returns to scale are the rule in manufacturing industries in the United States and in other countries.

In Chapter 7, we present an analysis of a firm's cost function. We then see how a solid background in the economic analysis of production will provide a better understanding of the cost structure of a firm, in both the short and the long run.

IMPORTANT CONCEPTS (Items with * are discussed in Appendix 6B or 6C.)

Average product (AP). The total product divided by the number of units of a particular input employed by the firm. (p. 183)

Cobb-Douglas production function. A power function in which total quantity produced is the result of the product of inputs raised to some power (e.g., $Q = aL^bK^c$). (p. 198)

Inputs. The resources used in the production process. Examples in economic analysis generally involve the inputs *capital* (representing the fixed input) and *labor* (representing the variable input). Other terms used in reference to these resources are *factors* and *factors of production*. (p. 181)

***Isocost.** A line representing different combinations of two inputs that a firm can purchase with the same amount of money. In production analysis, the isocost indicates a firm's budget constraint. (p. 237)

***Isoquant.** A curve representing different combinations of two inputs that produce the same level of output. (p. 231)

Law of diminishing returns. A law stating that as additional units of a variable input are added to a fixed input, at some point the additional output (i.e., the marginal product) will start to diminish. Because at least one input is required to be fixed for this law to take effect, this law is considered a short-run phenomenon. (p. 185)

Long-run production function. The maximum quantity of a good or service that can be produced by a set of inputs, assuming the firm is free to vary the amount of *all* inputs being used. (p. 182)

Marginal labor cost (MLC). The additional cost to the firm of using an additional unit of labor. This is more generally referred to as *marginal factor cost (MFC)* or *marginal resource cost (MRC)*. Labor is used in this term because it is the most commonly used variable input in the economic analysis of production. (p. 191)

Marginal product (MP). The change in output resulting from a unit change in one of the firm's variable inputs. (p. 183)

***Marginal rate of technical substitution (MRTS).** Given two inputs X and Y , the

marginal rate of technical substitution of X for Y represents the reduction in Y relative to the amount of X that a firm must add to replace Y to maintain the same amount of output. Mathematically speaking, it is represented by the slope of some given isoquant or $\Delta Y/\Delta X$ ($\partial y/\partial x$ for continuous isoquants). (p. 233)

Marginal revenue product (MRP). The additional amount of revenue resulting from the use of an additional unit of a variable input. It can be calculated by taking an input's marginal product and multiplying it by the market price of the product. For example, given some input i , that $MRP_i = MP_i \times P$. (p. 191)

MRP = MFC rule. A rule that guides a firm in its decision about how many units of a variable input it should use relative to its fixed input. The F in MFC stands for *factor*, another term used by economists for inputs in the production process. The rule states that the firm should employ a particular input up to the point at which the revenue contribution of the additional input is equal to the cost incurred by the firm to employ this particular input. In the case of more than one input, this condition applies separately for every input used by the firm. (p. 191)

MRP = MLC rule. A specific application of the MRP = MFC rule that is the focus of our analysis because labor is the most commonly used variable input in the economic analysis of production. (p. 191)

Power function. A mathematical function of the form $Y = X^n$, where n is a fixed number and X takes positive values continuously. (p. 197)

Production function. The maximum quantity of a good or service that can be produced by a set of inputs. Production functions are divided into two types: short run and long run. (p. 181)

Returns to scale. The increase in output that results from an increase in all of a firm's inputs by some proportion. If the output increases by a *greater* proportion than the increase in inputs, the firm is experiencing increasing returns to scale. If the output increases by the *same* proportion as the inputs, the firm is experiencing constant returns

to scale. Finally, if the output increases by a *smaller* proportion than the increase in inputs, the firm is experiencing decreasing returns to scale. (p. 193)

Short-run production function. The maximum quantity of a good or service that can be produced by a set of inputs, assuming the amount of *at least one* of the inputs used remains unchanged as output varies. (p. 182)

Stages of production. In a short-run production function, there are three stages of production. Stage I starts at zero and ends at the point where the firm has reached the maximum level

of *average* product. Stage II continues from this point on to the point at which the firm has reached the maximum level of *total* product. Stage III continues from this point on. Economic theory suggests that in competitive markets, the rational firm will try to produce in the short run in Stage II. In Stage I the firm would be underutilizing its fixed inputs, and in Stage III it would be overutilizing its fixed inputs. (p. 188)

Total product (TP). The firm's output for a given level of input usage, also referred to as *quantity*, or simply *Q*. (p. 183)

QUESTIONS

1. Explain the difference between a short-run and long-run production function. Cite one example of this difference in a business situation.
2. Define the *law of diminishing returns*. Why is this law considered a short-run phenomenon?
3. What are the key points in a short-run production function that delineate the three stages of production? Explain the relationship between the law of diminishing returns and the three stages of production.
4. Explain why a firm's adherence to the $MRP = MLC$ rule enables it to find the optimal number of units of a variable input to use in the short-run production process.
5. Define *returns to scale*. Why is this considered a long-run phenomenon?
6. According to the rule for optimal input usage, a firm should hire a person as long as her marginal revenue product is greater than her marginal cost to the company. It is well known that many companies have management training programs in which new trainees are paid relatively high starting salaries and are not expected to make substantial contributions to the company until after the program is over (programs may run between 6 to 18 months). In offering such training programs, is a company violating the optimality rule? Explain.
7. Explain the relationship between marginal product and average product. Why can we expect marginal product to equal average product at average product's maximum point?
8. Cite and discuss possible reasons a firm may actually find itself operating in Stage I or Stage III of the short-run production function.
9. Discuss the problems of measuring productivity in actual work situations. How might productivity be measured for each of the following industries?
 - a. Education (e.g., elementary and secondary education, higher education—undergraduate and graduate)
 - b. Government (e.g., the Social Security Office, the Internal Revenue Service)
 - c. Manufacturing (e.g., soap and toothpaste, computers, heavy machinery)
 - d. Finance and insurance (e.g., banks, insurance companies, brokerage houses)
10. For those of you with current or previous work experience, how is (was) productivity measured in your organization?
11. What are the two statistical methods most frequently used to estimate production functions? What are the advantages and disadvantages of each method?
12. Design a study of a production function for a steel mill and another one for a call center. Which variables would you use, and what statistical method would you select for each function? In general, compare and contrast the production function for a product and one for a service.

Refer to Appendices 6B and 6C to help you answer questions 13–16.

13. What are the properties of the Cobb-Douglas function $Q = aL^bK^{1-b}$. What conceptual change occurs when the equation is changed to $Q = aL^bK^c$?

14. In a power function $Q = aV^b$, how can you tell whether diminishing marginal returns are present?
15. When a Cobb-Douglas function with at least two inputs shows the existence of constant returns to scale, it implies that the marginal product of each input is diminishing. True or false? Explain.
16. Write a production function equation that expresses the existence of diminishing marginal returns. How will this equation differ from one that shows both increasing and decreasing marginal returns?

PROBLEMS

1. Indicate whether each of the following statements is true or false. Explain why.
 - a. When the law of diminishing returns takes effect, a firm's average product will start to decrease.
 - b. Decreasing returns to scale occurs when a firm has to increase all its inputs at an increasing rate to maintain a constant rate of increase in its output.
 - c. A linear short-run production function implies that the law of diminishing returns does not take effect over the range of output being considered.
 - d. Stage I of the production process ends at the point where the law of diminishing returns occurs.
2. The Oceanic Pacific fleet has just decided to use a pole-and-line method of fishing instead of gill netting to catch tuna. The latter method involves the use of miles of nets strung out across the ocean and therefore entraps other sea creatures besides tuna (e.g., porpoises, sea turtles). Concern for endangered species was one reason for this decision, but perhaps more important was the fact that the major tuna canneries in the United States will no longer accept tuna caught by gill netting.

Oceanic Pacific decided to conduct a series of experiments to determine the amount of tuna that could be caught with different crew sizes. The results of these experiments follow.

Number of Fishermen	Daily Tuna Catch (lb)
0	0
1	50
2	110
3	300
4	450
5	590
6	665
7	700
8	725
9	710

- a. Determine the point at which diminishing returns occurs.
- b. Indicate the points that delineate the three stages of production.
- c. Suppose the market price of tuna is \$3.50/pound. How many fishermen should the company use if the daily wage rate is \$100?
- d. Suppose a glut in the market for tuna causes the price to fall to \$2.75/pound. What effect would this have on the number of fishermen used per boat? Suppose the price rose to \$5.00/pound. What effect would this have on its hiring decision?

- e. Suppose the firm realizes that to keep up with the demand for tuna caught by the more humane pole-and-line method of fishing, each of its boats must catch at least 1,000 pounds of fish per day. Given the preceding data, what should it consider doing? Explain.
3. A firm has the following short-run production function:

$$Q = 50L + 6L^2 - 0.5L^3$$

where Q = Quantity of output per week
 L = Labor (number of workers)

- a. When does the law of diminishing returns take effect?
 - b. Calculate the range of values for labor over which Stages I, II, and III occur.
 - c. Assume each worker is paid \$10 per hour and works a 40-hour week. How many workers should the firm hire if the price of the output is \$10? Suppose the price of the output falls to \$7.50. What do you think would be the short-run impact on the firm's production? The long-run impact?
4. The owner of a small car rental service is trying to decide on the appropriate numbers of vehicles and mechanics to use in the business for the current level of operations. He recognizes that his choice represents a trade-off between the two resources. His past experience indicates that this trade-off is as follows (see Appendix 6B for help in answering this question):

Vehicles	Mechanics
100	2.5 (includes one part-timer)
70	5
50	10
40	15
35	25
32	35

- a. Assume the annual (leasing) cost per vehicle is \$6,000 and the annual salary per mechanic is \$25,000. What combination of vehicles and mechanics should he employ?
 - b. Illustrate this problem with the use of an isoquant/isocost diagram. Indicate graphically the optimal combination of resources.
5. An American company that sells consumer electronics products has manufacturing facilities in Mexico, Taiwan, and Canada. The average hourly wage, output, and annual overhead cost for each site are as follows:

	Mexico	Taiwan	Canada
Hourly wage rate	\$1.50	\$3.00	\$6.00
Output per person	10	18	20
Fixed overhead cost	\$150,000	\$90,000	\$110,000

- a. Given these figures, is the firm currently allocating its production resources optimally? If not, what should it do? (Consider output per person as a proxy for marginal product.)
- b. Suppose the firm wants to consolidate all its manufacturing into one facility. Where should it locate? Explain.

6. The owner of a car wash is trying to decide on the number of people to employ based on the following short-run production function:

$$Q = 6L - 0.5L^2$$

where Q = Number of car washes per hour
 L = Number of workers

- Generate a schedule showing total product, average product, and marginal product. Plot this schedule on a graph.
 - Suppose the price of a basic car wash (no undercoating, no wax treatment, etc.) in his area of business is \$5. How many people should he hire if he pays each worker \$6/hour?
 - Suppose he considers hiring students on a part-time basis for \$4/hour. Do you think he should hire more workers at this lower rate? Explain.
7. The Noble Widget Corporation produces just one product, widgets. The company's new economist has calculated a short-run production function as follows:

$$Q = 7V + 0.6V^2 - 0.1V^3$$

where Q is the number of widgets produced per day and V is the number of production workers working an 8-hour day.

- Develop a production schedule with V equaling 1 to 10.
 - Calculate average and marginal products.
 - Draw a graph.
8. Suppose Noble's production function (see Problem 7) is as follows:

$$Q = 7V - 0.5V^2$$

where Q is the number of widgets produced per day and V is the number of production workers working an 8-hour day.

- Develop a production schedule with V equaling 1 to 10.
 - Calculate average and marginal products.
 - Draw a graph.
 - Discuss the difference between the form of the production function in this problem and the form in Problem 7. Discuss, among other things, the implications for the three stages of production.
9. The International Calculator Company of China produces handheld calculators in its plant. It tries to keep the number of workers in the plant constant so the only variable factor that can be measured is materials. Over the last seven monthly periods, the data for materials and quantity produced were the following:

Materials	Quantity
70	450
60	430
80	460
95	490
77	465
100	550
85	490

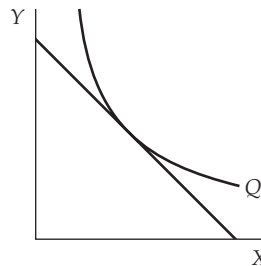
- Calculate a Cobb-Douglas production function of the form $Q = aM^b$.
- Discuss the important properties of your results.
- What is the marginal product of materials?

10. The Brady Corporation has eleven plants located around the world. In a recent year, the data for each plant gave the number of labor hours (in thousands), capital (total net plant assets, in millions), and total quantity produced:

Capital	Labor	Quantity
30	250	245
34	270	240
44	300	300
50	320	320
70	350	390
76	400	440
84	440	520
86	440	520
104	450	580
110	460	600
116	460	600

The plants all operate at a similar level of technology so a production function can be derived from the data.

- Use a Cobb-Douglas production function to calculate a regression, and discuss the important characteristics of your results, such as the form of the equation, R^2 , and the statistical significance of the coefficients.
- Calculate the estimated production for each plant.
- Does the result indicate constant, decreasing, or increasing returns to scale?
- What are the elasticities of production of labor and of capital? What is the meaning of the elasticities?
- Is the marginal product of labor decreasing? (See Appendix 6C to answer this question.)



Refer to Appendix 6B and 6C for help in answering Problems 11–15.

- Show what will happen to the diagram above as a result of the changes listed.
 - The firm's budget increases.
 - The price of Y decreases.
 - The price of X decreases.
 - Y becomes more expensive, and X becomes less expensive.
 - Technology makes the Y input more productive.
 - Technology increases the productivity of both inputs by the same proportion.
- Suppose you are given the following production function:

$$Q = 100K^{0.5}L^{0.5}$$

- a. Use this function to generate the data for the following table:

K										
8										
7										
6										
5										
4										
3										
2										
1										
0	1	2	3	4	5	6	7	8	L	

- b. Identify as many isoquants as you can.
 c. Comment on the returns to scale exhibited in this figure. (*Suggestion:* Start by assuming one unit of L and one unit of K are used. Then systematically increase both inputs by some given proportion.)
13. Following are different algebraic expressions of the production function. Decide whether each one has constant, increasing, or decreasing returns to scale.
- $Q = 75L^{0.25}K^{0.75}$
 - $Q = 75A^{0.15}B^{0.40}C^{0.45}$
 - $Q = 75L^{0.60}K^{0.70}$
 - $Q = 100 + 50L + 50K$
 - $Q = 50L + 50K + 50LK$
 - $Q = 50L^2 + 50K^2$
- g. Based on the answers for the preceding equations, can you make any generalizations about the functional form of a production equation, the relative magnitudes of the coefficients, and the nature of the returns to scale? Explain.
14. Use the production matrix presented to answer the following questions.

Y	8	31	67	101	133	161	184	202	213	
	7	30	62	93	122	147	168	184	193	
	6	27	54	82	108	130	149	168	163	
	5	23	45	69	91	108	126	137	142	
	4	17	34	54	72	89	101	108	111	
	3	12	25	38	54	65	74	79	79	
	2	6	14	24	33	44	54	47	43	
	1	3	7	11	17	27	19	16	8	
		1	2	3	4	5	6	7	8	X

- Determine the returns to scale for this matrix. (Start with one unit of X and one unit of Y .)
 - Suppose the firm has a budget of \$100, and the price of Y is \$20 and the price of X is \$10. What is the optimal combination of inputs X and Y for this firm?
 - Suppose the prices of Y and X are now \$10 and \$20, respectively. What effect will this have on the firm's optimal input combination?
 - Illustrate the answers to the preceding questions with the use of an isoquant/isocost diagram.
15. The economist for the ABC Truck Manufacturing Corporation has calculated a production function for the manufacture of their medium-size trucks as follows:

$$Q = 1.3L^{0.75}K^{0.3}$$

where Q is number of trucks produced per week, L is number of labor hours per day, and K is the daily usage of capital investment.

- Does the equation exhibit increasing, constant, or decreasing returns to scale? Why?
- How many trucks will be produced per week with the following amounts of labor and capital?

Labor	Capital
100	50
120	60
150	75
200	100
300	150

- If capital and labor both are increased by 10 percent, what will be the percentage increase in quantity produced?
- Assume only labor increases by 10 percent. What will be the percentage increase in production? What does this result imply about marginal product?
- Assume only capital increases by 10 percent. What will be the percentage increase in production?
- How would your answers change if the production function were $Q = 1.3L^{0.7}K^{0.3}$ instead? What are the implications of this production function?

Productivity in Services³⁸

The vital importance of service activities to the world economy needs no introduction. The service sector is the fastest growing in the global economy as measured by GDP, employment, and exports. According to the World Trade Organization (WTO), it accounts for roughly two-thirds of global output, one-third of global employment (over two-thirds in the more developed economies), and almost 20 percent of global trade. The ability to innovate and drive productivity growth in service activities is essential for the success of firms and nations, and for raising living standards worldwide.

When economists first developed their theories of production in the first half of the last century, manufacturing was the key sector in the developed countries of North America and Europe, and agriculture still made up a large portion of a country's labor force and output. In the rest of the world, agriculture was the biggest sector. But regardless of which part of the world or which sector, it would be safe to assume that the output Q of the production function would be a tangible good. Now that the service sector is such an important part of the global economy, we believe that it is important to look at some of the special challenges that arise in the measurement and improvement of production when the output Q is a service rather than a good.

Of course in the production function for goods, the output Q is considered at a very high level of abstraction, because there are all types of tangible product outputs. By the same token, the production of services is equally diverse, with different components displaying vast differences in capital intensity, innovation, and prospects for growth. So in measuring and studying service productivity, it is very important to capture these differences and to understand how and why they change over time, across firms or industries, and across locations.

Productivity as a Measure of Health

The concept of productivity is very simple: how much output can be produced from a given set of inputs. The output can be as straightforward as a

single product or service that a firm sells, or it can be a weighted average of products and services aggregated over a firm or other institution, industry, country, or set of countries. Naturally, there are many kinds of inputs, but they are normally grouped into the broad categories of labor, capital, and intermediate inputs.³⁹ Most generally, therefore, productivity can be expressed as:

$$\text{Productivity} = \frac{\text{Output index}}{\text{Combined index for labor, capital, and intermediate inputs}} \quad (6A.1)$$

Three types of productivity are commonly studied:

1. **Labor productivity (LP):** Calculated by ignoring the capital and intermediate inputs in Equation (6A.1). The labor input can be measured in units of time or people and is often aggregated using weights to differentiate between different skill levels. LP is perhaps the most commonly discussed measure of productivity, and in fact many authors refer to it simply as *productivity*. LP is closely watched as a critical measure of economic health, because its growth is strongly linked to the standard of living.
2. **Partial productivity measures (PPMs):** Calculated by selecting some subcomponent of the output and a corresponding subcomponent of the input. Examples would include meals/server (restaurant); calls/person/hour (call center); computer servers managed/full-time employee (IT outsourcing). These kinds of productivity measures tend to be most helpful in targeted analyses of a business activity, or to compare one's execution of a business activity against industry benchmarks.
3. **Multifactor (or total factor) productivity (MFP or TFP):** Calculated by including all relevant outputs and inputs. This is the most complete, all-encompassing productivity measure.

³⁸This appendix was reprinted by permission of Matthew Denesuk, Ph.D., IBM Global Technology Services.

³⁹Intermediate inputs refer to a combination of energy, materials, and enabling services needed to produce the output. If the value of intermediate inputs is removed from the gross output measure, one is left with "value-added." If value-added is used as an output measure, then the intermediate inputs can be ignored.

Because it includes all the relevant inputs and outputs, MFP is the productivity measure that should correlate most closely with a firm's (or other economic system's) profitability and competitive advantage. In practice, however, LP and/or PPMs, which are easier to measure, are found to be good indicators of advantage if they are targeted at the most important elements of the economic activity in question. Because labor is such an important input in many businesses (especially in many services businesses), LP and/or PPMs based on labor inputs are often strategic measures to track and improve.

It is important, nevertheless, to distinguish productivity from cost efficiency. Because productivity measures typically look at input and output *quantities* (rather than *costs* or *prices*), variations in costs and prices (e.g., with activities operating in different countries) can result, at least temporarily, in lower productivity but higher profit. For example, suppose hourly labor productivity decreased and at the same time the cost of labor (i.e., the hourly wage rate) decreased. In such a situation, a firm's profit could actually increase if the decrease in the wage rate more than offset the decline in labor productivity.

How one analyzes productivity depends specifically on the kind of economic system being analyzed, the goals of the analysis, and the kinds of data available. A macroeconomist might be interested in the sustainability of economic growth in a national economy and will employ broad measures of the labor force, capital equipment, and aggregate business outputs to evaluate LP and MFP measures. Partial productivity measures are more relevant in managerial economic decision analysis. For example, in this chapter, we saw how economic theory states that increases in a particular input, holding other inputs constant, result in output changes that follow the three stages of production.

Improving Productivity

In principle, what is needed to improve productivity in service systems (economic systems with a high component of service activities) is the same as in manufacturing: to increase the level of assets (or capital) per worker, where "asset" needs to be understood in its broader sense to include intellectual assets and other "nonrival"⁴⁰ assets. In some

⁴⁰Nonrival assets can be described roughly as assets that are not depleted as they are used (for example, a business process, a manufacturing process, deep industry knowledge, software for which one has full rights).

cases, the path may be a straightforward extension of existing strategies. For example, an airline may improve labor productivity by flying larger planes. But in other cases, a more radical rethinking of the service process may be required to make the service amenable to an increase in assets.

A basic mechanism for productivity improvement, understood at least since the time of Adam Smith, is rooted in the concept of *specialization*. As a worker is able to focus on a relatively narrow task, he or she tends to get progressively better at it (provided job satisfaction does not diminish appreciably). The intensity of time devoted to this task tends to result both in accumulating skill, as well as to the development of better means (assets) for performing the task.

Effective specialization, however, requires that the work be broken into a set of components, each of which can separately support some amount of specialization. This in turn requires that the overall work process be substantially standardized, so that the same components, interacting in essentially the same ways, are employed in most deliveries of the service. Investment in technology (new processes, tools, machinery, etc.) for improving the performance of each component can be justified on the basis of its extent of repeated usage. Scale economies and other efficiencies that correspond to growing productivity inevitably result (see Figure 6A.1). Thus a level of standardization ultimately enables effective increases in assets (in the broad sense) to drive growth in an economic system.

The Challenges of Improving Productivity in Services: "Baumol's Disease"

One cannot really talk about the challenges associated with productivity in service systems without referring to William Baumol. In a famous paper in the latter half of the 1960s,⁴¹ he posited that there

⁴¹W. J. Baumol, "Macroeconomics of Unbalanced Growth: The Anatomy of an Urban Crisis," *American Economic Review*, 57, no. 3 (June 1967): 415–426. Also see: Barry P. Bosworth and Jack E. Triplett (2003), *Productivity Measurement Issues in Services Industries: "Baumol's Disease" Has Been Cured*. The Brookings Institution and James Heilbrun (2003), *Baumol's Cost Disease* (PDF), originally in Edward Elgar, *A Handbook of Cultural Economics*, 2003.

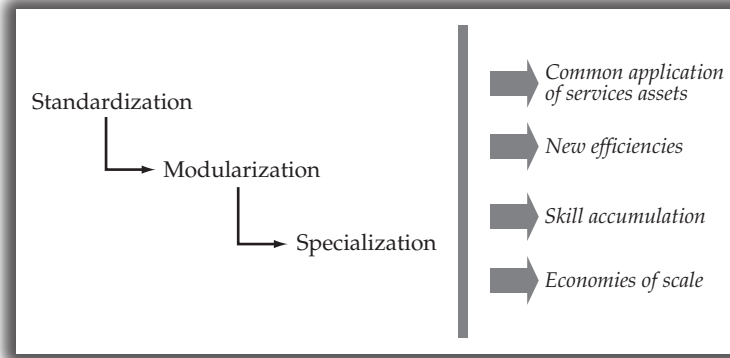


Figure 6A.1 Following a process of standardization → modularization → specialization enables a service activity to consistently leverage various types of assets that enable improving productivity and scalability.

were basically two types of economic activities, distinguished by their “technological structure”:

1. “Technologically progressive” activities, in which capital investment and the potential for economies of scale enable continuing increases in labor productivity
2. Other activities that, while they might experience occasional spurts in labor productivity, fundamentally resist sustained productivity growth

Naturally, one can always divide things into two buckets based on the value of a particular variable, but Baumol created a stir by arguing that to a large extent, *service industries tend to fall into the growth-resisting category*. He and others have articulated arguments for this and their nuances over the years, and they can be understood as emanating from the fact that with most service industries, *the labor input, and its interaction with the customer, is an integral part of the output*.

Several growth-inhibiting factors are associated with this intimate labor-output linkage:

1. **Direct linkage between labor quantity and output quality:** Reducing the labor quantity very often reduces—in fact or perception—the quality of the output, and thereby the effective level of the output. This contrasts with manufacturing, where labor activities are largely shielded from the customer, and where the manufacturer therefore has more latitude in changing the production system in ways that reduce labor.⁴² For example, most people probably have no

idea—and don’t really care—how many people it took to make their refrigerator. But parents do focus on the teacher/pupil ratio at a school, people notice or get annoyed when it seems as if no one is available to help them at the department store, and a business customer will be unhappy if only one junior associate shows up for the kick-off meeting of a consulting engagement.

2. **Customer as input:** The customer’s labor (and sometimes other customer assets) is often an important input, and a firm normally has much less control over this input than over its own labor, making it more difficult to standardize and manage the process effectively.
3. **Customization:** Service offerings are often customized to some extent for each customer (which is typically seen as value-adding), and this limits the extent to which standardization and its efficiency and scale benefits can be leveraged.⁴³
4. **Intangibility of the output:** Because the output of many services activities can be difficult to define consistently (which is in part due to the three points above), it can be very difficult to measure. And, as a general principle, one cannot improve what one cannot measure. This is addressed in more detail in the next section.

Other factors include the fact that services cannot generally be stored (so they cannot be premade), and that their production often requires being located physically near the client.

⁴²In a recent book (*Service Is Front Stage: Positioning Services for Value Advantage*, Insead Business Press, 2006), James Teboul distinguishes services from manufacturing by the relative amount of these “front stage” activities, in which there is direct interaction with the customer.

⁴³It should be noted that made-to-order or just-in-time manufacturing processes (e.g., as pioneered in the personal computer industry by Dell Computer) effectively incorporate a large service aspect into what had been seen as more of a manufacturing business, and this brings some of the challenges in improving production productivity.

Measurement of Services Productivity

Output intangibility, customization and other sources of output heterogeneity, accounting for the customer's inputs, and difficulty in adjusting output measures to include quality factors all contribute to difficulties in measuring service productivity. Although there are also challenges in productivity measurement for more manufacturing-oriented activities, the fact that these latter activities produce a more tangible product, for which the quality factors can generally be well defined and measured in-house before the product is sold, makes these challenges much more manageable.

Consider the commercial airline industry. Because of the substantial degrees of standardization, relatively little customer input quantity, and a high level of technology leverage, this might be considered one of the service industries more amenable to measurement. Conceptually, the purpose of their services is to move people and their luggage from one place to another. Thus, a straightforward measure for output could be revenue passenger-miles (i.e., the number of miles flown by paying customers), perhaps with revenue weightings for international versus domestic and first class versus business class versus coach. In addition, one might want to segment further to account for the major economic differences between shorter and longer flights.

But even with such a refined output definition, investments that result in higher quality service—and thus greater output—will not be reflected in the output measure (except insofar as this quality creates additional demand, which available capacity can fulfill). For example, investment in more flight attendants, gate agents, and/or better meal service can improve the customer experience and benefits substantially. The same can be said for investments that reduce travel delays, such as in-process management and technology, or redundant staff and equipment. At the firm level, a given airline might be able to generate such measures, but it is clearly quite difficult for industry analysts or government agencies to create measures that capture all these components of output and track them accurately over long time periods.

Output measurement is even more challenging in businesses such as consulting, where (1) the customer and his or her organization is a substantial participant (i.e., input) in the value creation process; and (2) each unit of output the service

provider sells can be unique, described by a customized, co-written contract. So how does a \$50K strategy consulting engagement three years ago compare to a \$60K engagement today? Do they correspond to essentially the same level of output, with the difference due to simple inflation-type forces? Or have new technologies improved the ability to execute such engagements, such that twice the level of effective output can now be delivered with modest additional labor input? Or have there been variations in the inputs (data, person-hours, equipment, etc.) that the customer provided in their participation in the process?

In principle, one can simply take revenue as a proxy for output, and this is commonly done in industries for which the government has published “deflators” that correct for inflation and other sources of changes in output level per dollar (e.g., in computer hardware, real output per dollar has grown enormously over the last few decades, and the deflator must account for this). Clearly, however, this just passes the output measurement problem along to the government agency or other group that will publish the deflator.

Although deflators exist for some service industries, there is not sufficient history in many of them, and the relatively coarse granularity is difficult to apply at the firm or market level. On the other hand, analysis of service industries that do not prodigiously employ technology should not be as sensitive to the specific deflator used. Gronroos and Ojasalo⁴⁴ have described many of the difficulties in measurement of service productivity and proposed a simple, easy to calculate measure equal to the total revenue divided by the total cost. Note that this can be seen as closely aligned with an MFP measure (where dollar amounts are used to combine the inputs and outputs), and contains the same information as the profit margin.⁴⁵

Why Service Productivity Matters

With service activities being such a large fraction of developed economies (and growing rapidly in the developing world), an intrinsic resistance to productivity growth is troubling. But in his work

⁴⁴Gronroos and Ojasalo, “Service Productivity: Towards a Conceptualization of the Transformation of Inputs into Economic Results in Services,” *Journal of Business Research*, 57 (2004), pp. 414–23.

⁴⁵Mathematically, this measure = $1/(1 - \text{margin})$.

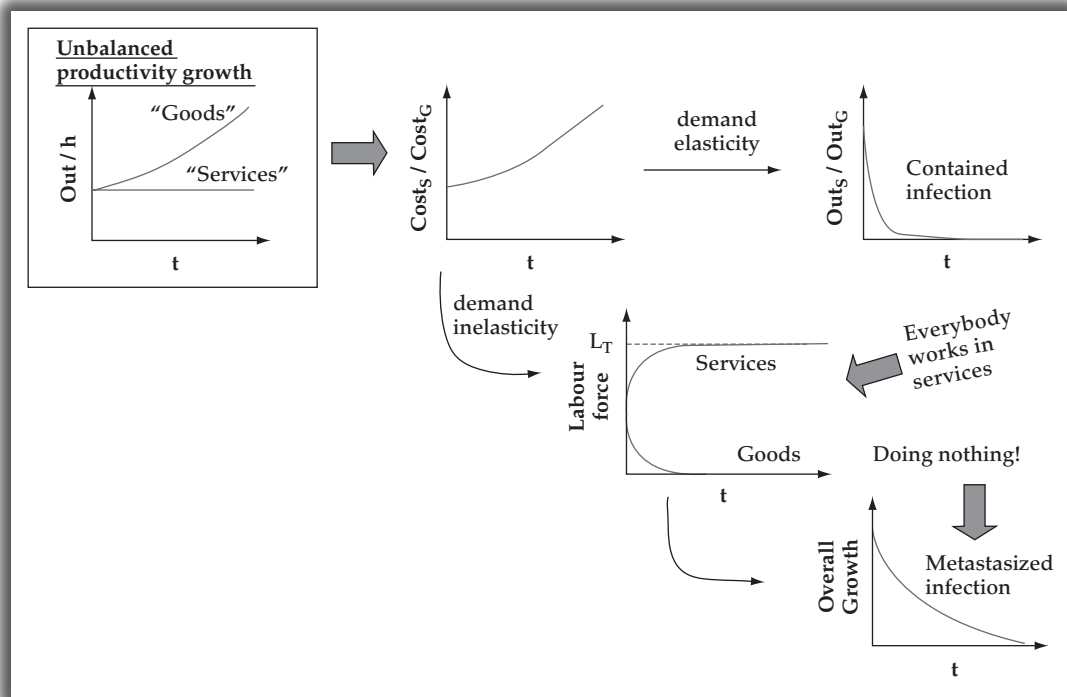


Figure 6A.2 Schematic depiction of “Baumol’s Disease,” in which many services are posited to be intrinsically low-growth, and whereby under certain economic conditions, such low-growth activities can depress an entire economic system.

referred to earlier, Baumol went still further. Based on the price elasticity of demand for a given service, he posited two scenarios (Figure 6A.2):

1. **Elastic demand:** Because productivity stagnation in an underperforming activity will raise its relative costs, if the price elasticity of demand is large (such as for many nonessential goods), then these rising costs should cause demand for the activity to continually diminish. In this case, it reduces the productivity of the overall economy in proportion to its size (a “contained infection,” if you will).
2. **Inelastic demand:** If the price elasticity of demand is small (such as for health care, many foodstuffs), the rising costs will cause the activity to continually absorb more and more of the available labor resources. This will deprive the more productive activities of needed inputs and concentrate resources in the less productive activities, thus producing a “double whammy” effect that can sharply diminish overall economic growth (a “metastasized infection”).

Because services are such a large part of most economic systems, either scenario bodes ill for overall economic growth, but if the inelastic case applies to a substantial fraction of these underperforming services, the deleterious effects are effectively

amplified. Thus a focus on large, relatively inelastic areas (such as health care) is doubly important.

Service Productivity: The Data

Overall productivity growth in the goods sector has traditionally been higher than that in the services sector, and this continues to be true as of this writing. However, the size of the gap between them, which is what spurred the deep concerns associated with Baumol’s Disease, has been narrowing sharply, leading some to claim that “Baumol’s Disease has been cured.”⁴⁶ In fact, recent studies demonstrate that increases in labor productivity growth (i.e., productivity acceleration) has been associated *mostly* with the services sector in the latter part of the 1990s, and due *entirely* to services in the first half of the 2000s—when goods productivity actually declined (Figure 6A.3).

To understand the drivers of this growth, one can use a technique known as *growth accounting*,

⁴⁶For example, see Jack E. Triplett and Barry P. Bosworth in *The New Economy and Beyond: Past, Present and Future*, pp. 34–71, ed. Dennis W. Jansen (Cheltenham, UK: Edward Elgar Publishing, 2006).

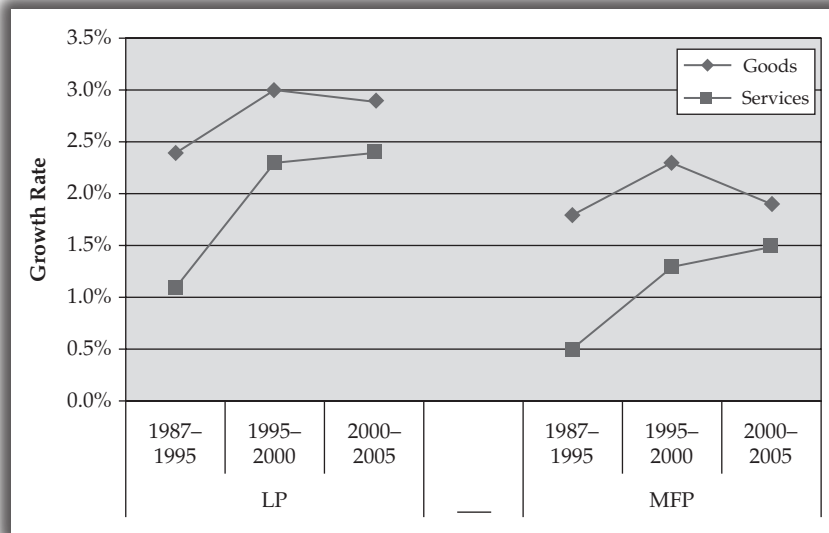


Figure 6A.3 Productivity (LP and MFP) growth rates for the goods sector and the services sector over three time periods between 1987 and 2005. Goods still exhibit stronger productivity growth, but acceleration has been much stronger in services. Data from Barry P. Bosworth, and Jack E. Triplett, “The Early 21st Century U.S. Productivity Expansion Is Still in Services,” *International Productivity Monitor*, 14 (2007), pp. 3–19.

using data describing labor, capital investment, and prices (wages, effective rental prices) to estimate how much of the observed growth is coming from various types of capital deepening and from MFP. A commonly used approach demonstrates that MFP changes can be estimated by subtracting weighted changes in capital intensity from changes in the LP.⁴⁷ The interpretation is that MFP is the “unexplained” portion of the changes in LP and is thus associated with whatever important production factors the analyst has not included. Traditionally, these are grouped under the label “technological changes” and can include, for example, different kinds of intangible capital, improvement in processes, or methods available broadly.⁴⁸ For the study referred to in Figure 6A.3, the estimated relative contributions of capital and labor tell an interesting story (Figure 6A.4). In the goods sector, despite the recent downtrend, MFP plays a dominant role in overall productivity growth. In services, MFP and capital are much more comparable, but the growth in MFP has clearly been outpacing that of the capital

contribution. Whether and to what extent this trend continues, it may be an important indicator of the sustainability of services productivity growth.

Although the services sector overall is exhibiting promising productivity growth, there is a considerable amount of heterogeneity—the sector is made up of many industries that exhibit widely varying productivity characteristics. For example, in the same study referred to in Figure 6A.3, it was found that, despite a jump in the 1995–2000 period, the “food services and drinking places” industry has displayed low and shrinking LP (Figure 6A.5). Both capital investment and MFP have dropped off, and one can wonder whether the jump observed was a one-time event associated with a unique opportunity to deploy new technology. The critical question in a case like this is whether the decline in investment was due to a lack of opportunities for improving productivity—which would bode ill for the future—or due to some other constraint on investment (such as a temporary decline in demand).

For a quite different example, the “professional and business services industry” has shown rapid growth in measured LP, exceeding the sector average by about 0.7 percentage points in the latest period. This was apparently driven by accelerating investments in capital (mostly information

⁴⁷For example, see Robert J. Barro and Xavier Sala-i-Martin, *Economic Growth*, 2nd ed., MIT Press (2004), pp. 434–60.

⁴⁸Note that, defined this way, MFP estimates also include whatever measurement errors may have been made in calculating LP, traditional capital, or the weighting factors.

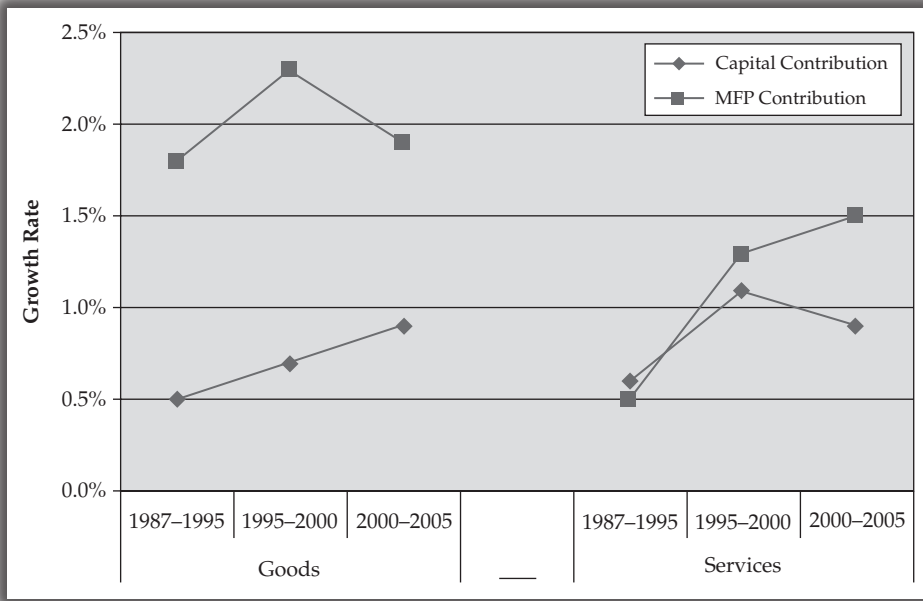
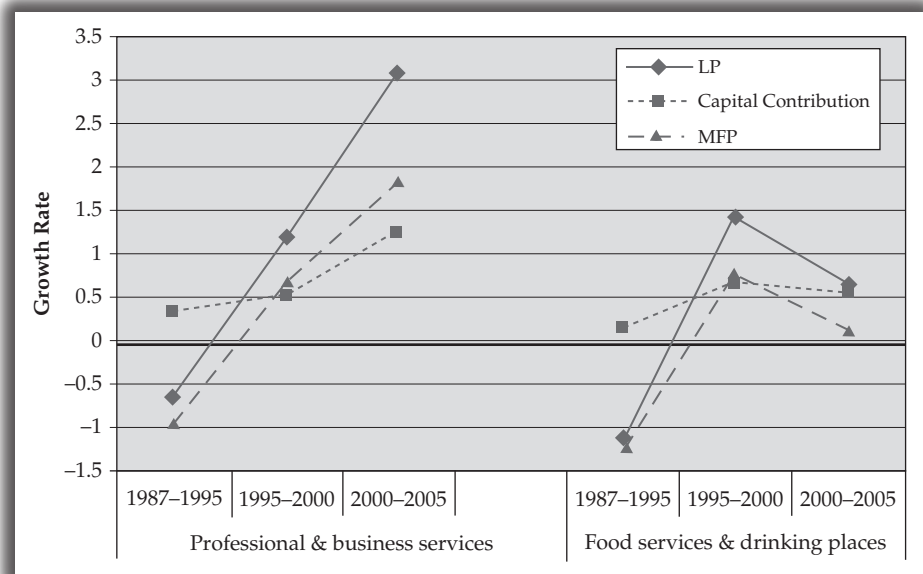


Figure 6A.4 Contributions of MFP and of capital to the observed growth rate in labor productivity over three time periods between 1987 and 2005. Despite a recent decline in capital contribution, the services sector is exhibiting increasing MFP. The data are taken from the same study referred to in Figure 6A.3.

Figure 6A.5 Two examples of labor productivity growth and the contributions of capital and MFP. “Food services & drinking places” shows overall poor productivity performance, despite what appears to have been a short-term boost in the 1995–2000 period. “Professional & business services” has exhibited very strong productivity growth, with no indication that continued growth is not sustainable. The data are taken from the same study referred to in Figure 6A.3.



technology) as well as steadily increasing MFP, and there are no obvious indications that this growth cannot be sustained.

Service Productivity: The Opportunities

General Principles

As stated earlier, productivity in service systems can be increased by increasing the level of asset uptake (where “asset” is to be understood in its broad sense, to include new processes, methods, tools, IT software/hardware, as well as more conventional equipment), and/or by increasing the intrinsic skill level of the people involved (e.g., through training, good HR practices). The assets can be targeted to either replace or enhance labor, or to make labor fundamentally more efficient.

Replace or Enhance Labor In situations where service activities are highly customized, labor can be replaced, for example, with predeveloped output components. These output components may be mass-produced or be nonrival in nature, so that their use increases labor productivity by providing incremental output with essentially no incremental labor. For example, consider the IT services industry. From a firm’s experience-based understanding of difficulties with fraud detection in the insurance industry, it may develop a set of detection algorithms, committed to software code, that it can employ in a nonrival manner to multiple future customers.⁴⁹ Similarly, such a firm may use its experience gathering customer requirements to develop a simplified template to quickly gather the requirements of a new customer. Going further, such a template can lead to the creation of a standard “menu” of service offerings from which the customer may pick and choose.⁵⁰ Such a reduction in scope permits additional opportunities for service firms to predevelop content, tools, and processes and, through specialization effects, to more rapidly improve execution skills. And because the customer’s role in co-creating value in service activities is one of the factors that makes it difficult to

⁴⁹If there is enough demand for such a software asset, it can be made into a conventional product and be governed by traditional product economics (substantial up-front development and hardening, documentation, support infrastructure, etc.). The services software asset must typically be governed by a model in between a product and a service.

⁵⁰Menu-oriented approaches have been successfully applied in many areas, for example, automotive service, financial services, insurance, and (obviously) restaurants.

increase (and even measure) productivity, the use of menus and related constructs serves to more tightly define the customer’s role and permits a more industrialized service. With such “productized” services, the customer may have less choice, but in return may receive higher value through better time-to-value, lower cost, and/or higher quality.

Another example relates to self-service, whereby some amount of the customer’s interaction with a person is replaced by interaction with assets. ATMs used for self-service banking are a common example, as are the many self-help kiosks that are increasingly seen in retail environments. Even in IT services, a customer can be provided with detailed instructions for performing an initial assessment of the current state of their IT infrastructure and for gathering requirements for the desired end-state. From the perspective of the service provider, self-service can increase labor productivity substantially, because less provider labor is needed, and because any additional labor provided by the customer is generally not counted as a formal input.

As discussed earlier, a service often needs to be substantially standardized for tools or predeveloped output components to have a sufficient utilization rate to be effective. Without this standardization, the assets cannot generally apply across enough instances of the provided service offerings for there to be a substantial impact on the overall productivity. There are exceptions, however. For example, providing consultants with general-purpose tools (such as information access tools, collaboration tools, expense reporting tools, or travel planning tools) can improve labor productivity without having to standardize the way these consultants do their core work. Or providing improved excavation equipment or nail guns in the construction industry can improve productivity without reducing the level of customization associated with the work. Often, however, such an approach is limited in terms of the extent to which it can provide sustained productivity growth, or in its ability to fundamentally improve the scalability characteristics (i.e., how output scales with the number of consultants) of the business.

Make Labor Fundamentally More Efficient It has been argued (for example, by Levitt⁵¹) that the apparent resistance of service activities to productivity growth is due more to social factors—such as

⁵¹Theodore Levitt, “The Industrialization of Service,” *Harvard Business Review* (September–October 1976), pp. 63–71.

people's perception of what a given type of service should look like—as well as by a lack of imagination in applying technology and “industrial thinking” to service systems. The fundamentals of labor efficiency can be improved, it is argued, by undertaking a dramatic redesign of the entire service system and/or selecting a narrow market selection in which such redesign is more feasible. Examples include fast-food restaurants (like McDonald's)⁵² and specialized automotive repair/maintenance facilities (like Jiffy Lube), which generally provide reduced menus of service offerings, but are able to provide a level of convenience, speed, reliability, and/or other value attributes that enable them to grow productivity while pleasing their customers. Such industrialized service firms are generally not arrived at through a gradual adoption of new assets, but through a radical redesign requiring substantial up-front investment and a greatly changed (and often untested) value proposition—not for the faint of heart.

In summary, labor productivity in service systems can be improved by enhancing or replacing labor, or by radically changing the structure of the service to use labor more efficiently. These

⁵²McDonald's and other fast-food chains have even started outsourcing the order-taking function in their drive-through operations.

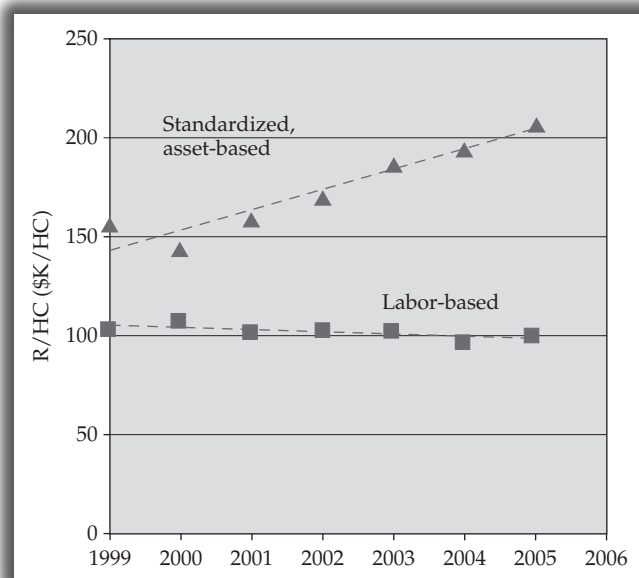
approaches, which require up-front investment, can leverage predeveloped output components, new tools, and rapid skills accumulation through specialization to drive productivity growth.

Real-World Example: IT and Business Services

The ratio of a firm's revenue (suitably deflated) to its headcount is easy to measure and can be taken as a rough proxy for labor productivity. Figure 6A.6 shows panel data for three categories of service firms over the time period from 1999 to 2006. The labor-based group comprises a set of firms that provide service offerings including IT and business process outsourcing, systems and applications management, and consulting—but on a predominantly custom and labor-based basis. Their flat labor productivity (R/HC) behavior indicates that they can grow only by hiring and training large numbers of people.

The standardized, asset-based group includes companies that have generally narrowed their focus and standardized and specialized on a set of asset-based offerings, including payroll and benefits management services, credit and risk management, insurance process services, and financial transaction services. These service activities include substantial labor components, but focus and standardization enable the firms to become more asset-based and to attain continuing productivity increases. These

Figure 6A.6 The behavior of deflated revenue per person—a rough proxy for labor productivity—for labor-based services firms and for asset-based services firms. Asset-based firms are able to increase their labor productivity through technology leverage.



businesses, with growing labor productivity, are fundamentally more scalable because they are not as dependent on large volumes of net new labor acquisition (which involves added costs of hiring, training, retaining) for growth.

The asset-based firms display a growing labor productivity proxy and the apparent ability to grow by making better use of technology. Labor-based firms exhibit a slightly decreasing labor productivity proxy, suggesting that growth requires substantial increases in labor contribution.

Interestingly, further analysis reveals that the labor-based firms invest more in traditional capital assets (which would include computers and other equipment), while the asset-based firms focus more on intellectual and intangible assets, which arise from such activities as R&D and such organizational support activities as sales, marketing, and general administration.⁵³

A Hypothetical Example

A simple, if frivolous, example can be used to demonstrate important concepts related to assets and services productivity growth. Consider a business with a snow removal service offering, where the main asset workers employ is a simple snow shovel. Assume that on average about three hours are required to remove the snow from a customer property using this asset. Such a business could employ

relatively low-wage workers to achieve labor productivity of about three properties per day per person, and would be flexible enough to handle almost any type of job if sufficient labor is available.

If a snowblower appliance is acquired and employed, a worker will require higher skill (and salary), but might be able to complete on average about 10 properties per day per person, provided that the property characteristics were amenable to snow clearing with a snowblower. This improvement, however, requires higher capital investment, a higher-skill/higher-wage employee base, and maintenance expenses (e.g., repairs, gas). Thus, although labor productivity will clearly improve, the behavior of MFP (and profitability) will depend on the price of snowblowers and the cost of capital, maintenance costs, and the utilization of the assets, that is, the quantity of services engagements to which the asset can be applied. This latter factor in turn depends on the marketing/sales function's ability to find and secure enough customers whose properties are amenable to snow removal using a snowblower, and on the quantity of snow days.

If a snowplow is used, still higher skill levels may be needed, but a single worker might be able to complete about 40 properties, provided the property characteristics (e.g., shape, ground contours) permitted the snowplow to be effective. Again, however, to estimate the impact on MFP (and profit), one must evaluate the impact of the other factors of production and the extent to which they will be applied to instances of the firm's services. Two scenarios corresponding to different utilization rates are depicted in Table 6A.1 and Figure 6A.7.

⁵³Matthew Denesuk, "Modelling Productivity and Performance Growth in Labor-Based, Custom Services Firms" paper presented at the Frontiers in Service Conference, San Francisco, CA, October 4–7, 2007.

Table 6A.1 How Service Technology Assets Would Affect a Snow Removal Business

		10 days / year		40 days / year		10 days / year	40 days / year
	Unit price	Labor cost/day	Equip. cost/day	Equip. cost/day	Labor productivity (properties/ day)	Gronroos productivity (MFP)	Gronroos productivity (MFP)
Use shovel	\$30	\$75	\$0.00	\$0.00	3.3	1.3	1.3
Use snowblower	\$30	\$150	\$20.30	\$8.82	10	1.8	1.9
Use truck	\$30	\$200	\$840.00	\$266.00	40	1.2	2.6

Assumptions: unit price constant at \$30; labor markup = 25%; cost of capital = 10%; snowblower cost = \$600 + \$5/day; plow cost = \$30,000 + \$75/day; asset intensity = (asset value) / (revenue).

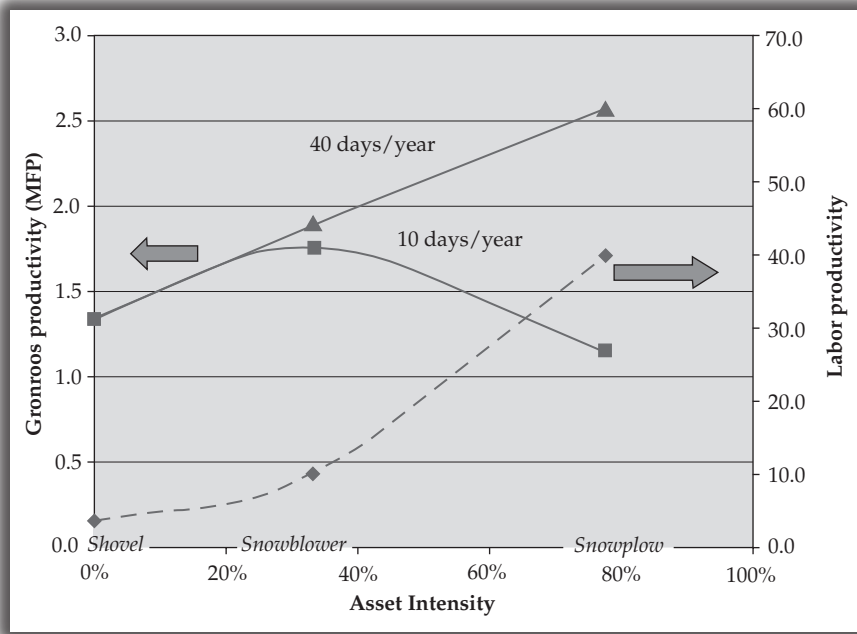


Figure 6A.7 The variation of labor productivity and a specific MFP proxy as a function of asset intensity for a hypothetical snow removal service offering. Although labor productivity increases monotonically with asset intensity, the behavior of the MFP proxy depends on the asset utilization rate.

In this particular example, although labor productivity increases continuously with asset intensity—and dramatically so at higher levels—the MFP proxy, which should relate more closely to overall business performance, exhibits behavior that is very sensitive to the asset utilization rate. In either scenario, the snowblower improves the MFP proxy, but the snowplow only does so at the higher asset utilization rate. Naturally, the actual behavior exhibited depends on the interplay of numerous economic factors, such as

the different wage rates, the cost of capital, maintenance costs, and demand for the different service activities. Clearly, in making business decisions, one must take a holistic view of productivity and its relation to business performance. Although this is a rather mundane example, the principles demonstrated apply more broadly to examples based on intellectual assets, where offering standardization is needed to achieve high asset utilization rates.

The Multiple-Input Case

In this appendix, we examine in greater detail the more general case in which a firm seeks the optimal combination of inputs, rather than simply the optimal level of one particular input. For explanatory purposes, we address the problem of determining the optimal combination of inputs using the two-input case. Mathematically, there is no problem in considering any number of inputs, but we use two inputs for ease of graphical illustration. Nonetheless, it should be noted that the decision rule for determining the optimal combination of inputs is the same whether two inputs or more than two are used in the production process.

From the standpoint of economic theory, the two-input case can be considered either a short-run or a long-run analysis, depending on what assumption is made about the nature of the firm's inputs. If we assume the firm has only two inputs (or, more realistically, that all its inputs can be divided into two basic categories), then the two-input case must be considered a long-run analysis because, in effect, all the firm's inputs are allowed to vary. However, if the firm is assumed to have other inputs that are being held constant while the two inputs are being evaluated, then the analysis must be considered short run. Readers should be able to discern from

the context of our discussion which case applies for our examples.

To illustrate the two-input case, we use the data in Table 6.1, reproduced as Table 6B.1. Suppose the firm produces 52 units of output ($Q = 52$). According to the table, the firm can employ the following combinations of inputs Y and X , respectively: 6 and 2, 4 and 3, 3 and 4, 2 and 6, and 2 and 8. Together, they form the isoquant shown in Table 6B.1. An **isoquant** is a curve representing the various combinations of two inputs that produce the same amount of output. A number of other isoquants are also shown in Table 6B.1. The isoquant for $Q = 52$ is plotted in Figure 6B.1.

A continuous production function, in which inputs are assumed to be perfectly divisible, is illustrated in Figure 6B.2. Here the isoquant appears as a smoothed-out version of that depicted in Figure 6B.1. Notice that both the discrete and continuous isoquants are downward sloping and convex to the origin.

The latter characteristic means that the slope of the isoquant becomes less steep as one moves downward and to the right. These characteristics pertain to the degree to which the two inputs can be substituted for one another.

Table 6B.1 Representative Production Table Illustrating Isoquants

Units of Y Employed	Output Quantity (Q)							
	1	2	3	4	5	6	7	8
8	37	60	83	96	107	117	127	128
7	42	64	78	90	101	110	119	120
6	37	52	64	73	82	90	97	104
5	31	47	58	67	75	82	89	95
4	24	39	52	60	67	73	79	85
3	17	29	41	52	58	64	69	73
2	7	17	29	40	48	52	54	52
1	4	8	14	20	27	24	21	17
	1	2	3	4	5	6	7	8

Units of X Employed

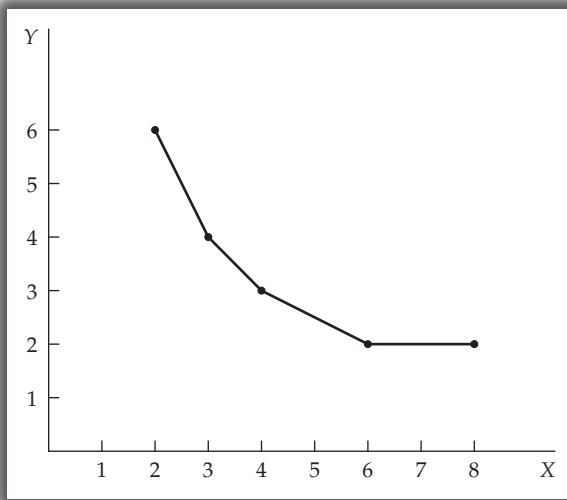


Figure 6B.1 Graph of Isoquant $Q = 52$

Substituting Input Factors

The degree of substitutability of two inputs is a measure of the ease with which one input can be used in place of the other in producing a given amount of output. To explain this further, let us use the example of the making of soft drinks. Consider the ingredients listed on the label of a typical soft drink:

Carbonated water, sugar and/or corn syrup, citric acid, natural flavoring, sodium benzoate as a preservative, and caramel coloring.

Notice that the sweetening component of the recipe is listed as “sugar and/or corn syrup.” This implies that the two ingredients are *perfect substitutes* for each other. The linear isoquant shown in Figure 6B.3a depicts the perfect substitutability between the two ingredients. The hypothetical numbers on the X- and Y-axes serve to caution you not to assume perfect substitutability means the two inputs must be substitutable at a ratio of 1 to 1. In this illustration, the substitution ratio is 2 to 1. That is, 2 grams of sugar can always be substituted for 1 gram of corn syrup, no matter how much sugar or corn syrup is used.

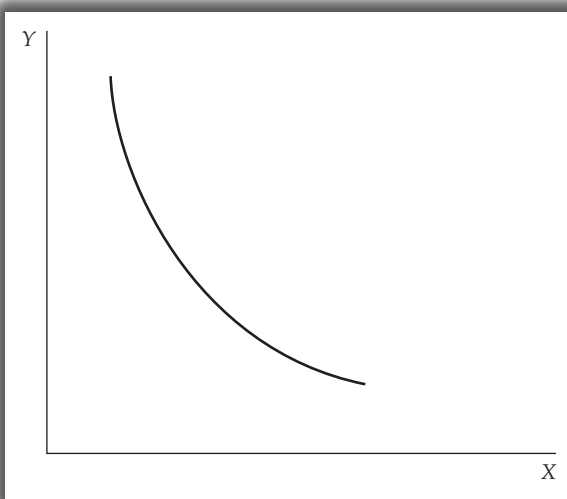


Figure 6B.2 Isoquant for Continuous Production Function

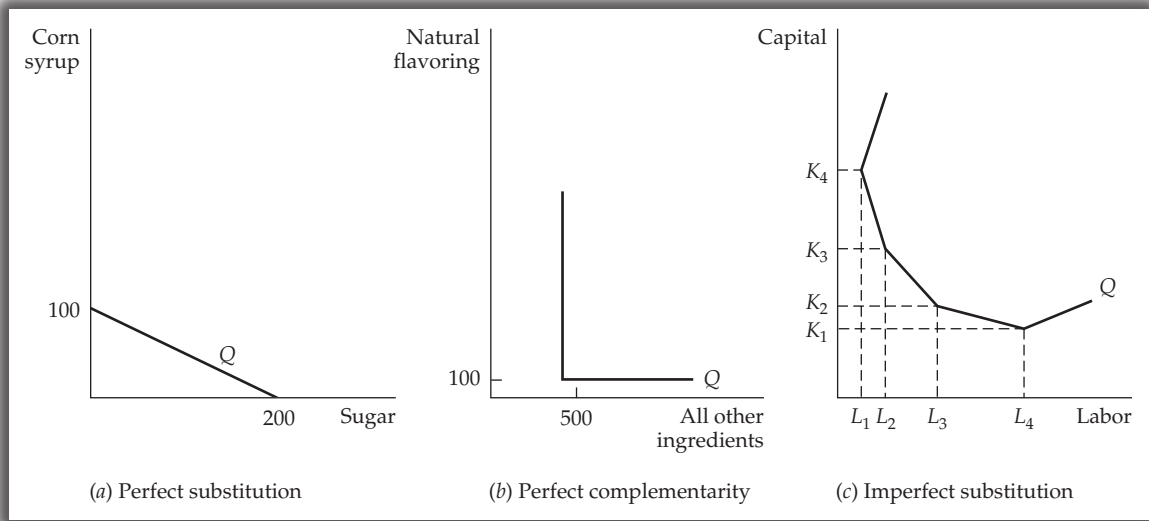


Figure 6B.3 Substitutability of Two Inputs in a Production Function

If we divide a soft drink into two components, “natural flavoring” and “all other ingredients,” we can illustrate the relationship between these two inputs with the isoquant shown in Figure 6B.3b. Presumably, what gives the soft drink its special flavor is the unique ratio of the flavoring to the rest of the components. The hypothetical numbers in Figure 6B.3b indicate that adding one more part of flavoring without increasing the other contents by five parts will not yield more output. Thus, we see that natural flavoring and the other ingredients are what economists term *perfect complements* to each other because they must always be used together in some fixed proportion (i.e., one part of flavoring and five parts of the other ingredients).

Between the extremes of perfect substitutability and perfect complementarity lies the case illustrated in Figure 6B.3c. Here labor and capital are *imperfect substitutes* for each other. That is, one can be substituted for the other, but only up to a limit. (As shown in the figure, any addition of labor beyond L_4 units requires more capital rather than less to maintain the level of output at Q .) Furthermore, as more of one input is used in place of the other, it becomes increasingly difficult to substitute. As can be seen in Figure 6B.3c, when the firm is employing L_1 units of labor and K_4 units of capital, it can substitute a relatively small amount of labor for capital and still maintain the same level of output. However, when it is using L_2 units of labor and K_3 units of capital, it must substitute a larger amount of labor to maintain

the same level of output. More will be said about imperfect substitution in the next section.

In terms of selecting the optimal combination of inputs, it is clear that the case of imperfect substitutability represents the greatest challenge to the firm. Perfectly complementary inputs must be used together in some fixed proportion. The case of perfectly substitutable inputs is equally trivial. If one input can always be substituted in some fixed amount for another (no matter how much of each input is being used), the only determinant of the optimal combination is input price.

In the case of imperfectly substitutable inputs, the optimal combination depends both on the degree of substitutability and on their relative prices. For example, suppose labor costs much less than capital. Does this automatically mean that the firm should use more labor relative to capital (i.e., a labor-intensive production process)? No, because the relative productivities of the two inputs also have to be considered. If capital is much more productive than labor, it will benefit a firm to use more capital relative to labor if the difference in productivity more than compensates for the difference in cost.

Before we explain how economists determine the optimal combination of two inputs that are imperfect substitutes, we should explain how the degree of imperfection is measured. The measurement itself is called the **marginal rate of technical substitution (MRTS)**. We consider an example in which we gradually substitute more of the X input

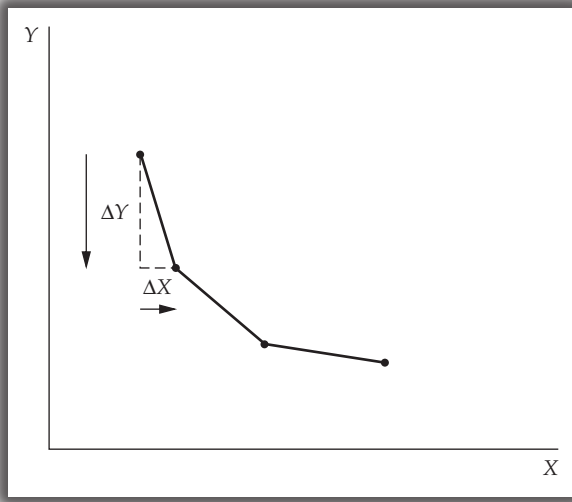


Figure 6B.4 MRTS (X for Y)

for the Y input. Algebraically, the marginal rate of technical substitution of X for Y can be expressed as:

$$\text{MRTS (X for Y)} = \frac{\Delta Y}{\Delta X}$$

Notice that the numerator shows the amount of Y that is removed from the production process, and the denominator indicates the amount of X needed to substitute for Y to maintain the same amount of output. Graphically, the MRTS can be represented by moving downward and to the right along any isoquant (Figure 6B.4). Indeed, if we look back at the algebraic expression of MRTS (X for Y), we see that it is the measure of the slope of the isoquant.

To see exactly how MRTS is measured along the isoquant, let us use the discrete case illustrated in Table 6B.1. The different input combinations that can be used to produce 52 units of output are summarized in Table 6B.2. The changes in Y and X as one moves from combination A to E are shown in Figure 6B.5. Moving from combination A to E, we measure the marginal rate of technical substitution of X for Y as follows:

Movement	MRTS (X for Y) = $\frac{\Delta Y}{\Delta X}$
A to B	$-\frac{2}{1}$
B to C	$-\frac{1}{1}$
C to D	$-\frac{1}{2}$
D to E	$\frac{0}{2}$

Notice that because the isoquant is downward sloping in the economic region of production, $\Delta Y/\Delta X$, or the MRTS, will always have a negative value. However, in discussing the economic significance of MRTS, it is easier to treat it as a positive value. Therefore, let us for the moment simply drop the negative sign. Thus, for example, the MRTS between A and B becomes: $\text{MRTS} = 2/1$. Between B and C and between C and D, the MRTS values are 1 and 1/2, respectively. Expressed as an absolute value, MRTS can be clearly seen to diminish as we move from combination A to combination E. Economists refer to this phenomenon as the *law of diminishing marginal rate of technical substitution*. As you might expect, this relates to the law of diminishing returns.

When we move from A to B, we substitute one unit of X for two units of Y. In other words, the loss in output resulting from the use of two fewer units of Y can be made up for by adding one unit of X. When we move from B to C, the loss in output

Table 6B.2 Input Combinations for Isoquant $Q = 52$

Combination	Y	X
A	6	2
B	4	3
C	3	4
D	2	6
E	2	8

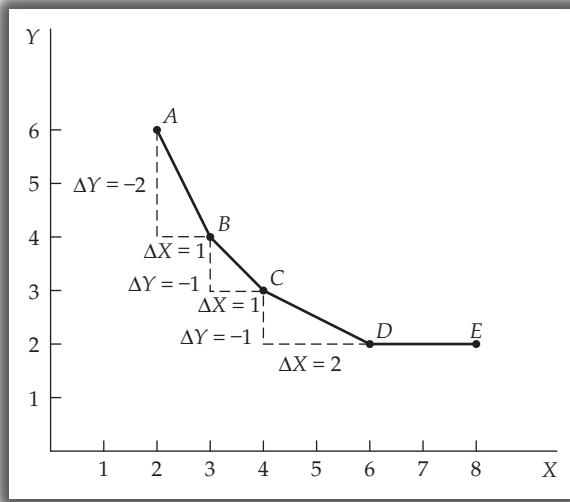


Figure 6B.5 Measuring MRTS along an Isoquant

resulting from the use of one fewer unit of Y must be made up for by adding one unit of X . From C to D we see that the loss in output resulting from the use of one fewer unit of Y must be made up for by adding two units of X . Finally, when we move from D to E , we find that rather than substituting X for Y , we must add more units of X . Because 54 units of output are produced with 7 units of X and 2 of Y , it must be the case that a somewhat smaller value of Y could be used together with 7 units of X to produce 52 units of output (assuming Y can be purchased in noninteger quantities). The move from D to E appears flat, but that is because a 2-unit change in X employment is under consideration. Considered as two 1-unit moves in X , the move from 6 to 7 units of X (from D) involves substituting Y for X along the isoquant. But the move from 7 to 8 units of X (to E) requires that we *increase* Y as we increase X to maintain the output level at 52.

Apparently, as we move from A to D , increasingly more X must be added relative to the amount of Y taken out of the production process to maintain the same amount of output. We started out by having to add only 1 unit of X to replace 2 units of Y and ended up having to add 2 units of X to replace 1 unit of Y . In other words, as more X is used relative to Y in the production process, the productivity of X diminishes relative to Y . This is a result of none other than the law of diminishing returns. Recall that this law states that as additional units of a variable factor are added to a fixed factor, at some point the additional output starts to diminish. If this holds true when one of the inputs is fixed, it must certainly hold when this same input is reduced.

To understand fully why increasingly more of input X is needed to compensate for the loss of a given amount of input Y to maintain the same output, we need to incorporate the concept of marginal product into our analysis. Looking back at Table 6.1 (reproduced as Table 6B.3 for your convenience), we observe that the movement from A to B actually involves two distinct steps. First, input Y is reduced by 2 units (from 6 to 4), resulting in a reduction in output by 13 units (from 52 to 39). Next input X is added to the production process to compensate for the reduction in Y . As can be seen in the table, an additional unit of X is required to restore the output to its original level of 52. This substitution of Y for X is shown as the first L-shaped arrow in Table 6B.3.

Recall from an earlier discussion that the marginal product is defined as the change in output relative to the change in some given input. In this case, the movement from A to B in two separate steps reveals that the marginal product of input Y is

$$\frac{\Delta Q}{\Delta Y} = \frac{-13}{-2} = 6.5$$

The marginal product of input X is

$$\frac{\Delta Q}{\Delta X} = \frac{13}{1} = 13$$

The ratio of these marginal products from A to B is therefore $MP_X/MP_Y = 13/6.5 = 2$. Proceeding next to combination C and then on to D using the same two-step process gives us the following ratios of the marginal products of X and Y : between B and C , $MP_X/MP_Y = 1$; between C and D , $MP_X/MP_Y = 1/2$.

Table 6B.3 Representative Production Table Illustrating Changing MRTS

Units of Y Employed		Output Quantity (Q)							
8	37	60	83	96	107	117	127	128	
7	42	64	78	90	101	110	119	120	
6	37	52	64	73	82	90	97	104	
5	31	47	58	67	75	82	89	95	
4	24	39	52	60	67	73	79	85	
3	17	29	41	52	58	64	69	73	
2	7	17	29	40	48	52	54	52	
1	4	8	14	20	27	24	21	17	
	1	2	3	4	5	6	7	8	
				Units of X Employed					

In reviewing the ratios of the marginal products of X and Y along the isoquant $Q = 52$, you probably have spotted an important link between the MRTS, the slope of the isoquant, and these ratios. Indeed, they are all equal. More specifically,

$$\text{MRTS} = \frac{\Delta Y}{\Delta X} = -\frac{\text{MP}_X}{\text{MP}_Y} \quad (6B.1)$$

This equation is illustrated in Table 6B.4.

Because Equation (6B.1) plays an important part in a later section of this appendix, let us briefly explain its derivation. Consider once again the movement along an isoquant between two given points. As illustrated using specific numbers in the previous pages, this movement involves two distinct steps: the reduction in one input (e.g., input Y) and the increase in the other input (e.g., input X). The decrease in the output resulting from a decrease in input Y can be stated as

$$-\text{MP}_Y \times \Delta Y$$

The increase in output resulting from an increase in input X can be stated as

$$\text{MP}_X \times \Delta X$$

Along the isoquant, the output level must be maintained at a constant level. Thus,

$$-\frac{\text{MP}_Y}{\Delta X} = \frac{\text{MP}_X}{\Delta Y}$$

Rearranging the terms in this equation and remembering that $\text{MRTS} = \Delta Y/\Delta X$ gives us

$$\text{MRTS} = \frac{\Delta Y}{\Delta X} = -\frac{\text{MP}_X}{\text{MP}_Y}$$

The Optimal Combination of Multiple Inputs

Earlier we state that the determination of the optimal combination of imperfectly substitutable inputs depends on both their relative prices and on the degree to which they can be substituted for one another. In

Table 6B.4 MP_X/MP_Y in Relation to MRTS (X for Y)

Combination	Q	Y	MP_X	X	MP_Y	MRTS (X for Y)	MP_X/MP_Y
A	52	6		2			
B	52	4	13	3	6.5	2	2
C	52	3	11	4	11	1	1
D	52	2	6	6	12	1/2	1/2

the previous section, you learn that the degree to which one input can be substituted for another is actually a reflection of the relationship between their marginal products. Therefore, the optimal combination of inputs depends on the relationship between the inputs' relative marginal products and their relative prices. In the case of two inputs, we can state this relationship mathematically as

$$\frac{MP_X}{MP_Y} = \frac{P_X}{P_Y} \tag{6B.2}$$

To demonstrate this relationship, we use **isocost** curves and combine them with the isoquant curves developed in the previous section. First, we rearrange Equation (6B.2) in the following way:

$$\frac{MP_X}{P_X} = \frac{MP_Y}{P_Y} \tag{6B.3}$$

In other words, two inputs are combined in the best possible way when the marginal product of the last unit of one input in relation to its price is just equal to the marginal product of the last unit of the other input in relation to its price.

Let us now return to the more formal economic analysis of optimal input combinations by explaining the optimality rule with the use of isoquants and isocost curves. Suppose $P_X = \$100$ and $P_Y = \$200$. Suppose further that a firm has a budget of \$1,000 to spend on inputs X and Y. At these prices and this expenditure limit, any of the combinations of X and Y in Table 6B.5 could be purchased.

Algebraically, the budget can be expressed as follows:

$$E = P_X \times X + P_Y \times Y \tag{6B.4}$$

Table 6B.5 Input Combinations for \$1,000 Budget

Combination	X	Y
A	0	5
B	2	4
C	4	3
D	6	2
E	8	1
F	10	0

where $E =$ Total budget allotment for inputs X and Y

$P_X =$ Price of X

$P_Y =$ Price of Y

$X =$ Quantity of input X

$Y =$ Quantity of input Y

In other words, the amount spent for X and Y is equal to the number of units of X multiplied by its price plus the number of units of Y multiplied by its price. In this case,

$$\$1,000 = \$100X + \$200Y \tag{6B.5}$$

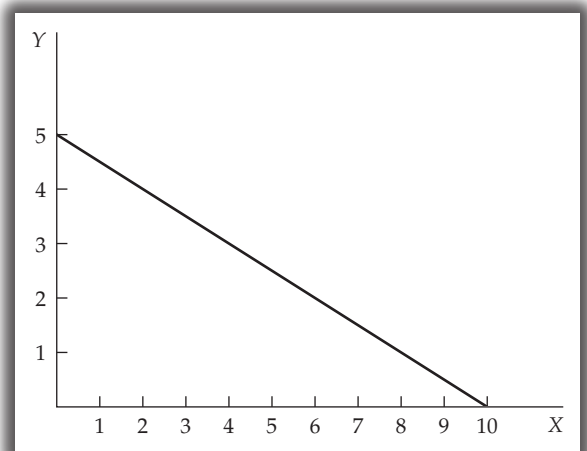
Using this equation to plot the numbers in Table 6B.5, we obtain the isocost curve shown in Figure 6B.6.

Note that the isocost curve is linear because the prices of the inputs are constant. A few algebraic manipulations of Equation (6B.4) indicate that the prices of the inputs in relation to each other (i.e., P_X/P_Y) are represented by the slope of the the isocost line:

$$\begin{aligned} E &= P_X \times X + P_Y \times Y \\ P_Y \times Y &= E - P_X \times X \\ Y &= \frac{E}{P_Y} - \frac{P_X}{P_Y} \times X \end{aligned} \tag{6B.6}$$

Figure 6B.7 combines the isocost curve shown in Figure 6B.6 and the isoquant shown in Figure 6B.5. Note that the isocost line and the isoquant are tangent to each other between points (4,3) and (6,2). This means that between these two points, the slopes of the two curves are identical. Therefore, if

Figure 6B.6 Isocost Curve for Inputs X and Y



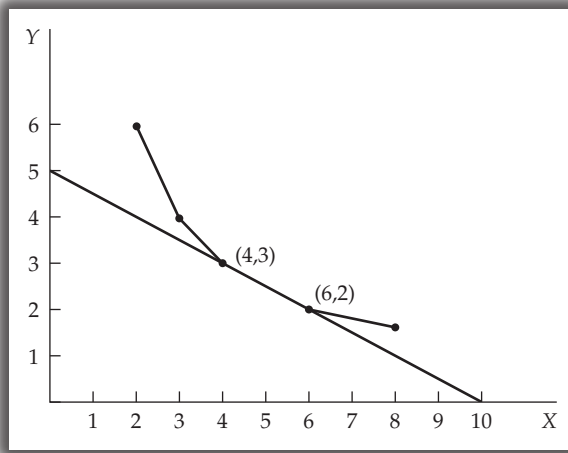


Figure 6B.7 Optimal Combinations of Inputs X and Y

the slope of the isocost line is $-P_X/P_Y$ and the slope of the isoquant is $-MP_X/MP_Y$, then between these two points

$$\frac{MP_X}{MP_Y} = \frac{P_X}{P_Y}$$

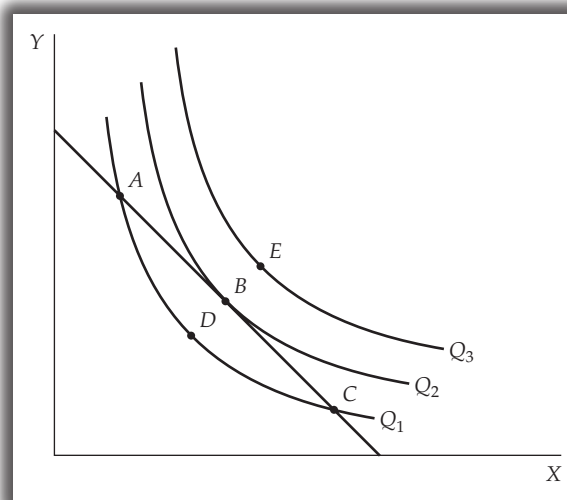
If we cancel the negative signs on both sides of the equation, we arrive at the optimality rule first stated in Equation (6B.2). Given a budget of \$1,000 and the input combinations represented by the isoquant in Figure 6B.7, the firm would employ the optimal combination of inputs if it used either four units of X and three of Y or six units of X and two of Y.

We are not able to find a unique combination of inputs because we have used a discrete set

of input combinations. Graphically, we can show quite easily how the use of a continuous production function enables us to find one optimal input combination. In Figure 6B.8, we have combined the isocost curve with a series of smoothed or continuous production isoquants. Point B represents the firm's optimal combination of inputs. Let us explain why.

To begin with, point D must be ruled out because at that point the firm would not be spending the full amount of its budget allotment. In contrast, point E represents a combination beyond the limits set by the budget. This leaves points A, B, and C, each representing a combination that can be purchased with the budget allotment. Of these three points, point B represents the best combination

Figure 6B.8 Selecting the Optimal Input Combination for a Continuous Production Function



because the firm would be producing the maximum amount given its budget limitation.

In terms of the marginal analysis developed previously, we can see that point *B* is the only one that fulfills the optimality condition expressed in Equations (6B.2) and (6B.3). That is, at point *B*, the slope of the isocost curve and the slope of the isoquant are identical. (Recall that the slope of a continuous isoquant is measured by the slope of the line tangent to the curve at a particular point.) Therefore, at point *B*, $MP_X/MP_Y = P_X/P_Y$.

The Optimal Levels of Multiple Inputs

By following the optimality condition first presented in Equations (6B.2) and (6B.3), a firm ensures it will be producing in the least costly way, regardless of the level of output. Therefore, Equation (6B.2) could be more specifically called the condition for cost minimization. But how much output should the firm be producing? The answer to the question, as was the case for a single input, depends on the demand for the product.

As you recall, the decision regarding the number of units of a single input to use was based on the condition $P_X = MFC_X = MRP_X$. That is, the firm should use the *X* input up to the point at which its cost (i.e., P_X) just equals the market value of the additional input's efforts (i.e., MRP_X). By the same token, the decision rule for two or more inputs requires that the firm use each input up to the point

$$P_i = MRP_i \quad (6B.7)$$

where P_i = Price of input *i*

MRP_i = Marginal revenue product of input *i*

If the firm is using two inputs (*X* and *Y*), its optimality condition is

$$P_X = MRP_X \text{ and } P_Y = MRP_Y \quad (6B.8)$$

We can explain the rationale for the optimality condition with two or more inputs simply by saying that what applies in the case of one input must apply to more than one input. However, there is a more formal explanation in microeconomic theory using several terms and concepts that are not presented in detail until Chapter 8. Nonetheless, following is a brief version of this theoretical explanation.

According to economic theory, the firm that wants to maximize its profit will always try to operate at the point where the extra revenue received from the sale of the last unit of output produced is just equal to the additional cost of producing this

output. In other words, its optimal level of production is at the point where marginal revenue (MR) is equal to marginal cost (MC). In Chapter 9, you learn in much greater detail about the rationale and application of the $MR = MC$ rule. For now, let us simply explain the justification for the rule governing the optimal use of more than one input.

Marginal cost, or MC, is the cost of producing an additional unit of output. Using the terms developed in previous examples in this chapter,

$$MC = \frac{P_i}{MP_i} \quad (6B.9)$$

where MC = Marginal cost of production

P_i = Price of the input *i* (i.e., the cost to the firm of using an additional unit of the input *i*)

MP_i = Marginal product of the input *i*

For example, suppose the input used is labor (measured in hourly units), and the price of labor is the wage rate given to the firm under perfectly competitive labor market conditions. Assume a wage rate of \$10 per hour. Also assume a particular hour of labor has a marginal product of 20 units of output. On a per-unit basis, these additional 20 units will cost the firm \$0.50 (i.e., \$10/20) to produce. In other words, at this point the marginal cost is \$0.50.

Let us assume the firm is operating at the profit-maximizing level of output—in other words, at the point where

$$MR = MC \quad (6B.10)$$

Let us also assume the firm employs two inputs, *X* and *Y*. Substituting Equation (6B.9) into (6B.10) gives

$$MR = \frac{P_X}{MP_X} \text{ for the } X \text{ input} \quad (6B.11)$$

$$MR = \frac{P_Y}{MP_Y} \text{ for the } Y \text{ input} \quad (6B.11)$$

Rearranging terms gives us

$$P_X = MR \times MP_X \text{ and } P_Y = MR \times MP_Y \quad (6B.12)$$

Because $MR \times MP_i = MRP_i$, the firm will be satisfying the optimality condition with that combination of *X* and *Y* at which

$$P_X = MRP_X \text{ and } P_Y = MRP_Y \quad (6B.13)$$

In short, the optimal level of multiple inputs occurs when the additional revenue that each input

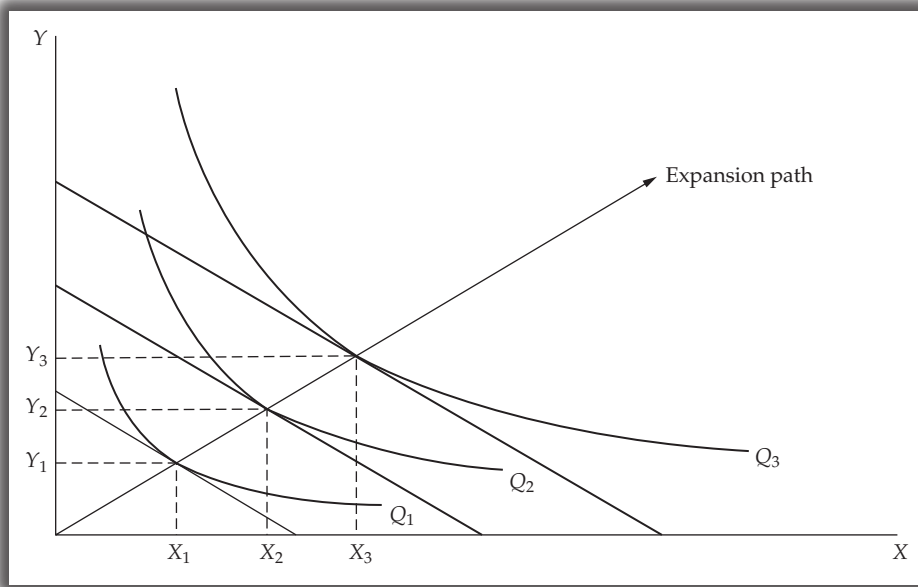


Figure 6B.9 Cost-Minimizing and Profit-Maximizing Input Combinations

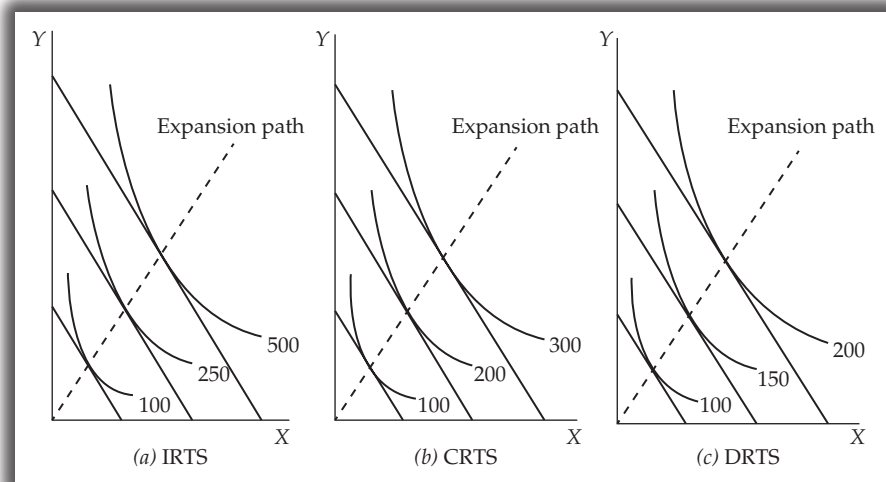
accounts for is just equal to the additional cost to the firm of using each input. Another way to view this optimality condition is to remember that it is actually derived from the assumption that the firm is already producing at the profit-maximizing level of output (i.e., where $MR = MC$). This, in turn, implies that the firm is combining its inputs in an optimal fashion. If not, then it could not possibly be maximizing its profit.

Figure 6B.9 illustrates the difference between the cost-minimizing and the profit-maximizing combinations of inputs. You can see that any one

of the points along the “expansion path” represents the cost-minimizing combination of inputs X and Y . However, suppose the $MR = MC$ rule for profit maximization dictates that a firm produce Q_2 units for sale in the competitive marketplace. As you can see, this implies that only one combination of inputs (X_3 and Y_3) should be used. All the other combinations of inputs will be cost efficient but will not enable the firm to maximize its profits.

Incidentally, when returns to scale are measured, economists always assume the firm is operating with the optimal combination of inputs. In Figure 6B.10,

Figure 6B.10 Optimal Input Combinations and Returns to Scale



we view returns to scale “from above” rather than “from the side” as we did back in Figure 6.5. In this figure, the different levels of output resulting from increases in input are found on a ray from the origin. (This ray is actually the locus of points indicating the optimal combination of inputs for different levels of

budgetary constraint.) As implied by the hypothetical numbers assigned to the isoquants, the values of these isoquants in relation to the values of the optimal input combinations indicate whether the firm is experiencing increasing, constant, or decreasing returns to scale.

Analyzing Production Functions with the Use of Calculus



In this chapter, we rely primarily on tables and graphs to illustrate our analysis of the production function. This appendix shows you how calculus can be used in the analysis.

A Brief Review of the Production Function

The production function expresses the relationship between the output and one or more inputs. The output is referred to as either *total product* (TP) or Q (quantity). We began our analysis of the production function by presenting the relationship between different combinations of two inputs (i.e., labor and capital) and output in a numerical table (see Table 6C.1). As already explained, the production function takes the general form:

$$Q = f(L, K) \quad (6C.1)$$

where Q = Quantity of output
 L = Labor (i.e., the variable input)
 K = Capital (i.e., the fixed input)

By stating this general functional form in more specific terms, we have an equation that can be used to generate a tabular model of the production function.

Two functional forms will be examined in this appendix, Cobb-Douglas and Cubic.

Cobb-Douglas Production

One popular form is the *Cobb-Douglas* production function. It can be stated as follows:

$$Q = AL^aK^b \quad (6C.2)$$

Q , L , and K have the same definitions as in Equation (6C.1). The values for A , a , and b determine the actual values in the production table. For example, suppose $A = 100$, $a = 0.5$, and $b = 0.5$. The production function would be

$$Q = 100L^{0.5}K^{0.5} \quad (6C.3)$$

This equation can be used to generate the numbers in Table 6C.1.

Marginal Product: The First Derivative of the Total Product Function

As we show in the discussion of isoquants, the marginal product expressed in terms of calculus is the partial derivative of the total product function. We can use the Cobb-Douglas form of the production

Table 6C.1 Cobb-Douglas Production Function

K	Output Quantity							
8	282	400	488	564	628	688	744	800
7	264	373	456	528	588	644	700	744
6	244	345	422	488	544	600	644	688
5	223	315	385	446	500	544	588	628
4	200	282	346	400	446	488	528	564
3	173	244	300	346	385	422	456	488
2	141	200	244	282	315	345	373	400
1	100	141	173	200	223	244	264	282
	$L1$	$L2$	$L3$	$L4$	$L5$	$L6$	$L7$	$L8$

function shown in Equation (6C.3) to illustrate how the derivative is used to find this marginal product.

Suppose we want to find the marginal product of the fourth unit of labor added to the production process, assuming we are currently using four units of capital. We simply take the partial derivative of the equation with respect to L , set $L = 4$ and $K = 4$, and find the resulting value of total product:

$$\begin{aligned} \frac{\partial Q}{\partial L} &= 100(0.5)L^{-0.5}K^{0.5} \\ &= 100(0.5)4^{-0.5}4^{0.5} \\ &= 50 \end{aligned}$$

In Table 6C.1 we note that the marginal product of the fourth unit of labor assuming 4 units of capital is 54 (i.e., $400 - 346$). The two values differ because the use of calculus actually enables us to compute marginal product right at the point at which four units of labor are being used, rather than that amount of output resulting from the addition of the fourth unit.

Converting the Cobb-Douglas Function into a Linear Form

1. The Cobb-Douglas function is nonlinear; it is an exponential function. However, it can be converted into a linear function in terms of logarithms:

$$\begin{aligned} Q &= aL^bK^c \\ \log Q &= \log a + b \log L + c \log K \end{aligned}$$

2. The original form of the function was

$$Q(L, K) = aL^bK^{1-b}$$

This model assumes constant returns to scale. In other words, if both labor and capital inputs are changed by a certain proportion s , Q will also change by s :

$$\begin{aligned} Q(sL, sK) &= a(sL)^b(sK)^{1-b} \\ &= a(s^bL^b)(s^{1-b}K^{1-b}) \\ &= (s^{b+1-b})Q(L, K) \\ &= s^1Q(L, K) \end{aligned}$$

Thus, the new quantity, $Q(sL, sK)$, will equal the old quantity, $Q(L, K)$, times s —the proportion by which both L and K , changed. This is the essence of CRTS.

The later version of the function relaxed the requirement of constant returns because it permitted $b + c$ to be less than, equal to, or greater than 1. In this case, the results would be as follows:

$$\begin{aligned} Q(sL, sK) &= a(sL)^b(sK)^c \\ &= a(s^bL^b)(s^cK^c) \\ &= a(s^{b+c})(L^bK^c) \\ &= (s^{b+c})Q(L, K) \end{aligned}$$

If $b + c > 1$, then for a proportional increase in inputs by $s > 1$, output will change by more than s (i.e., increasing returns to scale, IRTS). If $b + c < 1$, then the increase in output will be less than s (i.e., decreasing returns to scale, DRTS).

3. The marginal product of a factor is the partial derivative of quantity of output with respect to the factor:

$$\begin{aligned} MP_L &= \partial Q / \partial L \\ &= abL^{b-1}K^{1-c} \\ &= abL^bL^{-1}K^{1-c} \\ &= bL^{-1}Q \\ &= bQ/L \end{aligned}$$

Similarly, the marginal product of capital, MP_K , equals cQ/K or, in the CRTS case, $(1 - b)Q/K$.

4. The elasticity of production measures the sensitivity of total product to a change in an input in percentage terms:

$$E_Q = \frac{\text{Percent change in } Q}{\text{Percent change in input}}$$

or, in the case of labor,

$$\begin{aligned} \Delta Q/Q \div \Delta L/L &= \Delta Q/Q \times L/\Delta L = \Delta Q/\Delta L \times L/Q \\ &= \Delta Q/\Delta L \div Q/L \end{aligned}$$

$\Delta Q/\Delta L$ is, of course, the marginal product of labor; Q/L is the average product of labor (with capital held constant). Thus, the elasticity of production is equal to the marginal product divided by the average product. It was shown that the marginal product of labor is bQ/L and, as we have just pointed out, the average product of labor is Q/L . Dividing MP_L by AP_L ,

$$MP_L/AP_L = bQ/L \div Q/L = bQ/L \times L/Q = b$$

Thus, the elasticity of production for labor is b (and for capital, it is c). These are the original constant exponents of the Cobb-Douglas function. For any percentage increase (or decrease) in the quantity of a factor, holding the quantity of the other factor or factors the same, the increase (or decrease) in total product will be a constant percentage. Because the exponents are less than 1, the percent increase in total product is less than the increase in the quantity of the factor.

The Optimal Combination of Two Inputs

In the main body of this chapter, we show that if a firm is producing a level of output that maximizes its profit, then it must be using its inputs in such

a way that the marginal revenue product of every input used is equal to its price (or cost). In other words, if a firm uses k inputs, then

$$\begin{aligned} \text{MRP}_1 &= \text{Cost of input 1} \\ \text{MRP}_2 &= \text{Cost of input 2} \\ &\vdots \\ \text{MRP}_K &= \text{Cost of input } K \end{aligned}$$

Earlier we show why a rational firm would be using its inputs in the most cost-efficient manner if it combined them in such a way that the ratio of each input's marginal product relative to its price is equal for all inputs used:

$$\frac{\text{MP}_1}{P_1} = \frac{\text{MP}_2}{P_2} = \dots = \frac{\text{MP}_K}{P_K}$$

In this section, we employ calculus to demonstrate mathematically why this is so.

We begin by stating the profit function in the following manner:

$$\pi = \text{TR} - \text{TC} \tag{6C.4}$$

where π = Total profit
 TR = Total revenue
 TC = Total cost

By definition, total revenue is equal to price times quantity:

$$\text{TR} = P \times Q \tag{6C.5}$$

where P = Price of the output
 Q = Quantity of output sold

Total cost is equal to the amounts of the inputs used multiplied by their respective prices. Let us assume the firm is using two inputs, labor and capital, and their prices are the wage rate and some rental cost of using the capital. Thus, we can say that

$$\text{TC} = wL + rK \tag{6C.6}$$

where L = Labor
 K = Capital
 w = Wage rate of labor
 r = Rental cost of using capital

Substituting Equations (6C.6) and (6C.5) into Equation (6C.4) gives us

$$\pi = PQ - (wL + rK) \tag{6C.7}$$

As you know, the production function can be stated in general terms as

$$Q = f(L, K) \tag{6C.8}$$

Substituting this equation into Equation (6C.8) gives us

$$\pi = Pf(L, K) - wL - rK \tag{6C.9}$$

To find the level of input that will maximize the firm's profit, we can take the partial derivative of the profit function with respect to each input, L and K , while holding the other one constant and set each equal to zero:

$$\frac{\partial \pi}{\partial L} = Pf_L - w = 0 \tag{6C.10}$$

$$\frac{\partial \pi}{\partial K} = Pf_K - r = 0 \tag{6C.11}$$

If we express Equations (6C.10) and (6C.11) in terms of the prices of the inputs, we obtain

$$w = Pf_L \tag{6C.12}$$

$$r = Pf_K \tag{6C.13}$$

Recall that the definition of the marginal product of a particular input is the change in output with respect to a change in that input. In other words,

$$\text{MP}_L = \frac{\partial f}{\partial L} = f_L \tag{6C.14}$$

$$\text{MP}_K = \frac{\partial f}{\partial K} = f_K \tag{6C.15}$$

Recall further that the definition of the marginal revenue product (MRP) is the MP of a particular input multiplied by product price. Thus, Equations (6C.12) and (6C.13) are nothing more than a restatement of the conditions necessary for the optimal use of inputs discussed earlier in the main body of this chapter. That is,

$$Pf_L = \text{MRP}_L = w \tag{6C.16}$$

$$Pf_K = \text{MRP}_K = r \tag{6C.17}$$

Now that we have established the relationship, we can easily use the same equations and notations to show how the condition necessary for the most cost-efficient combination of inputs is derived. If we divide Equation (6C.12) by Equation (6C.13), we obtain

$$\frac{w}{r} = \frac{Pf_L}{Pf_K}$$

The P s in the denominator and the numerator on the right-hand side cancel out, giving

$$\frac{w}{r} = \frac{f_L}{f_K}$$

Using the definitions of marginal product stated in Equations (6C.14) and (6C.15) and rearranging the terms gives the condition for the most efficient combination of input usage:

$$\frac{MP_L}{w} = \frac{MP_K}{r} \quad (6C.18)$$

Or, by cross-multiplying, we have

$$MRTS = w/r$$

This simply says that efficient production requires tangency between the isoquant (whose slope is MRTS) and the isocost curve (whose slope is w/r).

Given the general form of Cobb-Douglas production, Equation (6C.18) simplifies to

$$bQ/(wL) = cQ/(rK)$$

By cross-multiplying, we obtain

$$rK = (c/b) \cdot wL$$

Substituting this into Equation (6C.6), we obtain

$$TC = wL + (c/b) \cdot wL$$

Regrouping and solving for wL , we obtain

$$wL = (b/(b + c))TC \quad (6C.19)$$

Substituting Equation (6C.19) into the rK equation and simplifying, we obtain

$$rK = (c/(b + c))TC \quad (6C.20)$$

Equations (6C.19) and (6C.20) say that efficient input usage will distribute resources to factors of production based on the relative size of the input elasticities.

Cubic Production



APP

The cubic production function allows varying returns to scale as well as varying marginal productivity; thus, it is a particularly useful functional form. Unfortunately, it is a bit cumbersome to work with algebraically. The function is defined by five parameters, a to e , all of which are non-negative. The function is:

$$Q(L, K) = aKL + bL^2K + cLK^2 - dL^3K - eLK^3 \quad (6C.21)$$

One point to initially note is that there are minus signs in front of the cubic terms, d and e . This means that eventually production will decline as either labor or capital increases with the other factor held fixed.

For fixed capital, the total product function is a cubic function of labor (because K , K^2 , and K^3 are all fixed in this instance). This is the functional form used to create the total product, and average and marginal product functions in Figures 6.6, 6.9, and 6.10. Average product of labor is obtained by dividing by L and the average product of capital is obtained by dividing by K :

$$AP_L(L; K) = aK + bLK + cK^2 - dL^2K - eK^3 \quad (6C.22a)$$

$$AP_K(K; L) = aL + bL^2 + cLK - dL^3 - eLK^2 \quad (6C.22b)$$

Note that each average product function is a quadratic function of the factor that varies (the cubic terms are cubic in the fixed factor of production).

Marginal products are similarly obtained by taking the partial derivatives of the production function with respect to labor and capital:

$$MP_L(L; K) = \partial Q/\partial L = aK + 2bLK + cK^2 - 3dL^2K - eK^3 \quad (6C.23a)$$

$$MP_K(K; L) = \partial Q/\partial K = aL + bL^2 + 2cLK - dL^3 - 3eLK^2 \quad (6C.23b)$$

These marginal products are reproduced in the Cubic Production 3D Excel App. These figures are based on $a = 0.3$, $b = 0.06$, $c = 0.005$, $d = 0.001$, and $e = 0.0005$. The short-run total product curve analyzed in the Excel files is based on $K = 10$. Given these values, $TP(L) = 3L + 0.6L^2 - 0.01L^3$, average product is $AP_L(L) = TP(L)/L = 3 + 0.6L - 0.01L^2$, and marginal product is $MP_L(L) = dTP(L)/dL = 3 + 1.2L - 0.03L^2$.

The economic region of production shown on the isoquant map on the 3D Graphs worksheet of the Cubic Production 3D Excel App is obtained from the marginal product equations. The green curve describes the set of bundles where $MP_L = 0$. This varies in a nonlinear fashion based on the amount of capital employed. The yellow curve describes the set of bundles where $MP_K = 0$. This varies in a nonlinear fashion based on the amount of labor employed. Both marginal products are zero at the top of the production hill.

Returns to Scale Given Cubic Production

Returns to scale issues are examined by varying both capital and labor. The Cobb-Douglas production function has a simple returns to scale result based on the size of the sum of exponents to the Cobb-Douglas production function. The cubic production function is more complex because it has varying returns to scale.

Returns to scale issues are examined in the Cubic Production 3D Excel App using a ray through the origin because this represents proportional change in all factors of production. The Excel file allows you to vary capital intensity k in the equation $K = kL$ over the range $0.4 \leq k \leq 2.5$. If kL replaces K in

Equation (6C.21), the resulting equation is a quartic function of labor. This function, upon regrouping terms, is:

$$Q(L) = akL^2 + (bk + ck^2)L^3 - (dk + ek^3)L^4 = fL^2 + gL^3 - hL^4 \tag{6C.24a}$$

Figure 6C.1 depicts this function for three values of k . The second version of this function describes the relation using $f = ak$; $g = (bk + ck^2)$; and $h = (dk + ek^3)$ in order to ease discussion. The average product along the ray is:

$$AP(L; K = kL) = fL + gL^2 - hL^3 \tag{6C.24b}$$

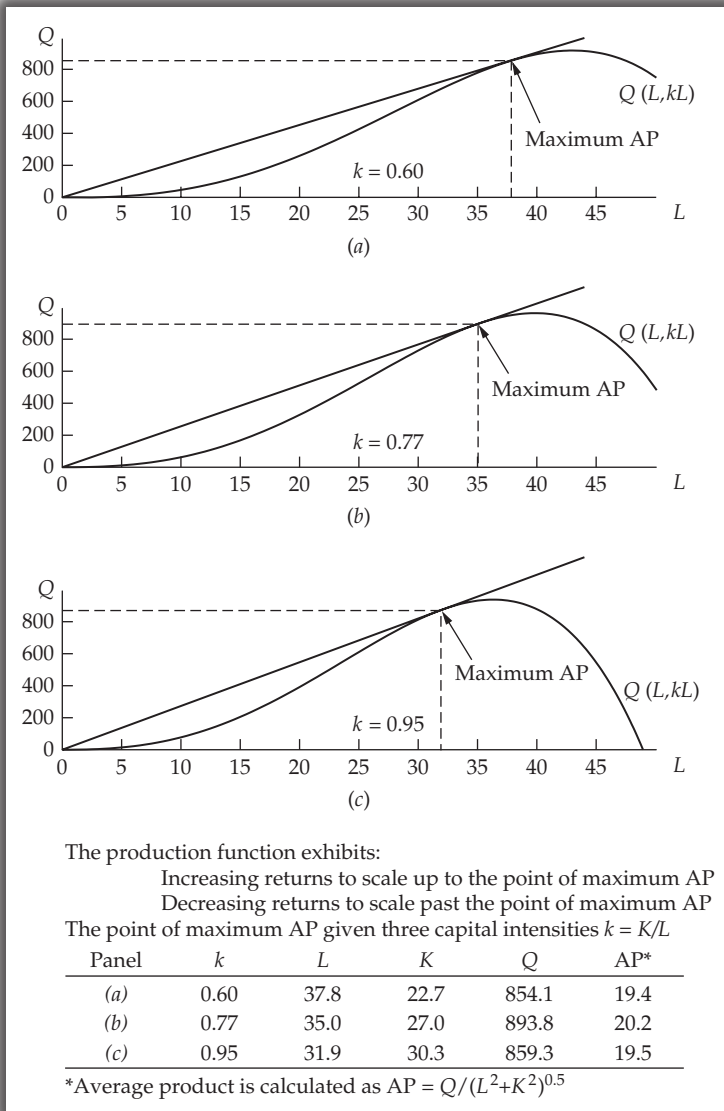


Figure 6C.1 Returns to Scale Given Cubic Production for Three Expansion Paths

The marginal product along the ray is:

$$MP(L; K = kL) = dQ(L)/dL = 2fL + 3gL^2 - 4hL^3 \tag{6C.24c}$$

Equations (6C.24b) and (6C.24c) are both metered in units of labor along the ray $K = kL$. To describe both as slopes along the ray, one should note that a one-unit increase in L means a movement of $(1 + k^2)^{0.5}$ along the ray. The steepness of the average product chord and the marginal product tangent in Figure 6C.1 are therefore AP and MP in Equations (6C.24b) and (6C.24c) divided by $(1 + k^2)^{0.5}$.

Average product along the ray is maximized when AP = MP. If Equation (6C.24b) is set equal to (6C.24c) and both sides are multiplied by $1/L$, we obtain:

$$f + 2gL - 3hL^2 = 0 \tag{6C.25}$$

This may be solved by using the quadratic formula. The labor value for which the cubic production function exhibits constant returns to scale given $K = kL$ occurs when:

$$L = (g + (g^2 + 3fh)^{0.5}) / (3h) \tag{6C.26}$$

If this is evaluated at $a = 0.3$, $b = 0.06$, $c = 0.005$, $d = 0.001$, and $e = 0.0005$ for $k = 0.60, 0.77$, and 0.95 , one obtains the unitary output elasticity points shown in each panel of Figure 6C.1. For example, along the ray $K = 0.95L$, the production function exhibits increasing returns up to $L = 31.9$ and decreasing returns thereafter, as seen in Excel Sheet 6C.1.

	A	B	C	D
1	0.3	<i>a</i>	0.285	$f = ak$
2	0.06	<i>b</i>	0.06151	$g = bk + ck^2$
3	0.005	<i>c</i>	0.00138	$h = dk + ek^3$
4	0.001	<i>d</i>		
5	0.0005	<i>e</i>	31.9043	L using Equation (6C.26)
6	0.95	<i>k</i>		
7	By varying value of capital intensity level <i>k</i> in cell A6, the resulting L value where CRTS occurs is determined in cell C5 using Equation (6C.26)			
8				
9				
10				

Excel Sheet 6C.1

Source: Cubic Production 3D Excel App, worksheet 6C.1

The Theory and Estimation of Cost

Learning Objectives

Upon completion of this chapter, readers should be able to:

- Define the cost function and explain the difference between a short-run and a long-run cost function.
- Explain the linkages between the production function and the cost function.
- Distinguish between economic cost and accounting cost.
- Explain how the concept of relevant cost is used in the economic analysis of cost.
- Define short-run total cost, short-run total variable cost, and total fixed cost and explain their relationship to each other.
- Define average cost, average variable cost, and average fixed cost and explain their relationship to each other in the short run. Do the same for average cost and average variable cost in the long run.
- Compare and contrast the short-run cost function and the long-run cost function and explain why economies of scale is considered to be a long-run phenomenon.
- Provide at least four reasons for the existence of economies of scale.
- (Appendix 7A) Use calculus to derive marginal cost from total cost and to find the minimums of each per-unit cost curve, and use Excel to depict the per-unit cost curves graphically.
- (Appendix 7B) Cite several studies that have estimated the cost functions in actual market situations.
- (Appendix 7B) Explain how to conduct empirical cost studies and cite several studies.

The Situation

Adam Michaels, the plant manager at Shayna Soda Company, was going through his daily routine of opening his mail on Monday morning when he came across a marketing flyer from Lawrence Aluminum Products, a company he had never heard of. The company was in New Jersey, about 400 miles away from his plant in upstate New York. At first he didn't pay too much attention to it, but then he saw that the base cost of soda cans was almost 30 percent lower than what he was now paying. He read on about the company and was impressed with its offer. The company offered an even

(continued)

(continued)

higher discount if you integrated with their technology and took advantage of electronic purchase orders. Shayna Soda Company required technological upgrades, so maybe this would be a perfect opportunity to switch both suppliers and computer systems. Adam was always under pressure from Corporate to increase the profitability of its soda production so his excitement on the potential of this change was high.

Adam called in Terry Roberts, the product availability manager, and told her of his discovery. “Terry, this sounds like a great opportunity for us and I want you to investigate this company a little more and let me know what we need to do to get supply from them.” Terry did not look excited about the project.

“But, Adam,” she replied, “what about Kayla Containers? We are their main customer and this could put them out of business.”

“Terry, this is a business decision and, although I like Kayla Containers, we need to stay in business, too. Please get on it, and give me a report by Friday.”

Terry went back to her office and picked up the phone to call Lawrence Aluminum Products. She was directed to Joseph Matthews, the director of sales. Terry told Joseph who Shayna Soda was and the type of cans that it would need. Lawrence had the right specifications on the can and actually had a little higher grade of quality in the aluminum. She continued to tell him their needs and asked how it would work to get Shayna’s logo and artwork on the cans. Joseph explained that they could do the logo and artwork with no problem, but Shayna would need to pay a one-time fee to set up the specifications. Moreover, the base price of the cans would go up 10 percent due to the extra coloring and production time. Terry quickly figured that even with the increase, the cost would be lower than Kayla Containers’ product.

“OK,” Terry continued. “What about delivery?”

“Delivery is no problem,” Joseph responded. “We have trucks that deliver in your area once a week. We would get your order in and put it on the next truck going out. And by the way, if you need to change the delivery date and request it in time, we can do it for you but we charge a bit more. If you want our base delivery fee, you must take delivery when we schedule it.”

“Is it a flat fee or by the mile?” Terry asked.

“It’s by the mile.”

Terry began to be concerned with the delivery aspect of the offer. “What about lead time? What do you require?”

“We have a pretty good turnaround time, so we only need five days,” Joseph replied.

“Five days!” Terry shot back. “That’s not great.”

“Well, to be honest, most of the time we can give that, but sometimes we may be able to meet your needs sooner, if we know your general plans in advance,” Joseph continued. “Ideally, we like to have a monthly forecast. In fact, we can offer you a discount if you integrate with our current technology for submitting forecasts and purchase orders and you make payment through the automated clearinghouse.”

What Joseph was saying now really began to pique Terry’s interest because the company was planning on improving its supply chain management anyway, particularly the procurement link of the chain.

“What would be the cost of that?” Terry asked.

“Well, you would need to purchase the software and the necessary support to set it up and integrate it with ours,” Joseph stated nonchalantly.

“You don’t help with any of the costs?” Terry asked.

“No. That would be your responsibility. But we can refer you to the contractors that helped us and offer any advice along the way. It would be of benefit to you in the long run, though. Just think of the efficiencies you could have.”

Joseph concluded the conversation by giving Terry all the numbers to do the necessary calculations. Terry spent the next couple of days figuring out all the aspects of the calculations.

THE IMPORTANCE OF COST IN MANAGERIAL DECISIONS

In the decade of the 1990s, the most commonly used way to contain or to cut costs was to reduce the number of people on a company’s payroll. *Restructuring*, *downsizing*, *rightsizing*, *redundancy*, and *force management* are all terms used in reference to this action. Over this past decade, these terms have taken on an additional dimension with the concept referred to as *shared services*. This typically refers to the consolidation of such “back office” and support functions as finance and accounting (e.g., purchasing and accounts payable), information technology (IT), and human resources (HR). This is

not a new concept. However, what is relatively new is where such consolidations or close tie-ins take place between the different “Centers of Excellence” (as shared service centers are often called) and the rest of the company. This is largely because of the proliferation and falling cost of broadband communications. Now a company can have its support function relating to headquarters as well as to all its other locations in the world where it operates.

Depending on the location, terms such as *onshore*, *near-shore*, and *offshore* are used to describe the situation. Companies recognize that they still require people to run their operations. So the idea is not to reduce labor as much as it is to optimize labor by reconfiguring where the labor being used is located and how it is organized. In addition to this particular trend in cost reduction, we have also noticed more companies trying to cut costs by reducing their assets (often referred to in the popular press as an “asset-light” strategy). We will provide examples of all these current efforts to reduce cost in the latter part of this chapter, after we discuss the economic theory of cost. (See pp. 276–279.)

An additional way in which companies have tried to cut costs to remain competitive is to merge and consolidate. Over the past several decades, there have been major mergers in industries such as airlines, pharmaceuticals, telecommunications, and financial services. Cost savings is often one of the major reasons cited for such mergers. For example, when Bank of America bought MBNA, one of the largest issuers of credit cards, this is what the press had to say:

With the acquisition, Bank of America, the nation’s third-largest bank, becomes its largest credit-card issuer, with \$143 billion in managed outstanding balances and 40 million active accounts. Bank of America said it will add more than 20 million new customer accounts as well as affinity relationships with more than 5,000 partner organizations and financial institutions.

*Bank of America said it expects to eliminate 6,000 jobs from the merged companies.*¹
[Emphasis added]

In 2004, the merger of Bank One and JP Morgan Chase made the headlines. In the pharmaceutical industry, Pfizer purchased Pharmacia (a company which only a few years before was the result of a merger between Upjohn and Pharmacia). Another example of a merger of companies which themselves were the result of previous mergers is Glaxo Wellcome and SmithKline Beecham. This company operates under the simplified name GSK. In telecommunications, there were some key mergers among the “Baby Bells” (the original seven operating companies resulting from the breakup of AT&T in 1984). Verizon is a result of the merger of Nynex and Bell Atlantic and GTE (the only major telecom company not part of the original AT&T). SBC is a result of the merger of SBC, Ameritech, and Pacific Telesis. In 2005, SBC acquired AT&T and took on the name AT&T. In that same year, Verizon acquired MCI. In 2006, AT&T acquired Bell South, thereby also consolidating the Cingular Wireless brand that the two companies had owned jointly. Soon afterward, Cingular was rebranded AT&T Wireless. The airline industry has also experienced many mergers and acquisitions. Two big mergers in recent times involved Delta Airlines’ takeover of Northwest Airlines in 2008 and United Airlines’ takeover of Continental in 2010. At the time the United-Continental merger was announced, company officials hinted at layoffs due to an “overlap of jobs” in certain locations and areas of work.²

¹www.consumeraffairs.com/news04/2005/bofa_mbna.htm.

²<http://www.csmonitor.com/Business/Latest-News-Wires/2010/0917/Continental-Airlines-UAL-to-merge>

As explained earlier, the economic analysis of cost begins with the production function. We therefore begin the analysis of cost by showing the links between production and cost functions. But before doing so, we review briefly the particular way in which cost is defined and used in economic analysis.

THE DEFINITION AND USE OF COST IN ECONOMIC ANALYSIS

In a typical business organization, cost is generally considered the domain of the accounting department. Its presentation to the outside world (e.g., to bankers, bond holders, investors, and suppliers) is based on generally accepted accounting principles (GAAP). For purposes of internal analysis and decision making, the definition of cost is based on the concept of relevancy. By definition, a cost is considered to be relevant if it is affected by a management decision. Any cost not affected by a decision is considered irrelevant. Both economists and managerial or cost accountants (as opposed to financial accountants) use the concept of **relevant cost** when analyzing business problems and recommending solutions. Following are some important ways to distinguish between relevant and irrelevant cost.

Historical Versus Replacement Cost

Suppose a manufacturer of a video game system has an inventory of \$750,000 worth of 16-bit chips left over from a discontinued system. Strong protectionist measures by Congress have created a shortage of these chips, driving their market value up to \$1,000,000. Meanwhile, the firm decides to reenter the video game market. (This time, it will manufacture the product in Thailand and will begin production with the leftover inventory of chips.) How much will it cost the firm to use this inventory? Although the **historical cost** is \$750,000, the replacement value is \$1,000,000. According to the principle of relevant cost, the firm should use the latter figure in computing its cost of reentering the video game market.³ Let us see why this is so.

If the firm were to decide not to proceed with the project but instead to sell its inventory of chips in the open market, it could receive the full market value of \$1,000,000. Therefore, by using the chips, it is forgoing the opportunity to receive \$1,000,000 for their sale. Moreover, if it decided to buy chips rather than use its inventory, it would have to pay \$1,000,000 for the same quantity that it holds in inventory. The \$1,000,000 is the relevant sum because it is the amount that has an impact on the alternatives being considered.

Opportunity Cost Versus Out-of-Pocket Cost

Previous discussions pointed out that **opportunity cost** is one of the most important and useful concepts in economic analysis because it highlights the consequences of making choices under conditions of scarcity. We can now use this term in a more specific way to help explain the concept of relevant cost. Opportunity cost, as you recall, is the amount or subjective value that is forgone in choosing one activity over the next best alternative. This type of cost can be contrasted with “out-of-pocket cost.” On occasion, economists refer to opportunity cost as *indirect cost* or *implicit cost*, and refer to out-of-pocket cost as *direct cost* or *explicit cost*.

³The recording of the cost of an asset for purposes of financial reporting is subject to GAAP, which state that assets and liabilities should be recorded in financial statements at historical cost. Inventories may be reported at historical cost or current market value, whichever is lower. But regardless of the rule set for external reporting, economists and cost accountants (as opposed to financial accountants) recommend that a firm use current market value for internal management decisions.

In the case of the company with the inventory of computer chips, we can clearly see that the opportunity cost of using the inventory in its second attempt to penetrate the video game market involves the cost of not being able to resell the inventory for \$1,000,000. Seen in this way, this sum is the firm's "relevant opportunity cost." The \$750,000 is not relevant because it is not the opportunity cost of going ahead with the project. Incidentally, the firm's out-of-pocket cost of using chips would be the cost of buying additional chips for the production process. For example, if the firm decides that it needs \$1,500,000 worth of chips (at current market prices) in the first year of the project's operation, we can infer that this figure consists of \$1,000,000 in opportunity cost and \$500,000 in additional out-of-pocket cost.

Sunk Versus Incremental Cost

Let us evaluate the cost to the firm of using its inventory of chips with a reversal in market conditions. Instead of the firm's inventory increasing in value, suppose something happens to cause its value to fall to \$550,000. For example, the introduction of a 32-bit video game system would drastically reduce the demand for and price of the 16-bit system. Under these circumstances, how much will it cost the firm to use the inventory for which it originally paid \$750,000? To answer this question, we employ the distinction that economists make between *incremental* and *sunk* costs. **Incremental cost** is the cost that varies with the range of options available in a decision. **Sunk cost** is the cost that does not vary in accordance with the decision alternatives. Our computer manufacturer has already paid \$750,000 for the chips and can really do nothing about the fact that changes in market conditions have driven the value of the chips down to \$550,000. If the firm decides to sell the inventory, it will receive at the most only \$550,000. If it decides to go ahead with the project, the incremental cost (i.e., the part of its cost that is affected by the decision) of using the inventory of chips will be \$550,000 and not \$750,000. And, as you have probably already concluded, the \$200,000 difference between these two sums must be considered a sunk cost to the firm. As it turns out, the \$550,000 can also be considered the opportunity cost of using the chips instead of selling them. Therefore, in summary, the firm should consider \$550,000 to be the relevant cost of using its inventory of chips because it is an *incremental opportunity cost*.

A dramatic example of the use of incremental and opportunity costs to determine relevant cost is the case of a new technology suddenly rendering the entire inventory of chips obsolete. For example, the introduction of a 64-bit game system would virtually destroy the market for 16-bit games. In this event, no one would want to buy the chips at any price, and the value of this inventory would be reduced to zero. The entire \$750,000 investment in inventory would be considered a sunk cost to the firm. Furthermore, because the resale value of the inventory is zero, the firm's use of the chips in its project would not incur an opportunity cost.⁴ In economic terms, the chips would represent a "free resource" because there would be no amount forgone with their use in the project. Would the firm be tempted to use the chips simply because of this? Would you want to build a new video game system with obsolete chips? If people will not buy the product, there would indeed be a high opportunity cost in terms of forgone sales.

⁴We avoid using the term *indirect cost* in reference to opportunity cost because some readers may confuse it with another definition. In manufacturing, *indirect cost* usually refers to the cost of using labor not directly involved in the making of a product (e.g., finance, personnel, research and development, and other support or staff functions).

THE RELATIONSHIP BETWEEN PRODUCTION AND COST

The economic analysis of cost is tightly bound to the economic analysis of production discussed in Chapter 6. As a matter of fact, one can say that the cost function used in economic analysis is simply the production function expressed in monetary rather than physical units. Furthermore, all the limiting assumptions used in specifying the short-run production function apply to the short-run cost function. The only additional assumption needed to determine the short-run economic cost function pertains to the prices of the inputs used in the production process. Here we assume the firm acts as a “price taker” in the input market; that is, it can hire or use as many or as few inputs as it desires, as long as it pays the going market price for them.

Table 7.1 presents an example of the numerical relationship between production and cost in the short run. The cost of using the variable input is determined by multiplying the number of units by the unit price. In this case, each unit of labor is assumed to be the equivalent of a 40-hour work week. The weekly wage rate is \$500. As indicated in the table, when the total product (Q) increases at an *increasing rate*, **total variable cost (TVC)** increases at a *decreasing rate*. When Q increases at a *decreasing rate*, TVC increases at an *increasing rate*. Plotting these numbers on a graph makes it quite apparent that total variable cost is a “mirror image” of total product (Figure 7.1).

In Chapter 6, the marginal product was defined as the change in total product divided by the change in the amount of the variable input used in the production process. Similarly, the rate of change in total variable cost is called **marginal cost**. Expressed in symbols,

$$MC = \Delta TVC / \Delta Q \text{ or } MC = \Delta TC / \Delta Q$$

Note that marginal cost is either the change in total variable cost or the change in total cost with respect to the change in output. This is because the **total fixed cost** component of **total cost** never changes as output increases. Using marginal cost and marginal product, we can restate the relationship shown in Figure 7.1 in this way: When the firm’s marginal product is increasing, its marginal cost of production is decreasing; when its marginal product is decreasing (i.e., *when the law of diminishing returns takes effect*), its marginal cost is increasing.

Table 7.1 Relationship Between Production and Cost, Short Run

Total Input (L)	Q	TVC (L × \$500)	MC (ΔTVC/ΔQ)	Reference Point in Figure 7.1
0	0	\$ 0		
1	1,000	500	\$ 0.50	A(A')
2	3,000	1,000	0.25	B(B')
3	6,000	1,500	0.16	C(C')
4	8,000	2,000	0.25	D(D')
5	9,000	2,500	0.50	E(E')
6	9,500	3,000	1.00	F(F')
7	9,850	3,500	1.42	G(G')
8	10,000	4,000	3.33	H(H')

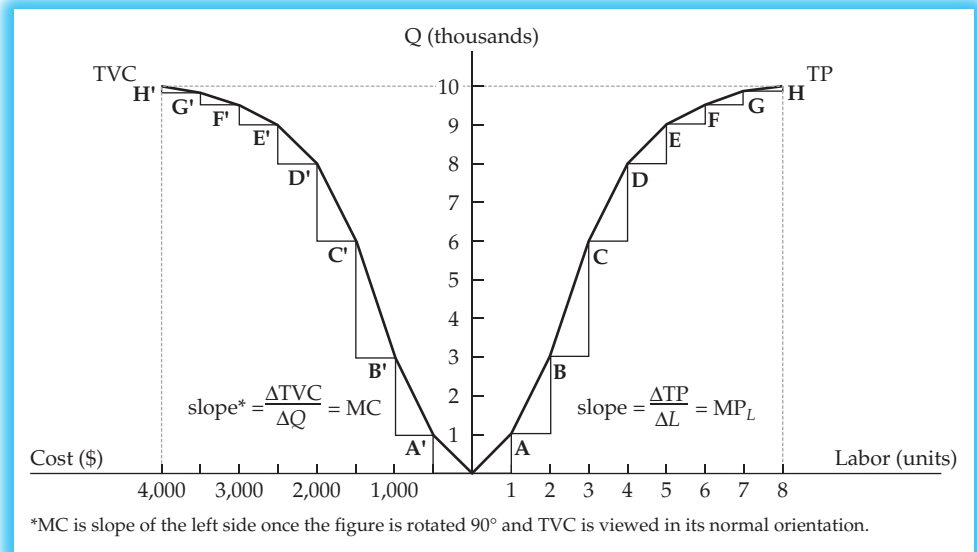


Figure 7.1 The Relation between Short-Run Production and Total Variable Cost

The relationship between diminishing returns and increasing marginal cost can also be illustrated algebraically. First, assume the variable input is labor (L), and its unit cost is some given wage rate (W). Now let us start by defining marginal cost as:

$$MC = \frac{\Delta TVC}{\Delta Q} \quad (7.1)$$

Because $TVC = L \times W$, we can say that

$$\Delta TVC = \Delta L \times W \quad (7.2)$$

Substituting Equation (7.2) into Equation (7.1) gives us

$$MC = \frac{\Delta L \times W}{\Delta Q} = \frac{\Delta L}{\Delta Q} \times W \quad (7.3)$$

Recalling the definition of MP, we know that $MP_L = \Delta Q / \Delta L$. Incorporating this observation into Equation (7.3) gives us

$$MC = \frac{1}{MP_L} \times W = \frac{W}{MP_L} \quad (7.4)$$

Clearly, Equation (7.4) tells us that, assuming a constant wage rate, MC will decrease when MP increases and will increase when MP decreases (i.e., when the law of diminishing returns takes effect).

In economic theory, the relationship between diminishing returns and marginal cost represents a key link between a firm's short-run production function and its short-run cost function because it is the law of diminishing returns that gives the short-run cost function its distinctive nonlinear form. Consequently, as seen in ensuing sections of this chapter, the firm's total cost, total variable cost, average cost, average variable cost, and marginal cost functions are all constructed in accordance with this nonlinearity.

THE SHORT-RUN COST FUNCTION

This section deals with the focal point of this chapter: the firm’s short-run cost function. A numerical model of the behavior of the firm’s short-run cost is shown in Table 7.2. Before commenting on each column in the table, let us review all the assumptions that economists make in specifying a model of this kind:

1. The firm employs two inputs, labor and capital.
2. The firm operates in a short-run production period. Labor is its variable input, and capital is its fixed input.
3. The firm uses the inputs to make a single product.
4. In producing the output, the firm operates at a given level of technology. (Recall that in our discussion of the short-run production function, we assumed the firm uses state-of-the-art technology in the production process. The same holds when we talk about short-run cost.)
5. The firm operates at every level of output in the most efficient way.
6. The firm operates in perfectly competitive input markets and must therefore pay for its inputs at some given market rate. In other words, it is a price taker in the input markets.
7. The firm’s underlying short-run production function is affected by the law of diminishing returns.

As we proceed in this chapter, you will see why these assumptions are crucial to the understanding of the short-run cost function.

The variables listed in Table 7.2 are defined as follows:

Quantity (Q): The amount of output that a firm can produce in the short run. (*Total product* is also used in reference to this amount.)

Total fixed cost (TFC): The total cost of using the fixed input *K*.

Total variable cost (TVC): The total cost of using the variable input *L*.

Total cost (TC): The total cost of using all the firm’s inputs (in this case, *L* and *K*).

Average fixed cost (AFC): The average or per-unit cost of using the fixed input *K*.

Average variable cost (AVC): The average or per-unit cost of using the variable input *L*.

Table 7.2 Total and Per-Unit Short-Run Cost

Quantity (Q)	Total Fixed Cost (TFC)	Total Variable Cost (TVC)	Total Cost (TC)	Average Fixed Cost (AFC)	Average Variable Cost (AVC)	Average Total Cost (AC)	Marginal Cost (MC)
0	\$100	\$ 0.00	\$ 100.00				
1	100	55.70	155.70	\$100.00	\$ 55.70	\$155.70	\$ 55.70
2	100	105.60	205.60	50.00	52.80	102.80	49.90
3	100	153.90	253.90	33.33	51.30	84.63	48.30
4	100	204.80	304.80	25.00	51.20	76.20	50.90
5	100	262.50	362.50	20.00	52.50	72.50	57.70
6	100	331.20	431.20	16.67	55.20	71.87	68.70
7	100	415.10	515.10	14.29	59.30	73.59	83.90
8	100	518.40	618.40	12.50	64.80	77.30	103.30
9	100	645.30	745.30	11.11	71.70	82.81	126.90
10	100	800.00	900.00	10.00	80.00	90.00	154.70
11	100	986.70	1,086.70	9.09	89.70	98.79	186.70
12	100	1,209.60	1,309.60	8.33	100.80	109.13	222.90

Average total cost (AC): The average or per-unit cost of using all the firm’s inputs.

Marginal cost (MC): The change in a firm’s total cost (or, for that matter, its total variable cost) resulting from a unit change in output.

The important relationships among these various measures of cost can be summarized as follows:

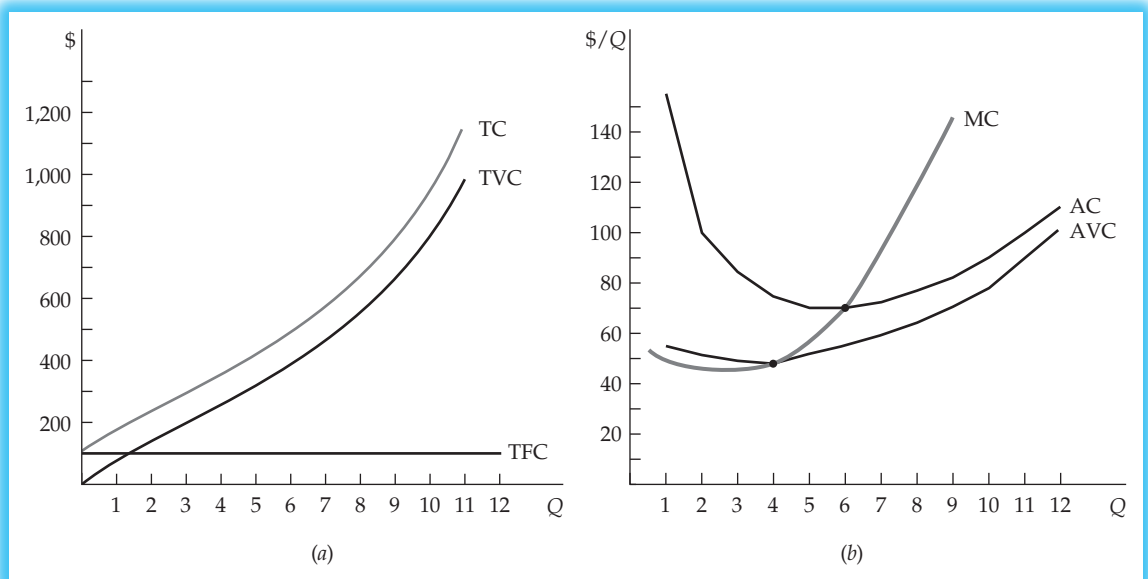
$$\begin{aligned}
 TC &= TFC + TVC \\
 AC &= AFC + AVC \text{ (or } = TC/Q) \\
 MC &= \Delta TC/\Delta Q \text{ (or } = \Delta TVC/\Delta Q) \\
 AFC &= TFC/Q \\
 AVC &= TVC/Q
 \end{aligned}$$

In evaluating the schedule of numbers in Table 7.2, note that as a matter of convenience, we have considered unit changes in output over the range of production being considered. Because ΔQ will always be equal to 1, we can quickly figure the marginal cost as output increases. For example, the marginal cost of the second unit of output is simply the change in the firm’s total cost (or total variable cost) between 1 unit and 2 units of production. In Table 7.2, we observe that this amount is \$49.90.

As output increases from 0 to 12, observe what happens to the various measures of cost. Total fixed cost, as expected, remains constant at \$100 over the range of output. Total variable cost increases at a decreasing rate, but when the fourth unit of output is produced, it starts to increase at an increasing rate. The same is true for total cost. When the numbers are plotted on a graph (see Figure 7.2a), the rate of change in total cost can be seen as the slope of the TC curve. The constancy of fixed cost is depicted by the horizontal line emanating from the corresponding point on the Y-axis.

As for the per-unit cost measures, we see that **average fixed cost (AFC)** declines steadily over the range of production. This is to be expected, because a constant sum of \$100 is being divided by the larger amounts of output. **Average variable cost (AVC)**

Figure 7.2 Total Cost, Total Variable Cost, Total Fixed Cost, Average Cost, Average Variable Cost, and Marginal Cost



declines, reaches a minimum at 4 units of output, and then starts to increase. **Average total cost (ATC or AC)** behaves in a similar fashion but reaches its minimum point at 6 units of output. Marginal cost declines and then starts to increase once the third unit of output is produced.

Notice in Table 7.2 that the values for marginal cost are placed between the output intervals, indicating that this measure of cost shows how total cost changes as a result of a unit change in quantity. For this same reason, marginal cost data are also plotted on a diagram between the output intervals. This particular way of showing marginal cost graphically can be seen in Figure 7.2*b*, along with the average total cost and average variable cost curves.

Figure 7.2*b* shows a particular relationship between marginal cost and the two other per-unit cost measures that is not as evident in Table 7.2. Notice that when marginal cost is equal to average variable cost, the latter measure is at its minimum point. (This occurs at 4 units of output.) When marginal cost is equal to average cost, average cost is at its minimum point. (This occurs at 6 units of output.) Another way to describe these relationships is to state that as long as marginal cost is below average variable cost, average variable cost declines as output increases. However, when marginal cost exceeds average variable cost, average variable cost starts to increase. The same relationship holds between marginal cost and average total cost. The economic significance of these relationships is explained in Chapter 9. But for now, it is important at least to note these relationships among the different per-unit measures of cost. Summarizing these relationships in abbreviated form:

One of three alternatives is always true: Marginal must be below, above, or equal to average

The relation between MC and AVC mirrors the one between MP_L and AP_L .
Cost depends on productivity and average depends on marginal

1 If $MC < AVC$, then AVC is falling.	This only occurs when $MP_L > AP_L$, then AP_L is rising.
2 If $MC > AVC$, then AVC is rising.	This only occurs when $MP_L < AP_L$, then AP_L is falling.
3 If $MC = AVC$, then AVC is flat (and at its minimum).	This only occurs when $MP_L = AP_L$, then AP_L is flat (and at its maximum).

This table describes two critical relationships.

- A. Productivity and cost are inversely related.
- B. In general, marginal pulls average up or down depending on if it is above or below average.

It is important to understand how the concept of relevant cost can be incorporated into the analysis of short-run cost. Suppose a firm is currently producing 6 units of output per period and is considering increasing this amount to 7. In deciding whether to produce the seventh unit of output, the relevant cost is the marginal cost. In other words, it is the *change* in total cost and not the total cost itself that must be considered. This is because whether the firm produces 6 or 7 units of output per time period, it still must pay the same amount of fixed cost. By evaluating the change in total cost, the firm automatically excludes fixed cost from consideration.

Increasing Cost Efficiency in the Short Run

As we shall see in Chapters 8 through 10, the short-run cost function plays a central part in the economic analysis of production and pricing. For now, we can appreciate the value of this model by considering the ways in which a company might attempt to become more economically efficient. The assumptions listed when the model was specified provide a convenient guide.

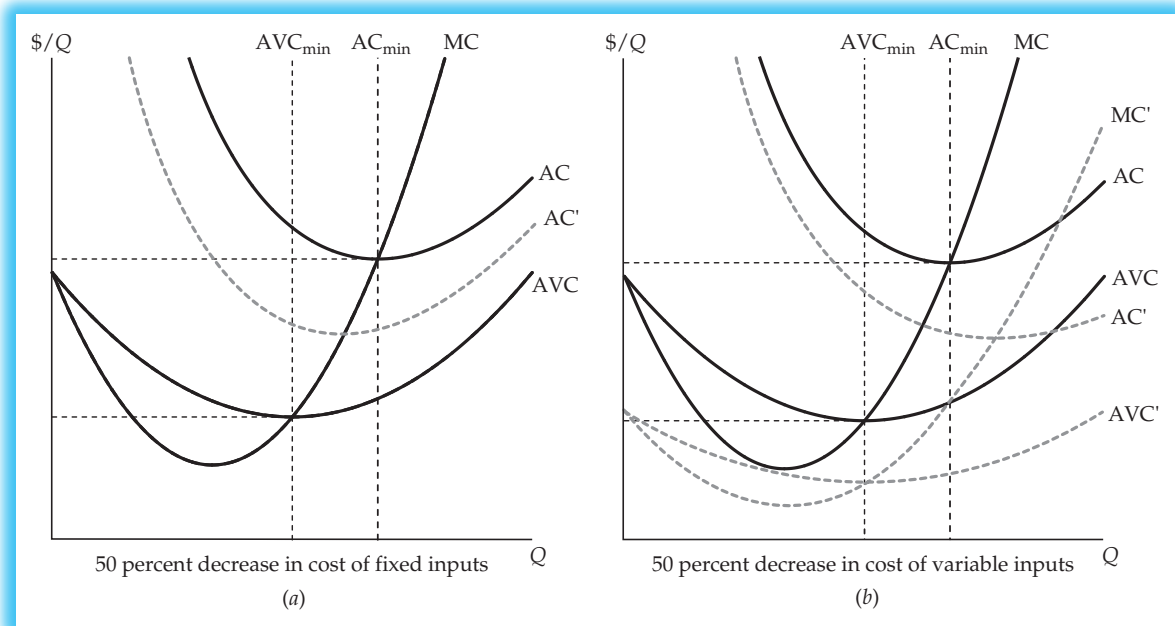


Figure 7.3 Effects on Short-Run Cost Structure of Price Changes in Fixed and Variable Inputs



To start with, the model assumes the firm is already operating as efficiently as possible. If we assume the firm is in fact operating as best it can with state-of-the-art technology, then the only possibility to reduce cost in the short run is for the inputs to decrease in price. If this were to happen, there would be a downward shift in the firm’s short-run cost curves. This effect is shown in Figure 7.3*a* and *b*. Notice that a reduction in the firm’s fixed cost (e.g., a reduction in rental payments) would simply cause the average cost curve to shift downward, whereas a reduction in the firm’s variable cost (e.g., a reduction in wage rates or raw materials costs) would cause all three cost curves—AC, AVC, and MC—to shift. Minimum MC' and AVC' occur at the same output levels as MC and AVC , but minimum AC' occurs at a larger output level than AC in this instance.

The purchase of capital equipment was not considered in the discussion because we assumed it to be our model’s fixed input. Actually, the addition of capital equipment could be considered a short-run change *if we assume labor to be the fixed input*. For example, a work force of a certain size might be given additional machinery with which to work. The point is that in the short-run analysis of cost, at least one of the inputs must be held constant.



Module 8A Alternative Specifications of the Total Cost Function

In economic analysis, the most common form of the short-run cost function is the one that has been used thus far in this chapter. That is, the **total cost function** is specified as a cubic relationship between total cost and output. As output increases, total cost first increases at a decreasing rate, and then at some point it increases at an increasing rate. By now you should be well aware that this is due to the underlying relationship between the firm’s variable input and the resulting output. Total cost increases at a decreasing rate because the firm is experiencing increasing returns to its variable input. When the law of diminishing returns takes effect, the firm begins to experience decreasing returns to its variable factor, thereby causing its total cost to begin increasing at an increasing rate.

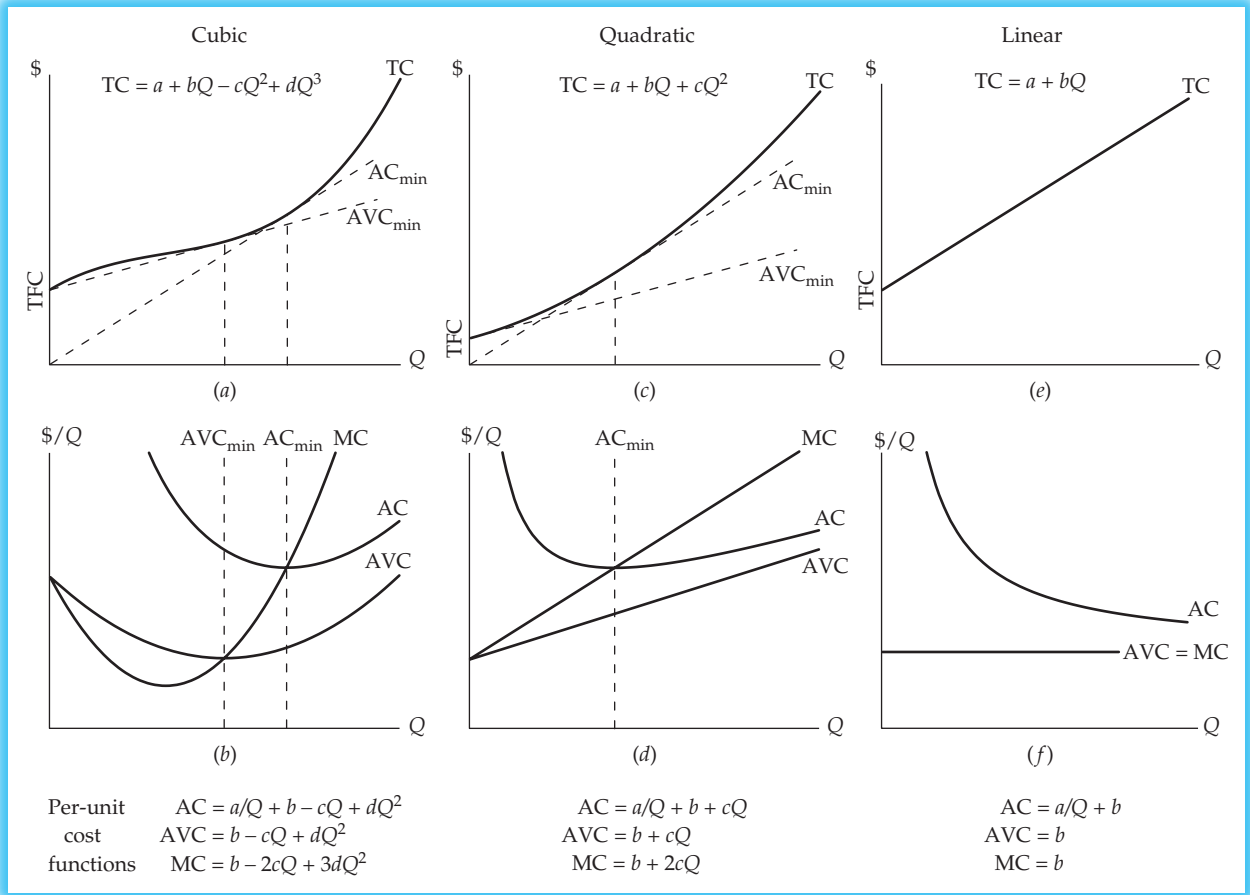


Figure 7.4 Alternative Representations of Total, Average, and Marginal Costs

In addition to this cubic form of the cost function, two other important functional relationships between total cost and output are considered in economic analysis. One is a quadratic relationship, and the other is a linear one. These relationships along with the cubic cost function are shown in Figure 7.4. This figure also shows the general algebraic expressions for the three forms of total cost. These variations can be explained once again by the underlying relationship between the firm’s variable input and the resulting output. In Figure 7.4c, we can see that the quadratic total cost function increases at an increasing rate from the outset of production. This implies that the law of diminishing returns takes effect as soon as the firm starts to produce. In contrast, the linear cost function, shown in Figure 7.4e, indicates that total cost increases at a constant rate. Think about this for a moment. What does this mean in terms of the underlying relationship between the firm’s variable input and its output? It means that the firm is experiencing neither increasing nor diminishing returns as its variable input is added to its fixed input. Each additional unit of variable input adds the *same* amount of additional output (marginal product) throughout the short-run range of production. Therefore, the change in the total cost relative to the change in output (which, as you recall, is indicated by the slope of the total cost line) is the same throughout the range of output considered.

Rather than discuss the different specifications of the firm’s short-run cost function in terms of total cost, it might be easier and more meaningful to discuss the alternative forms in terms of unit costs, both average and marginal. Thus, instead of

referring to total cost “increasing at an increasing rate,” we can simply say that the firm’s marginal cost is increasing. An increase in total cost at a constant rate means a constant marginal cost. We can also use Figure 7.4 to illustrate this point.

In Figure 7.4*b*, we see the familiar set of short-run cost curves depicted earlier. As output increases, marginal cost falls, reaches a minimum, and then starts to increase (as the law of diminishing returns takes effect). As it increases, marginal cost intersects with average variable cost and average total cost at their respective minimum points, for reasons explained earlier. In Figure 7.4*d*, we see that marginal cost increases as soon as production begins, just as we noted when evaluating the slope of the total cost function. As in the case of Figure 7.4*b*, marginal cost intersects average variable cost and average cost at their minimum points.

In the case of the linear cost function, shown in Figure 7.4*f*, the horizontal marginal cost line denotes that marginal cost remains constant as output increases. But also witness that marginal cost is equal to average variable cost, unlike the cases in Figures 7.4*b* and 7.4*d*. To illustrate this relationship, consider the following numerical example. Suppose you are given the following cost function:

$$TC = 100 + 0.50Q \quad (7.5)$$

Assuming the intercept and slope coefficients represent dollar units, this function tells us that the firm’s total fixed cost is \$100 and its marginal cost is \$0.50. That is, each additional unit of production adds \$0.50 to the firm’s total cost. Omitting the fixed cost component of the equation gives us

$$TVC = 0.50Q \quad (7.6)$$

Recall by definition that $AVC = TVC/Q$. Dividing Equation (7.6) by Q gives an AVC of \$0.50, which is the same as the marginal cost.

The mathematics of this example may seem rather trivial. Nonetheless, it is important to elaborate on this relationship between average variable cost and marginal cost because these two measures of cost are often used interchangeably in the business world. For example, when cost accountants use the term *standard variable cost*, they are usually referring to *both* marginal cost and average variable cost. In so doing, they are assuming the total cost function is linear, that the firm experiences neither increasing nor diminishing returns in the short run. The importance of these assumptions as well as the role of the linear total cost function in economic analysis are discussed further in Appendixes 7B and 8B.

THE LONG-RUN COST FUNCTION

The Relationship Between Long-Run Production and Long-Run Cost

In the long run, all inputs to a firm’s production function may be changed. Because there are no fixed inputs, there are no fixed costs. Consequently, all costs of production are variable in the long run. In most work situations, managers of firms make decisions about production and cost that economic theory would consider short run in nature. For example, they might have to decide how many labor hours are required for a particular project or whether the existing work force requires more machinery to meet increased demand.⁵ But from time to time, managers must make long-run

⁵Recall that we stated earlier that the “short run” in economic theory requires only that at least one of the inputs in the production function be held constant. Thus, we can treat the addition of capital equipment as a short-run change if the amount of factory space or the amount of labor, for example, remains unchanged.

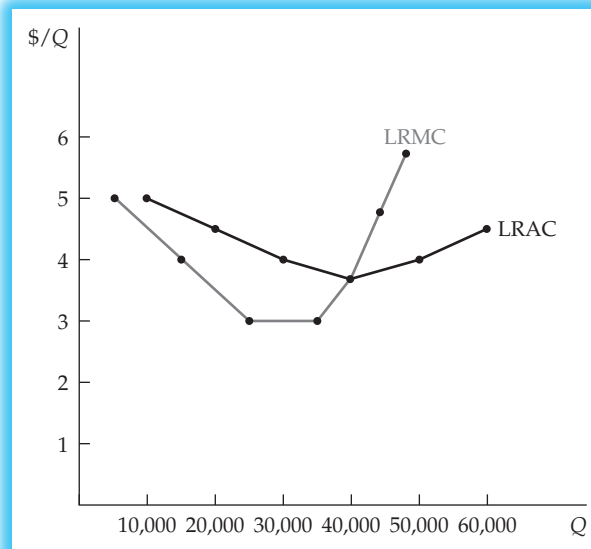
Table 7.3 Long-Run Cost Function

Scale of Production (Capacity Level)	Total Product (Output/Mo)	Long-Run Total Cost (LRTC)	Long-Run Marginal Cost (LRMC)	Long-Run Average Cost (LRAC)
A	10,000	\$ 50,000	\$5.00	\$5.00
B	20,000	90,000	4.00	4.50
C	30,000	120,000	3.00	4.00
D	40,000	150,000	3.00	3.75
E	50,000	200,000	5.00	4.00
F	60,000	260,000	6.00	4.33

production decisions. That is, they must consider possible changes in *all* the firm's inputs and, hence, changes in all the firm's cost of operation. Decisions of this kind are considered by economists to be part of a manager's planning horizon.

In explaining the nature of the long-run cost function, we begin with a schedule of numbers showing a firm's long-run cost function in relation to its long-run production function. Unlike the case of short-run cost presented in Table 7.2, the hypothetical numbers in Table 7.3 are based on the assumption that greater amounts of output are the result of increases in *all* the firm's inputs. Consistent with this assumption, this table implies that the firm would incur no cost if it chose not to produce any output, because in the long run there is no fixed cost.

As output increases, observe that total cost increases, but not at a constant rate. As with the short-run function, the rate of change of the long-run total cost function is called the marginal cost (long-run marginal cost, to be more precise). In looking at the long-run marginal cost column in Table 7.3, we see that this measure at first decreases, then is constant, and finally increases over the range of the output. The numbers in Table 7.3 are plotted on a graph in Figure 7.5. The rate of change in the

Figure 7.5 Long-Run Average and Marginal Cost

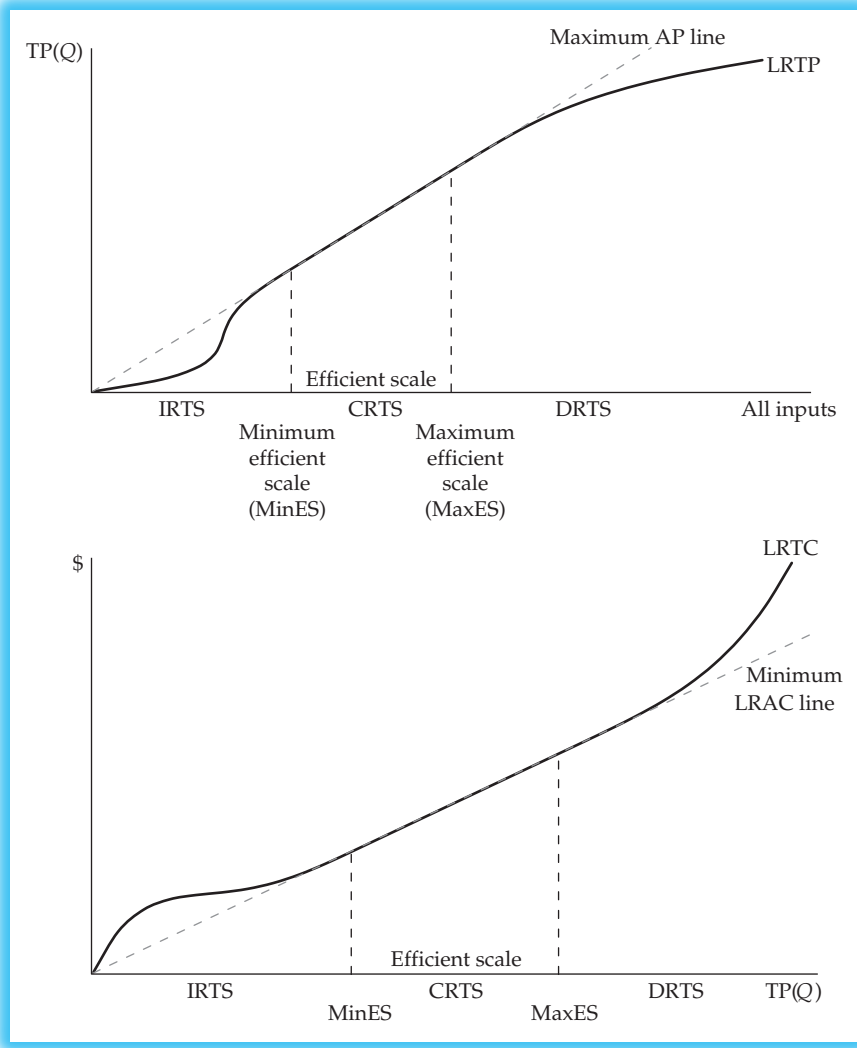


Figure 7.6 Returns to Scale for Long-Run Total Production and Long-Run Total Cost

long-run total cost can be seen by observing the slope of this curve, as well as by observing the behavior of the long-run marginal cost curve.

The reason for this particular behavior of the firm’s long-run marginal cost (or the rate of change in its long-run total cost) pertains to returns to scale. As explained in Chapter 6, economists hypothesize that a firm’s long-run production function may at first exhibit increasing returns, then constant returns, and finally decreasing returns to scale. This being the case, we would expect a firm’s long-run cost to change in a reciprocal fashion.

When a firm experiences increasing returns to scale, an increase in all its inputs by some proportion results in an increase in its output by some *greater* proportion. Assuming constant input prices over time, this means that if the firm’s output increases by some percentage, its total cost of production increases by some *lesser* percentage. The graphs in Figure 7.6 illustrate the reciprocal behavior of long-run cost and long-run production.

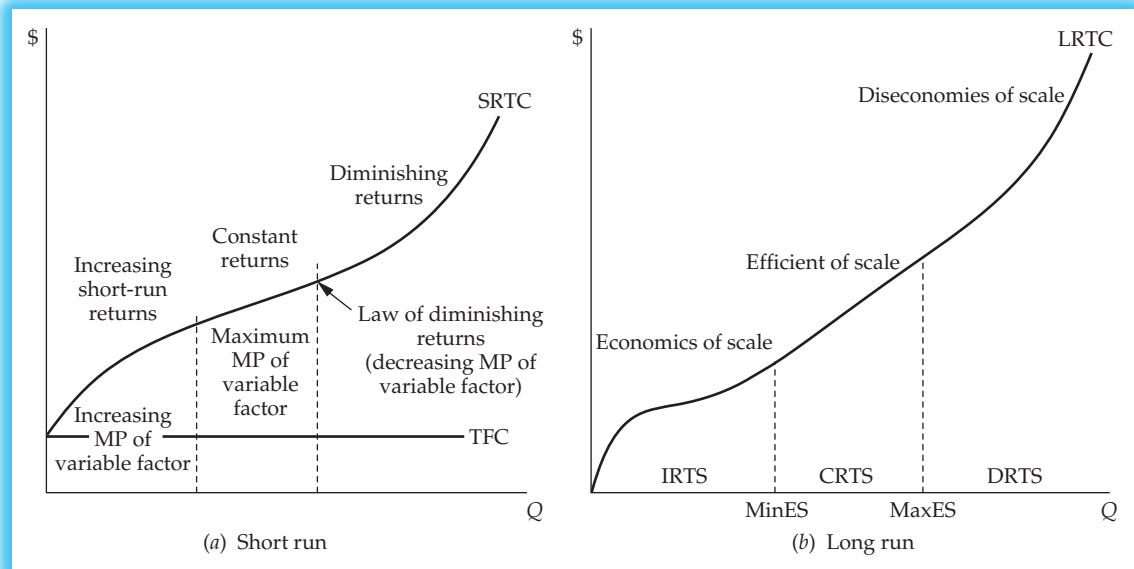


Figure 7.7 Short-Run Versus Long-Run Cost Functions

It is important to stress that although the long-run cost function appears to exhibit the same pattern of behavior as the short-run cost function, the reasons for their respective patterns are entirely unrelated. The short-run cost function is affected by increasing and diminishing returns to individual factors, a phenomenon that is assumed to take effect when at least one of the inputs is held constant, and the long-run function is affected by increasing and decreasing returns to scale, a phenomenon assumed to take effect when all the firm's inputs are allowed to vary. Figure 7.7 serves as a reminder of this distinction.

Economies of Scale

One of the measures of cost in Table 7.3 has yet to be discussed: long-run average cost. This variable is the key indicator of a phenomenon called **economies of scale** (or increasing returns to scale [IRTS]). If a firm's long-run average cost declines as output increases, the firm is said to be experiencing economies of scale. If long-run average cost increases as output increases, economists consider this to be a sign of **diseconomies of scale** (or decreasing returns to scale [DRTS]). There is no special term to describe the situation in which a firm's long-run average cost remains constant as output increases or decreases. We say that such a firm experiences efficient scale (or constant returns to scale [CRTS]) over a range of outputs. The smallest output where minimum LRAC is achieved is called **minimum efficient scale**. This is often shortened to MES or MinES to distinguish it from the largest output level for which minimum LRAC is achieved, called **maximum efficient scale** (MaxES). Figure 7.8 illustrates a typical U-shaped average cost curve reflecting the different types of scale economies that a firm might experience in the long run.

The primary reason for long-run scale economies is the underlying pattern of returns to scale in the firm's long-run production function. Further evaluation of Table 7.3 indicates that as long as marginal cost is falling, it is less than long-run average cost and in effect pulls the average down—a sure sign of economies of scale. However, once the firm starts to experience decreasing returns to scale, its long-run

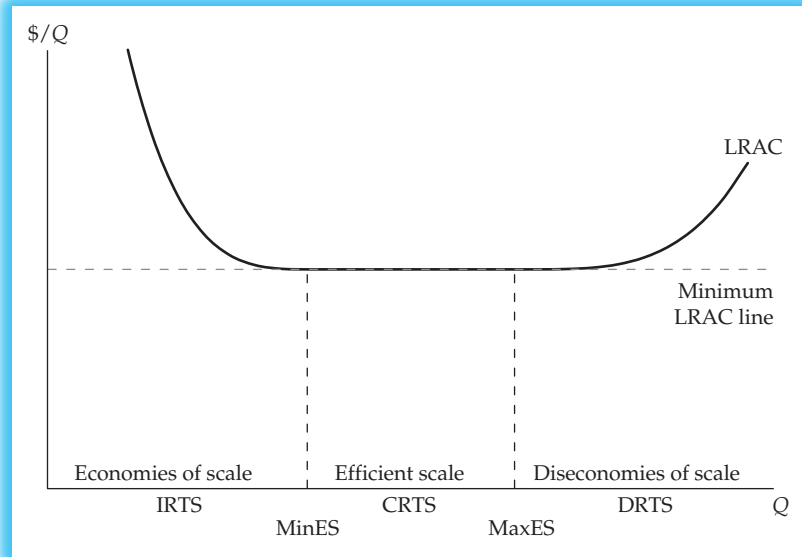


Figure 7.8 Long-Run Average Cost

marginal cost begins to rise. Eventually it becomes greater than long-run average cost, causing LRAC to rise and signifying diseconomies of scale.⁶

Scale economies and diseconomies may result from factors besides those relating to returns to scale. These other factors pertain primarily to the prices of the firm's inputs. For example, as the firm's scale of production increases, it may be able to exert some market power over its suppliers and thus receive volume discounts for bulk purchases of raw materials and component parts. Another example involves a firm's use of capital equipment with better *price–performance ratios*. As the firm increases its scale of production, it may be worthwhile to buy more cost-effective machinery whose price and capacity could not be justified at smaller scales of production. Only a few years ago, a typical illustration cited by economists of this type of scale economy was the computer. But with the advent of personal and midsize computers, there are now machines with prices and levels of computing power that are appropriate for firms of all sizes. Furthermore, with the tremendous advances in technology and software, the prices of the machines relative to their computing power (i.e., their *price–performance ratios*) do not differ much among the various sizes of computers.

We should briefly mention two other factors contributing to economies of scale. First, larger firms may be able to raise funds in capital markets at a lower cost than smaller ones. For instance, a large company may be able to secure short-term funds in the commercial paper market and long-term funds in the corporate bond market, whereas a small company may be able to secure borrowed funds only from a bank. Generally, the interest rates that firms must pay for funds in money and capital markets are less than for bank loans with comparable maturities. Second, a large firm may be able to take advantage of economies resulting from the spreading out of promotional costs. If a firm expands its scale of production, it may not have to expand

⁶You should recognize the relationship between LRMC and LRAC to be the same mathematically as the relationship between SRMC and SRAC. That is, regardless of whether a short- or a long-run time period is assumed, when the marginal is below the average, it brings the average down. When the marginal is above the average, it pulls the average up.

Table 7.4 Factors Affecting Economies and Diseconomies of Scale

Possible Reasons for Economies of Scale
Specialization in the use of labor and capital*
Indivisible nature of many types of capital equipment*
Productive capacity of capital equipment rises faster than purchase price
Economies in maintaining inventory of replacement parts and maintenance personnel*
Discounts from bulk purchases
Lower cost of raising capital funds
Spreading of promotional and research and development costs
Management efficiencies (line and staff)*
Possible Reasons for Diseconomies of Scale
Disproportionate rise in transportation costs
Input market imperfections (e.g., wage rates driven up)
Management coordination and control problems*
Disproportionate rise in staff and indirect labor*

*Indicates reason directly related to economies or diseconomies of scale in the long-run production function.

its advertising budget by the same proportion, if at all. The same can be said about research and development expenditures.

As far as diseconomies of scale are concerned, if the firm's scale of production becomes so large that it begins to substantially affect the total market demand for its inputs, it may start to increase the price of these inputs. A typical case is the expansion of a major employer in a local area with a relatively fixed supply of labor. If the firm's higher scale of production sufficiently increases its demand for labor, it could begin to drive up local wage rates. Economists also believe that, past some point, most firms exhibit bureaucratic diseconomies of scale related to coordination and control problems.

Another factor not related to the long-run production function that could cause diseconomies of scale is a firm's transportation costs. As a firm increases the production capacity of a particular manufacturing facility, per-unit transportation cost tends to rise rather than to fall. This is largely because transportation costs involve more than just the delivery of goods from one point to another. In addition, there are handling expenses, insurance and security expenses, and inventory costs (as goods await shipment). Increases in these types of expenses help to increase the total transportation cost to the extent that average transportation cost also increases. Furthermore, basic delivery expenses may rise at a faster rate than other kinds of cost if the firm has to ship the additional output to farther destinations. Economists hypothesize that eventually the increase in unit transportation cost will more than offset the fall in unit cost due to economies of scale. If this happens, then diseconomies of scale (i.e., rising average total cost) will result. Table 7.4 summarizes the key reasons for economies and diseconomies of scale. Factors that primarily relate to returns to scale are noted with an asterisk.

The Long-Run Average Cost Curve as the Envelope of Short-Run Average Cost

Up to now, we have been discussing long-run average cost as part of the firm's planning horizon. That is, the firm is assumed to be free to choose any level of capacity that it wants, because in our theoretical long-run time period, all inputs may vary. However, once a firm commits itself to a certain level of capacity, it must consider at least one of the inputs fixed as it varies the rest. In terms of production cost, this means that once a

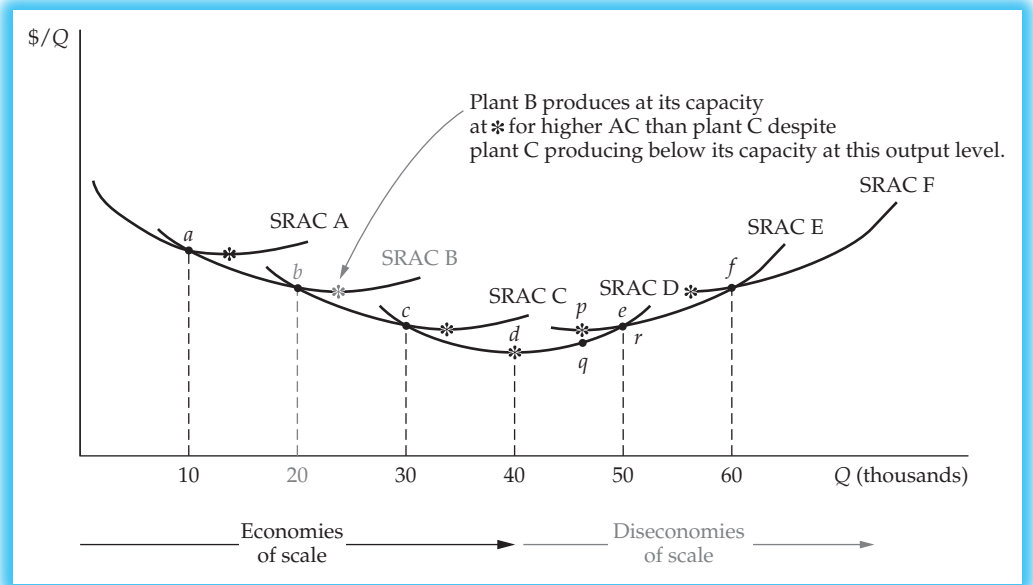


Figure 7.9 Capacity Level and Short-Run Average Cost

capacity level is decided on, the firm must work with a short-run cost function. We can illustrate this by returning to the example in Table 7.3.

Suppose the capacity levels shown in the table represent plants of increasing size. Figure 7.9 shows these capacity levels in relation to each plant's short-run average cost curve. The points labeled *a* through *f* represent the levels of output and average cost shown in Table 7.3. The lines passing through the points indicate the short-run average cost curves that we imagine to exist once the firm has locked into one of the plant sizes represented by the labeled points.

As expected, short-run average cost (SRAC) curves for the larger plants are positioned to the right of the curves for smaller ones, indicating greater production capacity. For example, SRAC B is to the right of SRAC A because plant B is larger than plant A. But, as noted in Figure 7.9, plants with larger capacities are greatly influenced by economies and diseconomies of scale. Because of the impact of economies of scale, plant B's SRAC curve is positioned below and to the right of plant A's, so the minimum point of B's SRAC curve is *lower* than that for A. The same can be said for plant C's minimum SRAC in relation to plant B's and for D's in relation to C's. However, because of the impact of diseconomies of scale, plant E's SRAC curve is positioned above and to the right of D's, and plant F's is above and to the right of E's. That is, the SRAC curves of plants larger than plant D have progressively higher minimum average cost points. For reference, the minimum SRAC points for all the plants are marked in Figure 7.9 with asterisks.

Another important aspect of Figure 7.9 is that none of the labeled dots is at the lowest point of its respective short-run average cost curve *except for the dot representing plant D*. For example, the asterisk marking plant B's minimum short-run average cost depicts a level above the average cost that would be incurred by plant C in the short run for a comparable level of production (see arrow in Figure 7.9). A logical extension of this illustration is the observation that if a firm wants to produce between 20,000 and 30,000 units of output per month, it would be better off using the manufacturing capacity provided by plant C than to try to increase the usage of the smaller plant B.

To understand the full economic implication of this observation, it is necessary to take a slight detour in our discussion to explain the particular way in which economists represent plant capacity with the use of short-run average cost curves.

You will recall that the typical short-run average cost curve (based on a cubic total cost function) declines, reaches a minimum point, and then rises as the firm produces more with some fixed amount of input. By definition, economists consider the lowest point of the short-run average cost curve to represent the firm's maximum capacity. Although "maximum capacity" usually denotes the physical limit of production (i.e., a plant simply cannot produce any more output), Figure 7.9 shows that a firm is clearly capable of producing beyond the output level at which average cost is at its lowest point. Rather than try to figure out the reason for this particular use of the term, just remember that for economists, "maximum plant capacity" coincides with a level of output that costs a firm the least amount per unit to produce in the short run.

Thus, Figure 7.9 shows that over certain ranges of output, a firm is better off operating a larger plant at less than maximum capacity than a smaller plant at maximum capacity. The average cost of production in the larger plant is lower than the lowest possible average cost of production in the smaller plant. What causes this to happen? Economies of scale, of course. Because of this phenomenon, we can expect that over certain ranges of output, the reduction in average cost resulting from the economies of using a larger plant will be greater than the reduction in average cost resulting from operating a smaller plant at its most efficient (the economic "maximum") level of capacity!

As a cautionary note, we should add that once diseconomies of scale take effect, it is better for a firm to operate a plant of a given size beyond its "maximum capacity" than to build a plant of a larger size. We leave it to you to work through a detailed explanation of this observation. Suffice it to say that the reasoning is very similar to the explanation regarding the impact of economies of scale. (If you need help, refer to Figure 7.9, points p , q , and r .)

Looking at Figure 7.9 from another perspective, we can see that the firm's long-run average cost curve is actually the envelope of the various short-run average cost curves. As such, the long-run curve outlines the lowest per-unit costs that the firm will incur over the range of output considered, given the possibility of using plant sizes A through F. Figure 7.10*a* illustrates this point. As the number of choices of plant size approaches infinity, the envelope becomes a continuous version of the graph; this version is shown in Figure 7.10*b*. The Excel App for Figure 7.10 allows you to interactively examine the discrete versus continuous version of long-run average cost.



APP

Using Long-Run Average Cost as a Decision-Making Tool: The Importance of Coordinating Production Plans with Market Forecasts

In planning for its long-run capacity, which plant size should a firm choose? Using the schedule of numbers in Table 7.3, we can see that if a firm decides to build plant A, then the lowest per-unit cost of production it can expect is \$5. But if it decides to build a plant with greater capacity—for instance, plant C—the potential for reducing per-unit cost is considerable because of economies of scale. Using plant C, the firm could produce its output for as little as \$4 per unit. Indeed, this is where accurate forecasts of product demand are needed to help the firm plan for the best plant size. There would be no sense in investing in plant C, for example, if monthly demand is only 20,000. In this case, the smaller plant B would be more suitable. Actually, given the plant size

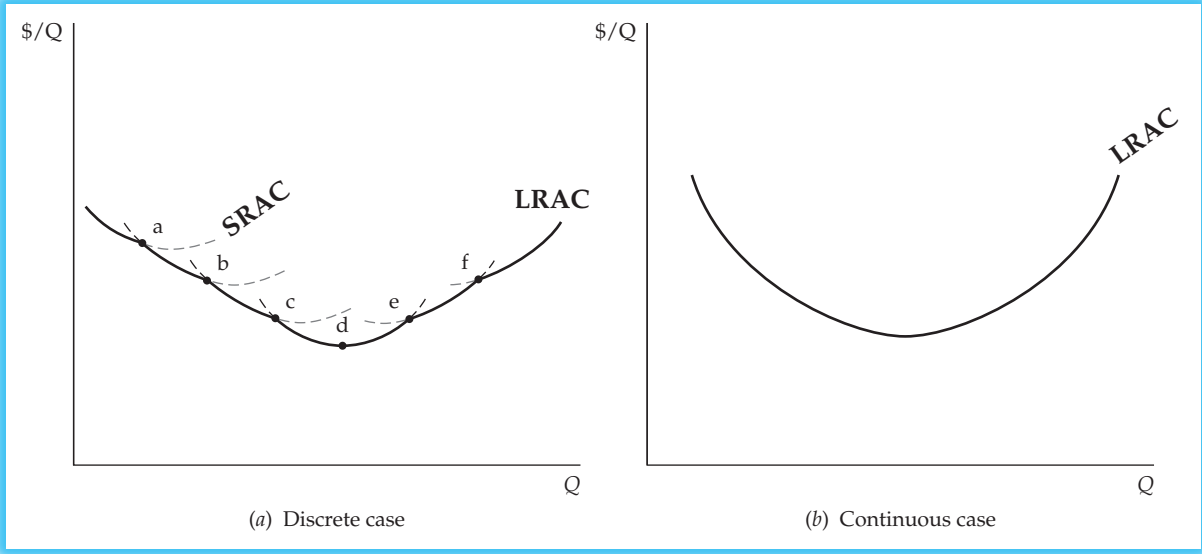


Figure 7.10 Forming LRAC Curve from SRAC Curves 

options shown in Table 7.3 and Figure 7.9, we see that it is not necessary for a firm to forecast a specific amount of demand, but it should at least have a good idea of what the range of demand will be in the future. Table 7.5 presents a schedule showing the most appropriate plant size for different ranges of demand and production.

You should recognize that because diseconomies of scale take effect when plant E is used, plant D is appropriate for output levels on the downward-sloping (from *c* to *d*) and upward-sloping part of its SRAC (from *d* to *e*). The penalty for selecting the inappropriate level of capacity is the incurrence of unnecessary cost, whether the actual demand turns out to be above or below the range used as a basis for the firm’s decision on long-run production capacity. In Figure 7.11, we can see that if the firm had decided to build plant B and the demand turned out to be 25,000, it would lose on a per-unit basis the amount depicted by the arrow. If demand were such that it required the firm to produce only 5,000 units per month, the firm would suffer a similar type of loss.

Table 7.5 Optimal Plant Size According to Expected Demand

Units of Output Produced Based on Expected Demand	Appropriate Plant Size
0 to 10,000	A
10,000 to 20,000	B
20,000 to 30,000	C
30,000 to 50,000	D
50,000 to 60,000	E
60,000 to 70,000 (not shown on graph)	F

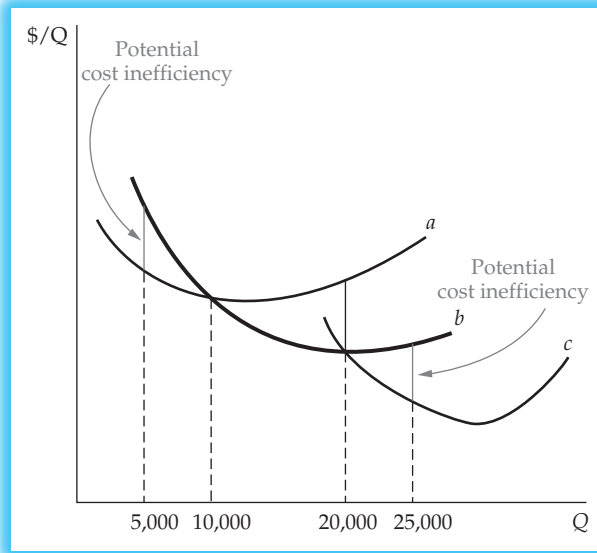


Figure 7.11 Unnecessary Costs Due to Inappropriate Plant Size

As a final comment on long-run cost, it is easy to put the onus for picking the best level of long-run capacity on the market forecasters. After all, their estimates of future demand should be the main guide for the decision makers who plan production capacity. Nevertheless, the production people in a business could become carried away with considerations of production capacity and efficiency for their own sake (e.g., “let us build a plant of the size that will enable us to benefit substantially from economies of scale”) and convince the management to over-build the company’s long-run capacity. Certainly, optimal long-run production decisions require a balanced contribution from both the engineers and the marketing people.

THE LEARNING CURVE

The **learning curve** is a line showing the relationship between labor cost and additional units of output. Its downward slope indicates that this additional cost per unit declines as the level of output increases because workers improve with practice. The reduction in cost from this particular source of improvement is often referred to as the *learning curve effect*.

Specifically, a learning curve is measured in terms of the percentage decrease in additional labor cost each time output doubles. Table 7.6 presents data for an “80 percent” learning curve. Each time the output doubles, the cost of producing the next increment of output decreases to 80 percent of the previous level (i.e., declines by 20 percent). As you can see in this table, the first unit costs \$100,000; the second unit costs 80 percent of this amount, or \$80,000; the fourth unit costs 80 percent of this amount, or \$64,000; and so on. Notice that the percentage reduction is actually with respect to labor hours. However, given some wage rate (in this case, \$10/hour), the labor cost decreases by the same percentage. The data in Table 7.6 are plotted in Figure 7.12.

Table 7.6 Numerical Example of the Learning Curve

Unit Number	Unit Labor Hours	Cumulative Labor Hours	Cumulative Average Labor Hours	Unit Labor Cost	Cumulative Average Labor Cost
1	10,000	10,000.0	10,000	\$100,000	\$100,000
2	8,000	18,000.0	9,000	80,000	90,000
4	6,400	31,421.0	7,855.3	64,000	78,553
8	5,120	53,459.1	6,682.4	51,200	66,824
16	4,096	89,201.4	5,575.1	40,960	55,751
32	3,276.8	146,786.2	4,587.1	32,768	45,871
64	2,621.4	239,245.3	3,738.2	26,214	37,382
128	2,097.2	387,439.5	3,026.9	20,972	30,269
256	1,677.7	624,731.8	2,404.6	16,777	24,046

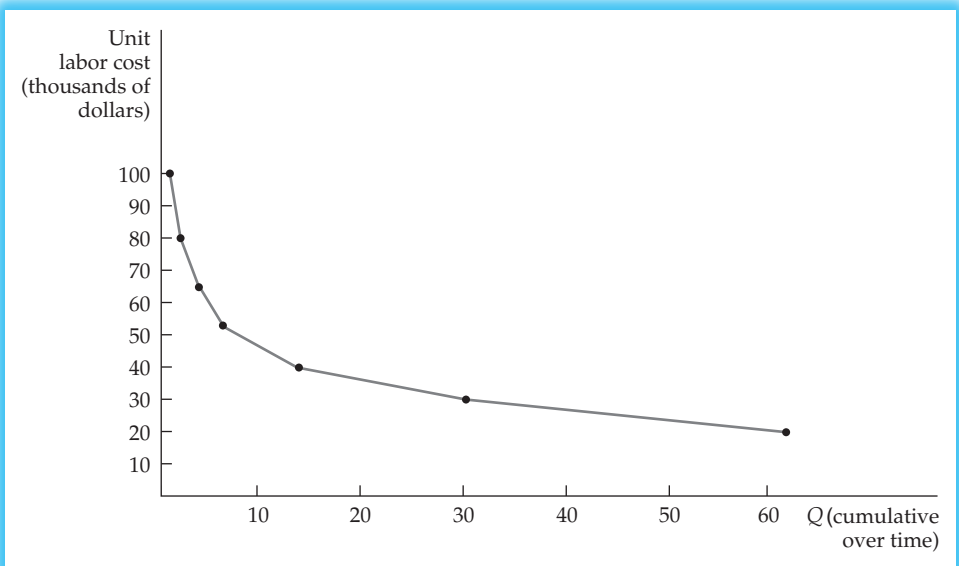
Wage rate = \$10/hr.

There is a mathematical formula for determining the pattern of reduction in labor cost based on a selected percentage decline. This formula is

$$Y_x = Kx^n$$

- where Y_x = Units of factor (labor hours) or cost to produce x^{th} unit
- K = Factor units or cost to produce k^{th} (usually first) unit
- x = Product unit (the x^{th} unit)
- $n = \log S / \log 2$
- S = Slope parameter ($S = 1 -$ the degree of cost reduction)

Figure 7.12 An 80 Percent Learning Curve



For an 80 percent learning curve, the number of direct labor hours required to produce the eighth unit of output is

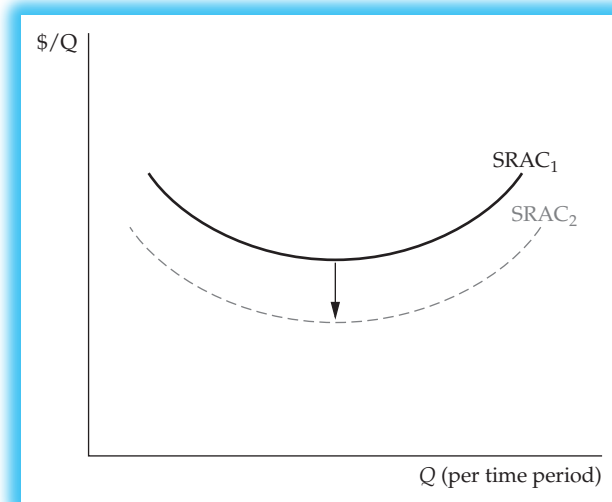
$$\begin{aligned}
 S &= 0.8 \\
 Y_8 &= 100,000 \cdot (8)^{\log(0.8)/\log 2} \\
 &= 100,000(8)^{-.322} \\
 &= \frac{100,000}{8^{.322}} \\
 &= \frac{100,000}{1.9535} = 51,200
 \end{aligned}$$

This answer conforms to Table 7.6 and Figure 7.12. You may want to try constructing other learning curves based on different percentages, or refer to learning curve tables in an engineering textbook.

Although the learning curve is expressed in terms of the marginal cost of production, the impact of improving with practice can also be seen in terms of the decline in average cost. Table 7.6 also shows the cumulative labor cost and the cumulative average labor cost of producing various levels of output. As can be seen, the average labor cost also decreases, although not as sharply as the marginal labor cost. In any case, the learning curve effect clearly has an impact on the short-run cost presented earlier. In particular, the learning curve effect causes the short-run average cost curve to shift downward. This is shown in Figure 7.13.

The Japanese have been frequently cited in academic studies and in the popular press for their use of the learning curve in driving down costs. This is most dramatically shown in their production of computer chips and consumer electronics. Their particular use of the learning curve involves accelerating production experience through aggressive price-cutting measures. The price cuts enhance sales and give them production experience more rapidly. This in turn helps to drive down costs of production faster. The tactic of learning curve pricing is illustrated in Figure 7.14.

Figure 7.13 Impact of Learning Curve Effect on Short-Run Average Cost per Time Period



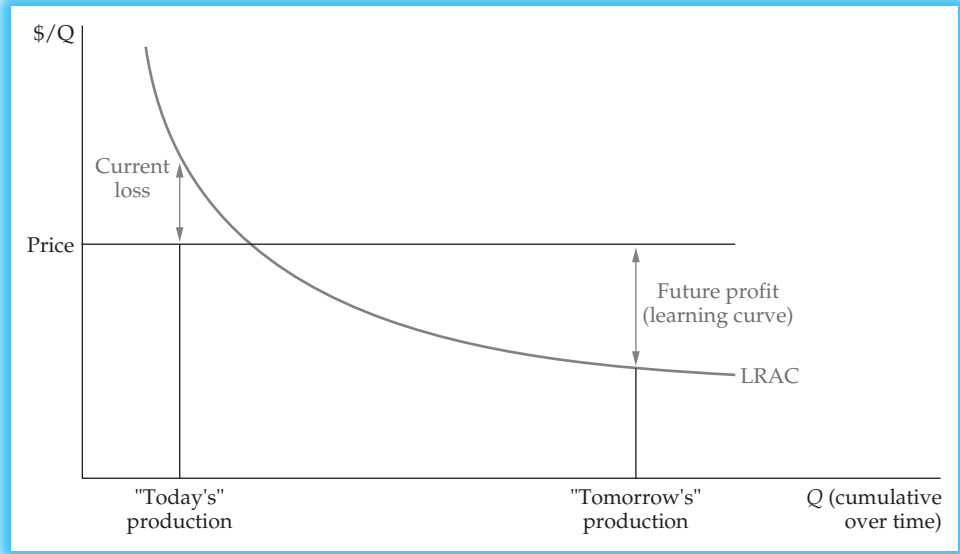


Figure 7.14 Pricing Based on Learning or Experience Curve

In concluding this section on the learning curve effect, we should note that this phenomenon was first observed in the production of aircraft more than 50 years ago. The reason cited for the learning curve effect was the repetition of tasks performed by workers actually manufacturing the product (e.g., direct labor). Later, the experience gained from repetition by those indirectly related to the production process (e.g., engineers, researchers) was also included in this phenomenon. Thus, such factors as the development of new process and engineering methods, the substitution of lower-cost materials or processes, and product redesign were also considered as factors causing unit costs to decrease as production levels increased. *The recognition* of these additional factors prompted the use of the broader term **experience curve**. Today, *experience curve* and *learning curve* are generally used interchangeably. However, some business managers and consultants still prefer to make a distinction between the two terms.

ECONOMIES OF SCOPE

Before summarizing this chapter, we should introduce briefly a concept often used in business but that does not quite fit into the conventional economic theory of cost, namely, **economies of scope**. This term can be defined as

the reduction of a firm's unit cost by producing two or more goods or services jointly rather than separately.

The effort by our hypothetical firm Global Foods, Inc., to increase its profits by expanding into the soft drink industry offers a good example of a company seeking to take advantage of economies of scope. The company already has the knowledge, experience, and skills to produce and distribute processed food items and hopes to use these attributes in the production and distribution of soft drinks.

In a sense, the concept of economies of scope is closely related to economies of scale. Engaging in more than one line of business may require a firm to have a certain minimum scale of operation. Another way to view this relationship between scale and scope is to consider that a company's expansion into different lines of business naturally increases its scale of operation.

A good example of economies of scope is the case of PepsiCo's beverage and snack business relative to its distribution channels. Potentially, negotiations with wholesalers on pricing, with retailers on shelf space and product placement, and with bottlers and distributors on truck routing and delivery schedules could all be done with both types of products at the same time and by people of similar skills. Another example of economies of scope occurs in the area of mass advertising. Consider the millions of dollars that advertisers such as beer, auto, and soft drink beverage companies spend for a 30-second commercial during the Super Bowl. Now consider two hypothetical beer companies, one that has 20 million regular customers and the other that has only 5 million. Each company must spend the same amount of money but the bigger one is able to remind 4 times as many customers to remain loyal to its brand or brands.

ECONOMIES OF SCALE: THE SHORT RUN VERSUS THE LONG RUN

Economies of scale essentially means that a company's average cost decreases at higher levels of output. In microeconomic theory, *scale* implies a long-run time period. Recall that the economic meaning of *long run* implies that firms have time to adjust their fixed factors of production. In terms of cost, this means in effect that a firm's "fixed" cost is really "variable" in the long run. Another possible action that a firm could take in the long run is to merge with or acquire another firm or be acquired by another firm. When potential mergers are announced, the CEOs often tell the analysts that one of the benefits is "economies of scale."

From time to time, we have observed that the term *economies of scale* is used in the context of a "short-run" time period. In other words, authors of articles might say that a firm tries to take advantage of economies of scale by operating its plant at or near full capacity. This type of scale economies occurs because in so doing, a firm is spreading out its fixed or overhead costs. This was not the original intention of economic theorists, but we cannot prevent writers from using the term in this way. Therefore, a useful distinction can be made between "short-run economies of scale" and "long-run economies of scale." The reason for the first type of economies is the spreading of fixed cost. The reasons for the second type are listed in Table 7.4.

SUPPLY CHAIN MANAGEMENT

Supply chain management (SCM) has become an important way for companies to reduce their costs. Readers will undoubtedly study SCM in great detail in an operations research or production course in an undergraduate or MBA curriculum. Certain elements of SCM pertain either directly or indirectly to managerial economics. As you can see, SCM played a part in the "The Situation" vignette presented at the beginning of this chapter. Therefore, we now present a brief introduction to the subject.

Supply chain management can be defined as "efforts by a firm to improve efficiencies through each link of a firm's supply chain from supplier to customer." This is done primarily by fostering better communication and cooperation within each link by all parties involved. The goal of SCM is to increase profits primarily by reducing costs. But SCM can also help increase profits indirectly because a more efficiently operating supply chain can increase customer satisfaction. This in turn may enable firms

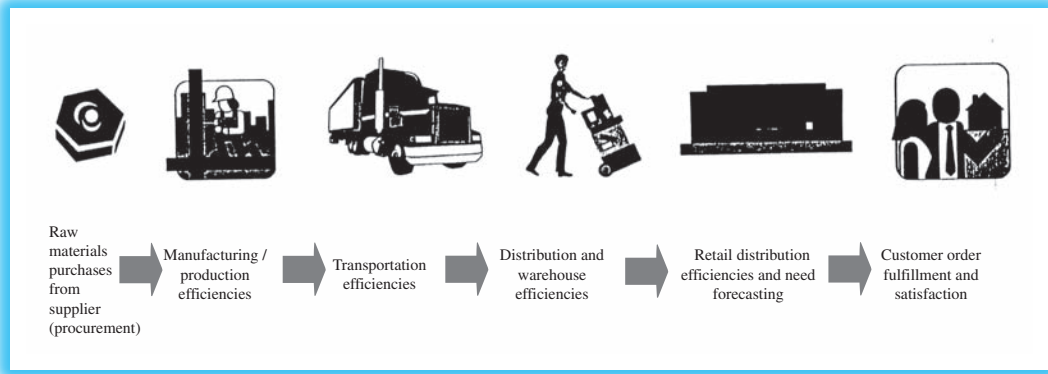


Figure 7.15 Supply Chain Management

to charge a premium price or at least help them retain their customers. Figure 7.15 shows a supply chain for a typical manufacturing company and the efficiencies that SCM wants to accomplish at each linkage point.

In the most general sense of the term, supply chain management includes all the firm's internal and external activities required to fulfill its customer's demand. Historically a typical firm, particularly in manufacturing, was a completely vertically integrated entity. U.S. Steel had its own coal mines that eventually fired up its mills. IBM made its own chips (and still does) that go into its computers. But in recent years, much more of a company's activities have been outsourced to other companies. Does Folgers have its own coffee plantations? Does General Motors manufacture antilock brakes or catalytic converters? Herein lies the reason for the increasing importance of SCM among practitioners, as well as the increasing attention given to it by academic economists. Presumably, a firm outsources certain activities to reduce costs. But once the activity is outside the direct control of the firm, there may be certain other costs that may in fact outweigh the cost savings of outsourcing. Let us continue with this discussion by focusing on the supply chain link between two separate businesses, one of which is supplying and the other of which is buying. In manufacturing, this could involve suppliers of component parts to the company that is making the final product (e.g., Bosch supplying antilock brakes to Mercedes-Benz). In retail, this could involve suppliers of final products to retail stores (e.g., Colgate-Palmolive supplying toothpaste to Wal-Mart). In the context of this particular link, our topic is essentially the "make versus buy" decision that is also discussed in accounting and finance texts, as well as economics texts.⁷

From the perspective of managerial economics, the "make versus buy" decision or the trade-off between cost savings by going outside the firm versus costs incurred by staying inside raises the most fundamental of all questions: What is a firm? Recall our discussion of this question at the outset of Chapter 2, "The Firm and Its Goals." You may also recall that the seminal work on this subject was done by Professor Ronald Coase, who received a Nobel Prize in Economics for this work. Essentially, Coase

⁷Potentially any or all links of a supply chain could be outsourced in any type of industry. Therefore, supply chain management costs discussed in this section have much broader implications. Think about the supply chain problems that occur when a bank issuing credit cards outsources its direct mail advertising or its billing or even its call centers. Think about what eBay is. Some might consider it the ultimate outsourced retail company.

postulated that a company compares the costs of organizing an activity internally with the cost of using the market system for its transactions. The cost incurred by using resources outside was what he called “transaction costs.” These costs include the original investigation to find the outside firm, followed by the cost of negotiating a contract, and later, enforcing the contract.

Following Coase’s path-breaking work, two other categories of costs of going outside the firm have been identified in addition to transactions costs: coordination costs and information costs.⁸ Coordination costs arise because of the uncertainty and complexity of such tasks as timing the deliveries, meeting customer specifications, making alternative payments arrangements, and forecasting needs. Uncertainty exists because it simply may be impossible to plan for or insert in a contractual agreement what to do for every possible case in which an outside supplier is not able to completely satisfy the firm’s requirements. Information costs arise because information is essential to the proper coordination of activities between the firm and its suppliers. For example, to meet the firm’s exact specifications, a supplier might require timely and detailed information about a firm’s sales. This may mean that a firm has to provide its suppliers with proprietary information. Short of sharing this type of information, the firm must be prepared to sacrifice some efficiency in the supplier’s delivery obligations. If this information is shared, there is a certain amount of cost associated with setting up and maintaining the necessary technology platform (e.g., data communications network, logistics software, etc.). As is the case with coordination problems, the need for information flows complicates the relationship between a firm and its outside supplier that also cannot be easily addressed with a contract, no matter how detailed are the terms and conditions.

In the linkage between the firm purchasing the input and the firm supplying the input, one critical indicator of supply chain efficiency is the level of inventory. Ideally, the firm purchasing the input wants to hold no inventory in excess of what it needs to fulfill what its customers demand. Some years ago, Japanese automobile manufacturers became well known for their “just in time” (JIT) inventory control system.⁹ What we are talking about is very similar, except we consider JIT to be a part of a bigger picture (i.e., SCM). If anything goes wrong in this linkage, firms purchasing the input will find themselves with more or less inventory than desired. Firms selling the inputs will find themselves with an inventory problem that is a mirror image of the one experienced by firms purchasing its product. As a result, the vendor may be faced with either too much or too little capacity in its production facilities relative to the purchasing firm’s mismatch of inventory levels and needs. In all cases, added costs are involved. How should this link of the supply chain be managed in such a way that neither party experiences these problems and added costs?

⁸One could argue that these costs are all part of transactions costs. However, supply chain management experts prefer to use these additional categories. See, for example, Ian Stuart and David McCutcheon, “The Manager’s Guide to Supply Chain Management,” *Business Horizons*, March/April 2000, pp. 35–44. Some of this section’s comments on contracts and supplier relationships were taken from this article.

⁹During his time working with Sterling Vineyards, Stephen Erfle came across an extreme example of JIT inventory control. Due to the footprint of Sterling’s bottling facility and its inability to get an expansion of that footprint approved by Napa County, with its strict agricultural zoning laws, Sterling had the bottles it used on a given day delivered the night before. A truck had to bring empty bottles “Up Valley” each night. That same truck went back “Down Valley” the same night with bottles that had been filled the day before. This arrangement circumvented zoning laws because it made it possible to have an operation of Sterling’s size (300,000 cases per year at that time) operate without any on-site finished bottle inventory or empty bottle storage spaces. Roughly 1,000 cases per night were moved in and out of the bottling line facility. This worked due to the coordination and long-standing relationship Sterling maintained with their bottle supplier, the same supplier used for all of Seagram’s products. At the time, Sterling Vineyards was owned by Seagram Classics Wine Company.

There are numerous consultants and software products out in the marketplace that promise firms an answer to this question not only for purchasing but for the entire supply chain. This is supposedly what a consultant implies when he or she uses the term *supply chain management* as a buzzword. Indeed, we can venture to say that all or most major consulting firms have their own supply chain management practice. Moreover, software companies such as Oracle and SAP offer software packages that help companies to become more efficient in some or all aspects of their supply chain.

EXAMPLES OF WAYS COMPANIES HAVE CUT COSTS TO REMAIN COMPETITIVE

The consideration of cost in business decisions often evokes mixed feelings among managers. This is because efforts to reduce costs may involve the downsizing of staff, the layoff of workers, reductions in discretionary expenses for such things as training and travel, or the elimination of executive perquisites such as corporate jets. In contrast, efforts to increase revenues involve growth and all the activities that accompany it (e.g., new hires, better equipment, new offices, increases in discretionary spending and perquisites). Yet, from an objective standpoint, it is clear that a dollar saved, given some level of output and revenue, adds as much to the bottom line as an additional dollar of revenue, given some level of output and cost. Following are examples of cost-cutting methods or procedures that firms have implemented to remain competitive.

1. **The strategic use of cost:**¹⁰ As we explain in Chapter 9, a firm may choose to use cost leadership as the core of its strategy for competing in the marketplace. In the retail industry, the example that immediately comes to mind is Wal-Mart. In the airline industry, it is Southwest Airlines. Let us look closer at how this airline used cost leadership as an integral part of its strategy.

Southwest Airlines began in 1967 as a small operation aimed at improving air service between Houston, Dallas–Fort Worth, and San Antonio, Texas. Today it is a formidable competitor in the short-haul (500 miles or less) air transportation market. Indeed, there are now certain routes (e.g., Burbank, California, to Las Vegas, Nevada) that the major airlines such as American Airlines have withdrawn from since Southwest’s entry into the market. Southwest Airlines has achieved this success by being the low-cost provider of short-haul service. Its everyday fares can be as low as one-third of the standard fares charged by the major airlines on short-haul routes.

It is more meaningful to look at Southwest’s cost advantage not as a series of ad hoc cost-cutting measures, but as the consequence of its overall market strategy. When the founders of Southwest Airlines devised their strategy of providing low-priced air transportation in the short-haul market, they anticipated being able to keep costs down by virtue of the very strategy that they selected. By flying only on short-haul routes, Southwest needed only one type of aircraft (the Boeing 737). This meant less training costs for flying and maintenance, as well as a lower inventory of replacement parts. Furthermore, short hauls require no in-flight meals. In addition, because it has positioned itself in the minds of its customers as a low-cost provider of air transportation, it can save more money by cutting out services or conveniences that customers generally expect from the major airlines. For example, Southwest does not take reservations, thereby saving money on staff and computer services. Whenever possible, it also uses secondary or alternative airports, thereby avoiding the higher landing fees charged by the major-hub airports.¹¹

¹⁰For a good historical review of Southwest Airlines’ start-up and operations, see Christopher H. Lovelock, “Southwest Airlines,” Harvard Business School Case Study #575–060, 1975.

¹¹Most of the material in this section was taken from Gene Walden and Edmund O. Lawler, *Marketing Masters: Secrets of America’s Best Companies*, New York: HarperCollins, 1993, pp. 17–22.

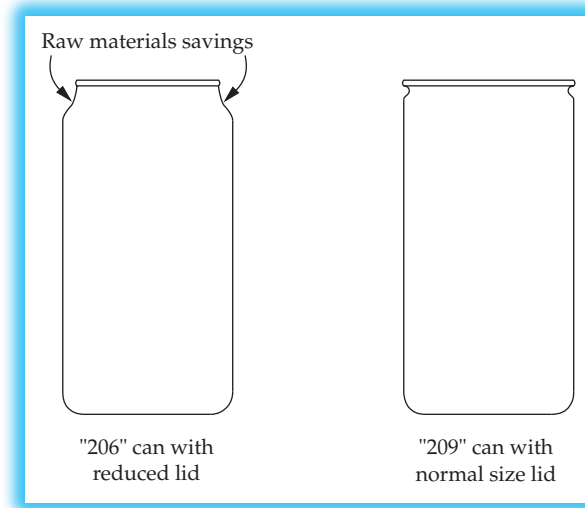


Figure 7.16 Using Reduced-Lid Can to Achieve Raw Materials Savings

Over the years, Southwest Airlines' low-cost approach continues to outperform the "legacy" airlines in earning a profit. In 2001, the year of the 9/11 tragedy, all of the major U.S. airlines lost money except for Southwest. During the Great Recession of 2008 and onward, Southwest earned a profit while the other major airlines were losing money. In fact, by the end of 2011, Southwest had achieved a record thirty-nine straight years of profitability.¹²

- 2. Reduction in cost of materials:** Manufacturers are constantly seeking to cut costs by materials substitution or modification. In the beverage industry, one noticeable example of this is the reduction in the size of the lid of a 12-ounce aluminum can. Toward the end of the bottling process, each can on the conveyor belt passes a measuring device called a "fill height detector" that checks if there are at least 12 ounces of liquid in the container. Those with less are discarded. This measuring device shoots a gamma beam about one-quarter inch below the top of the can's lid. Because it is not necessary for the product to be filled to the very top of the can, the lid can be made slightly smaller by crimping the top portion of the can. The savings from the reduction in the aluminum content are considerable, given the millions of cans that are produced each year (Figure 7.16).

Cost savings can come for other reasons besides using fewer materials or less expensive materials. Considerable savings can occur when lighter materials reduce related energy costs. For example, the use of lightweight composite materials (i.e., carbon fibers meshed together by epoxy instead of aluminum) is one of the main features of Boeing's new 787 Dreamliner. This is one of the reasons why the Dreamliner is expected to burn 20 percent less fuel than airplanes of similar size.¹³ Another good example can be found in the processed food and snack business. Sonoco is the leading manufacturer of consumer packaging. In an interview with one of this textbook's authors, a representative of the company said that the new lightweight plastic bottles that it developed for Planter's Peanuts enables the company to use 25 percent fewer trucks to transport the product to retail stores.¹⁴

¹²http://www.slate.com/articles/business/operations/2012/06/southwest_airlines_profitability_how_the_company_uses_operations_theory_to_fuel_its_success_.html

¹³<http://articles.latimes.com/2011/oct/01/business/la-fi-cover-jet-revolution-20111002>

¹⁴Go to www.sonoco.com for more examples of the type of packaging that this company provides for many well-known food and snack products.

3. **Using information technology (IT) to reduce costs:** From its early beginnings as a tool for the mechanization of routine business tasks such as accounting, IT has played an increasingly greater role in the efforts by companies to reduce costs. The real breakthrough in using IT to reduce costs began in the early 1990s with *enterprise resource planning* (ERP). SAP, a German company, was one of the first companies to offer services for this tool and continues today to be a leader in the industry. Another company offering similar products is Oracle.

ERP is the industry term for software that enables companies to integrate their planning, scheduling, synchronization, and tracking of the entire chain of their production process. The use of ERP has helped companies cut costs and increase revenues by getting the right products into the hands of their customers in the most timely and efficient way. SAP's website (www.sap.com) provides examples of how companies have saved money by using its products and services.

4. **Reduction of process costs:** Cost savings stemming from the use of ERP has been getting most of the attention in the popular press and the literature on business studies. However, there are also more specific but equally important ways that appropriate software can help cut costs. Most companies are starting to require that their employees use software packages to file travel and expense reports electronically. American Express conducted one of the first studies showing how ERP can help companies cut their costs. Back in 1998, it found that U.S. companies spent about \$175 billion on travel and expenses in 1998 and that the use of electronic filing can cut the cost of processing expense reports from \$36.46 per paper report to as low as \$7.91.¹⁵
5. **Relocation to lower-wage countries or regions:** Over the past decade, much of the manufacturing by U.S. (as well as western European and Japanese) companies has shifted to China. But manufacturing for companies around the world is also being expanded throughout Southeast Asia (e.g., Thailand, Vietnam, and Malaysia) and eastern Europe (e.g., Poland, Hungary, Czech Republic, Slovakia, and Romania).

At the same time that manufacturing was shifting to China, service operations such as call centers were shifting to India. Information technology (IT) support and software development have also increased tremendously in India. A decade or so ago, wages for China and India were considerably lower than wages for comparable work in the United States, Japan, and western Europe. But the wage differentials have begun to close.¹⁶ As a result, certain companies, such as apparel companies, have started to shift production out of China to countries such as Vietnam. There are now more call center workers in the Philippines than there are in India.¹⁷

6. **Mergers, consolidations, and subsequent downsizing:** This particular approach to cost cutting was cited earlier in this chapter. We also discuss it briefly in Chapter 14.
7. **Layoffs and plant closings:** Sometimes companies find it necessary to resort to simply laying off workers and closing down certain operations. Although downsizing and restructuring have become practically a way of life for corporate America, they have become much more common in many other parts of the world. For example, in Japan, a country formally known as one that provided lifetime employment for its workers in large corporations, downsizing has become much more common. As of the writing of this edition of this textbook, Olympus, the well-known Japanese maker of cameras, announced a layoff of 2,700 of its staff.¹⁸
8. **Reduction in fixed assets:** Fixed assets are long-term investments in tangible resources (e.g. property, plant, and equipment) and intangible resources (goodwill, intellectual property, and software). Finance and accounting courses cover the use, measurement, and management of assets in great detail. In managerial economics, a firm's cost of

¹⁵"Software Is Easing Expense-Form Headaches," *Wall Street Journal*, December 28, 1998.

¹⁶http://articles.economicstimes.indiatimes.com/2011-09-29/news/30213297_1_india-salary-expat-salaries-cisco-india

¹⁷<http://www.cnn.com/2012/04/30/world/asia/philippines-call-centers/index.html>

¹⁸<http://www.irishtimes.com/newspaper/breaking/2012/0608/breaking11.html>

buying and managing fixed assets tends to be covered at a rather high level of abstraction. We simply treat it as part of the fixed cost of doing business. More specifically, this fixed cost includes depreciation as an accrued expense on a company's income statement and a cash outlay on its statement of cash flows. The addition of new capital investment in assets plus existing assets minus accumulated depreciation of the existing investments would rest on a company's balance sheet as net assets. Interestingly, economic theory, financial accounting, and corporate finance all converge in their coverage of capital budgeting, which we cover in Chapter 12.¹⁹

CAUTIONARY NOTE TO MANAGERS ABOUT THE USE OF COST-CUTTING AS A STRATEGY

In Chapter 9, we introduce the ideas on strategy developed by Professor Michael Porter.²⁰ One of the main ideas introduced in an early chapter of his book was about a company seeking a “differentiated” versus a “cost leadership” approach to strategy.²¹ We clearly saw an example of this in the case of Southwest Airlines discussed in the previous section. But in a later monograph, Porter warned about the long-run sustainability of such a strategy when a company focuses primarily on “operational excellence.”²² In the end, he believed that cost leadership strategies are more easily duplicated by rivals. Therefore, companies that want to attain a sustained competitive advantage should focus on a differentiated strategy. In other words, increasing profit by growing revenue and profit margin is more effective in the long term than increasing profit by only cutting cost.

We agree with Professor Porter about this. However, in times of crisis, companies cannot help but turn to cost-cutting to stay competitive. Moreover, the tremendous global competition today makes it imperative that all companies stay focused on being cost efficient, regardless of the level of demand for their products. But perhaps another disadvantage of focusing too much on the cost leadership approach, besides being easily copied, is the danger that this will distract senior management from closely watching and adjusting to consumer demand. Still worse, it could lead to developing a sense of complacency or what consultants might refer to as a “competitive blind spot.” We believe such was the case at Dell Computer.

Over the years, Dell has become a textbook case of how a company can successfully compete using the low-cost approach. Its model of direct selling, primarily to businesses, combined with the outsourcing of much of the manufacture of its PCs (in the early days rivals would scoff at Dell by calling it a “bag and tag” shop), resulted in a true powerhouse in this market. But then, suddenly, the market started to shift. Businesses were starting to buy PCs more as replacements than as incremental acquisitions. Furthermore, the big boost in demand in the market was coming from people who wanted to buy laptops, not desktops for their home use, because of lower cost broadband services and the general desire to be more mobile. Purchasing agents of companies may feel comfortable buying laptops as well as desktops on the Internet,

¹⁹We make no attempt to explain in any detail the relationship between the economic theory versus financial accounting in the treatment of assets. Another chapter would probably have to be written for this. But we remind readers that you must always look for connections between the different courses that you take in your business curriculum.

²⁰Porter's ideas on strategy are probably the most quoted in the business world today. His first book on this subject is highly recommended and remains a classic text on strategy. It is in its 60th printing and has been translated into 19 languages. See Michael Porter, *Competitive Strategy: Techniques for Analyzing Industries and Competitors*,

²¹See Chapter 9, pp. 375–377.

²²Michael Porter, “What Is Strategy?” HBR OnPoint Enhanced Edition, February 1, 2000.

but not so with consumers. They would much prefer to go to a store and try out a laptop for themselves. Furthermore, Dell's models and designs were geared more toward the office than the family room. When Dell's profits began to dive partly because of this, an industry analyst, Sam Bhavasi, was quoted as saying, "Dell was not paying attention to the competitive landscape."²³

Dell tried to recover from its slip-up in the consumer market for laptops by selling its laptops through major retailers such as Wal-Mart and Staples, thereby modifying its direct sales business model. But in a broader, more strategic move, Dell has decided to renew its focus on business markets with products and services other than personal computers. Currently, it is concentrating on selling servers and host computing services to businesses with sales of less than \$500 million a year. As CEO Michael Dell stated, "Within the context of a \$3 trillion global market in information technology, consumer is \$250 billion. The commercial business is \$2.75 trillion. We tend to focus more on that [commercial business]."²⁴

GLOBAL APPLICATIONS: LI & FUNG WILL DO IT ALL FOR YOU

In this chapter, we talk about a company outsourcing parts of its supply chain. How about outsourcing the *entire* supply chain? You may never have heard of Li & Fung, but we are sure you know many of the brands of clothing or consumer goods that this company has produced, distributed, and perhaps even designed. This Hong Kong-based global company is described by its CEO as a "supply chain manager and network orchestrator on a global basis."²⁵ The CEO goes on to say that Li & Fung is an "Asian-based multinational operating in 40 countries around the world."

The company was started in 1906 by Li and Fung in Canton, China, as an exporter of porcelain, fireworks, jade, and silk, primarily to the United States. In the 1930s, it moved to Hong Kong, exporting products, such as toys, wigs, garments, and plastic flowers, that gave rise to the well-known label at the time "Made in Hong Kong." In the 1970s, because of pricing pressures from both their suppliers and customers (i.e., the importers), the company transformed itself into a full-service manufacturing services company.²⁶ This means that the company offers help to other companies in all aspects of the supply chain, from design, sourcing, marketing, and distribution. For example, if a clothing designer wants to manufacture his or her apparel in China, Li & Fung will find the contract manufacturer, assist in the selection of fabrics and manufacturing process, monitor the quality of production, and develop the distribution system. The company also buys the licenses to use a particular brand and then conduct the entire production and distribution operation of the branded apparel on their own behalf. In recent years, Li & Fung has started conducting sourcing and supply chain operations in countries other than China such as those in South Asia, Central America, and parts of the Middle East.

²³Quoted in Crayton Harrison, "Dell Shares Drop 10% on Warning: Computer Maker Loses \$5 BIL. in Value on Profit Slowdown," *Dallas Morning News*, July 22, 2006.

²⁴<http://bits.blogs.nytimes.com/2012/02/27/dell-is-hoping-its-more-than-a-server-company/>

²⁵From a video shown on the company's website: www.lifung.com

²⁶Some of the historical highlights of the company can be found on its website and also at: http://en.wikipedia.org/wiki/Li_%26_Fung

A listing of the company's recent acquisitions can be found in its annual report. Most of them are not well-known outside the apparel industry, but perhaps you have heard of French luxury designer Sonia Rykiel. In early 2012, Li & Fung bought an 80 percent equity stake in the company. Li & Fung's parent company, the Li & Fung Group, owns the Toys 'R' Us stores that operate in China. One of the authors of this textbook once visited one of these stores in Shanghai and found it amusing to see all of these toys sold in the country of manufacturing origin.

The Solution

Terry now needed to put all the pieces together to figure out if Lawrence Aluminum was really the right way to go. Even before crunching the numbers, she knew she had a somewhat delicate situation to consider. Her company had been doing business with Kayla Containers for many years, and the two companies had developed a close relationship (as well as friendly rivalry in the local softball league). If Shayna Soda stopped ordering from Kayla, Kayla could very easily go out of business. Lawrence was just a start-up, eager to grab new business by "low-balling" its prices. But what if Lawrence went out of business or decided not to source to upstate New York bottlers? Then what would her company do?

With this in mind, she turned to the facts. There was no denying that Kayla Containers had a higher base fee on soda cans, but that included the artwork. Furthermore, Kayla was always willing to change the artwork at no extra cost. Lawrence Aluminum's cans would be more than 20 percent cheaper in the base fee, but the artwork required an extra one-time fee. Terry took that fee and divided it out for the year to figure what the impact would be on the average cost of the cans for the first year. This fee increased the base cost of the cans another 2 percent. The cans were still considerably cheaper.

Next Terry considered the delivery charge. Kayla Containers was only 20 miles from Shayna's production plant. The delivery fee for Kayla was basically the same as Lawrence's base fee. But Kayla's fee was for only 20 miles delivery and Lawrence was for 400 miles delivery. In addition, Lawrence charged extra for specific time delivery and could even require the company to take inventory earlier than it wanted. Terry calculated the average cost of delivery annually using Kayla, and then for Lawrence. She also had to determine what the warehouse carrying costs would be for the inventory that would have to be held on early delivery. For each day that the cans were early, Shayna would incur extra costs for having that inventory on the floor. If the inventory needed to be moved for any reason, then there was a labor-hour issue involved. Now the costs really changed. Taking into consideration all these additional costs, Lawrence's base cost now jumped to 8% higher than Kayla.

Terry wanted to immediately run into Adam's office to tell him that Kayla was the better choice, but she first had to make sure that she covered all angles. She admitted that Lawrence Aluminum had a better quality aluminum, but Kayla's quality was perfectly acceptable and in fact better than Shayna's corporate minimum requirements. As far as technology was concerned, Terry knew that her company's computer systems needed to be upgraded in order for them to use a sophisticated supply chain management process regardless of which supplier they used.

The technology issue actually proved to be a critical factor in helping Terry to make her decision. She figured that if her company went with Lawrence Aluminum, then it would have to set up its systems according to this new vendor's specifications. Kayla, in contrast, did not have a sophisticated system in place and in fact was also thinking that it needed to upgrade its system. If Shayna stuck with Kayla, then Kayla would have to set up its system in a way that would integrate with Shayna's. It would be the other way around for Shayna if Lawrence Aluminum became the supplier. Terry e-mailed her friend Ryan over at Kayla to see if this would be a good move from Kayla's point of view. As soon as Ryan got her e-mail, he picked up the phone to talk to her in person. Given the possibility of losing a major customer and the fact that his company had to improve its systems anyway, there was no question in Ryan's mind that Kayla would work in every way possible to keep Shayna as its customer.

Now Terry was ready. It was Friday and she had all the answers to bring to Adam Michaels. As soon as she walked in, Adam asked, "When are we switching?"

Terry replied confidently, "We're not, we're sticking with Kayla."

SUMMARY

This chapter is devoted to the analysis of the cost structure of the firm, both in the short run and the long run. In the short-run analysis, it is important to keep in mind that “behind the scenes,” the law of diminishing returns causes the firm’s marginal cost to increase, and that this increase in turn affects the pattern of behavior of the firm’s average variable cost and average total cost. In the long-run analysis, it is important to recognize that certain factors may cause the firm’s unit costs to decrease as its scale of operation increases. However, certain other factors may actually cause the unit cost to increase if the scale of operations becomes too large. In addition, other factors such as the learning curve effect and economies of scope must be taken into account in the long-run analysis of cost.

The material that is covered in this chapter may seem rather mundane because of the great emphasis placed on defining the various cost terms used in economic analysis. However, your patience and thoroughness in reading this chapter will be rewarded when you read the next four chapters. There you will find that the material in this chapter, together with the analysis of demand presented in Chapters 3 and 4, forms the core of the economic analysis of the firm. As you will learn, all the decisions involving the production and pricing of goods or services—and, in a much broader sense, the very desirability for a firm to be in a particular line of business—depend on the demand for the particular product, the cost at which the firm is able to provide this product, and the competitive structure of the market in which it is operating.

IMPORTANT CONCEPTS (Items with * are discussed in Appendix 7B.)

Average fixed cost (AFC). The fixed cost per unit of output. (p. 256)

Average total cost (AC or ATC). The total cost per unit of output. (p. 257)

Average variable cost (AVC). The variable cost per unit of output. (p. 256)

Constant returns to scale, CRTS. If in the long-run total cost increases by the same proportion as output, then the production process exhibits CRTS. In this instance, unit costs are constant and are equal to minimum LRAC. (p. 263)

Diseconomies of scale, also called decreasing returns to scale, DRTS. The increase in the unit cost of production as the firm increases its capacity. Like economies of scale, this is considered to be a long-run phenomenon. Among the more important reasons a firm may experience rising unit costs as its scale of production increases are (1) management coordination and control problems and (2) a disproportionate increase in staff in relation to indirect labor. (p. 263)

Economies of scale, also called increasing returns to scale, IRTS. The reduction in the unit cost of production as the firm increases its capacity (i.e., increases all its inputs). It is considered a long-run phenomenon. Among the reasons a firm experiences economies as its scale of production increases are (1) the ability to exact volume discounts from vendors, (2) the ability to use more fully specialization and division of labor,

(3) the ability to justify the use of certain types of capital equipment or technology appropriate only for very large scales of production, and (4) management efficiencies resulting from an increased span of control at all levels of management. (p. 263)

Economies of scope. The reduction in cost resulting from the joint production of two or more goods or services. (p. 272)

***Engineering cost estimating.** A method of estimating long-run costs. Professionals familiar with production facilities calculate optimal combinations of inputs needed to produce given quantities of output. Monetary values are then assigned to obtain cost. (p. 303)

***Estimation of long-run cost functions.** Analysis that assumes all factors, especially capital, are variable. Cross-sectional regression analysis is customarily employed in the estimation. (p. 299)

***Estimation of short-run cost functions.** An analysis in which certain factors are assumed to be fixed during the period analyzed. Time series regression analysis is customarily employed in the estimation. (p. 296)

Experience curve. The relationship between the unit cost of labor and all inputs associated with the production process (i.e., both direct labor, such as factory workers, and indirect labor, such as design engineers). (p. 272)

Historical cost. The cost incurred in a past activity (e.g., the purchase price of an asset or the cost

of a project incurred up to the point at which a decision is to be made). (p. 251)

Incremental cost. The total cost associated with a particular decision (e.g., the cost of building an additional wing to an office building, the cost of going into the soft drink business). If incremental cost is considered on a per-unit basis, it becomes marginal cost. For example, incremental cost can be considered the change in total variable cost, whereas marginal cost can be considered the change in total variable cost divided by the change in output. (p. 252)

Learning curve. The relationship between the unit cost of labor and the total amount of output produced by labor that is directly associated with the production process (i.e., “direct labor”). Essentially, this concept is based on the principle that one improves with practice. The resulting productivity gains lead to a reduction in the direct labor cost of producing a unit of output. (p. 269)

Marginal cost (MC). The cost to a firm of producing an additional unit of an output. (p. 253)

Opportunity cost. The amount of subjective value forgone in choosing one activity over the next best alternative. (p. 251)

Relevant costs. Costs that are affected by a current decision alternative and that must therefore be taken into account in the decision. Variable costs and incremental costs are considered to be relevant costs. (p. 251)

Sunk cost. A cost incurred in the past that is not affected by a current decision. If a resource has no opportunity cost (i.e., it has no market value in an alternative use), it is considered to be sunk. (p. 252)

Supply chain management. Efforts by a firm to improve efficiencies through each link of a firm’s supply chain from supplier to customer. (p. 273)

***Survivorship technique.** A method for estimating long-run cost curves. The proportion of total industry output by firms of different sizes is observed over a period of time. That size segment of the industry that gains in proportion of industry output over time is deemed most efficient (lowest cost). (p. 304)

Total cost (TC). The total cost of production, including both total variable and total fixed costs. (p. 253)

Total cost function. Economic analysis considers three basic functional forms of total cost: cubic, quadratic, and linear. The microeconomic theory of the firm relies almost entirely on the cubic function, and break-even analysis generally uses the linear form. (p. 258)

Total fixed cost (TFC). A cost that remains constant as the level of output varies. In a short-run analysis, fixed cost is incurred even if the firm produces no output. (p. 253)

Total variable cost (TVC). The total cost to a firm of using its variable inputs. (p. 253)

QUESTIONS

- Define and compare the following types of cost:
 - Sunk cost versus incremental cost
 - Fixed cost versus variable cost
 - Incremental cost versus marginal cost
 - Opportunity cost versus out-of-pocket cost
- Point out which costs in the preceding question are considered “relevant” and which are considered “irrelevant” to a business decision. Explain why.
- Explain the relationship between a firm’s short-run production function and its short-run cost function. Focus on the marginal product of an input and the marginal cost of production.
- “If it were not for the law of diminishing returns, a firm’s average cost and average variable cost would not increase in the short run.” Do you agree with this statement? Explain.
- Explain the distinction made in economic analysis between the short run and the long run.
- Define *economies of scale*. How does this relate to returns to scale? Cite and briefly discuss the main determinants of economies of scale.
- Define *diseconomies of scale*. Cite and briefly discuss the main determinants of this phenomenon.

8. Define *economies of scope*. Is this concept related to economies of scale? Explain.
9. Explain the relationship between the learning curve and a firm's cost function. Would economists consider the learning curve a short-run or a long-run phenomenon?
10. Define the *experience curve*. Compare its impact on a firm's cost function with that of the learning curve.
11. "Because of economies of scale, it is sometimes more cost effective for a firm to operate a large plant at less than maximum efficiency than a small plant at maximum efficiency." Do you agree with this statement? Explain.
12. When a company states its financial results in its annual report, it typically presents its income statement in the following way:

	Revenue
–	Cost of Goods Sold (including some depreciation)
	Gross Profit
–	Selling, General, and Administrative Expenses
–	Research and Development
–	Depreciation
	Operating Profit
+ / –	Net Interest (income and expense)
	Net Profit before Income Taxes
–	Taxes
	Net Profit after Income Taxes

"Cost of goods sold" includes all costs directly associated with making a product or providing a service. In retail merchandising this cost is essentially the wholesale cost of goods sold.

Discuss the differences between cost of goods sold and the concept of relevant cost used in this chapter. Are there any situations in which selling, general, and administrative expenses, or research and development expenses might be considered as part of a firm's relevant costs? Explain.

13. Overheard at the water cooler: "I think our company should take advantage of economies of scale by increasing our output, thereby spreading out our overhead costs." Would you agree with this statement (assuming this person is not your boss)? Explain.

Refer to Appendix 7B for help in answering questions 14 through 16.

14. Discuss the estimation of short-run cost functions. Which regression method is most frequently used, and what are some of the problems a researcher will encounter? What adjustment factors may have to be employed?
15. Discuss the estimation of long-run cost functions. Which regression method is most frequently used, and what are some of the problems a researcher will encounter? What adjustment factors may have to be employed?
16. Comment briefly on the following methods of cost estimation:
 - a. Engineering costs
 - b. Survivorship principle
 Discuss the strengths and shortcomings of these methods and the circumstances under which each can be applied.

PROBLEMS

1. Based on your knowledge of the definition of the various measures of short-run cost, complete this table.

Q	TC	TFC	TVC	AC	AFC	AVC	MC
0	120	—	—	x	x	x	x
1	—	—	—	265	—	—	—
2	—	—	264	—	—	—	—
3	—	—	—	161	—	—	—
4	—	—	—	—	—	—	85
5	—	—	525	—	—	—	—
6	—	—	—	120	—	—	—
7	—	—	—	—	—	97	—
8	—	—	768	—	—	—	—
9	—	—	—	—	—	97	—
10	—	—	—	—	—	—	127

2. Mr. Lee operates a green grocery in a building he owns in one of the outer boroughs of New York City. Recently, a large chemical firm offered him a position as a senior engineer designing plants for its Asian operations. (Mr. Lee has a master's degree in chemical engineering.) His salary plus benefits would be \$95,000 per year. A recent annual financial statement of his store's operations indicates the following:

Revenue	\$625,000
Cost of goods sold	325,000
Wages of workers	75,000
Taxes, insurance, maintenance, and depreciation on building	30,000
Interest on business loan (10%)	5,000
Other miscellaneous expenses	15,000
Profit before taxes	\$175,000

If Mr. Lee decides to take the job, he knows that he can sell the store for \$350,000 because of the goodwill built with a steady clientele of neighborhood customers and the excellent location of the building. He would still hold on to the building, however, and he knows he could earn a rent of \$50,000 on this asset. If he did sell the business, assume he would use some of the proceeds from the sale to pay off his business loan of \$50,000. He could then invest the difference of \$300,000 (i.e., \$350,000 – \$50,000) and expect to receive an annual return of 9 percent. Should Mr. Lee sell his business and go to work for the chemical company?

In answering this question, also consider the following information:

- In his own business, Mr. Lee works between 16 and 18 hours a day, 6 days a week. He can expect to work between 10 and 12 hours a day, 5 days a week, in the chemical company.
 - Currently, Mr. Lee is assisted by his wife and his brother, both of whom receive no salary but share in the profits of the business.
 - Mr. Lee expects his salary and the profits of his business to increase at roughly the same rate over the next 5 years.
3. Joe enjoys fishing and goes out about 20 times per year. One day, his wife, Sarah, told him that fishing is simply too expensive a hobby. "I think that you should stop going fishing," she exclaimed. "I did a little calculation, and I figured that it costs us about \$28.75 for every

fish that you catch because you usually catch about 20 fish per trip. Besides, I always end up having to clean them. We would be much better off buying ready-to-cook fish from the local fish market.”

Comment on Sarah’s remarks. Do you agree with her argument? Explain. (Following below are her cost estimates.)

Boat (cost = \$30,000, usable for 10 years, 20 outings per year)	\$150
Boat fuel	45
Dock fees and insurance for the boat (average per trip)	130
Travel expenses to and from the lake (100 miles @ \$0.25 per mile: gas, oil, and tires, \$0.18, and depreciation and insurance, \$0.07)	25
New fishing equipment purchases this year (prorated over 20 trips)	25
Annual fishing license	35
Bait and miscellaneous expenses	50
Food	40
Beverages	35
Traffic fine received on the way to the lake	40
Total cost per trip	<u>\$575</u>

4. You are given the following cost functions:

$$TC = 100 + 60Q - 3Q^2 + 0.1Q^3$$

$$TC = 100 + 60Q + 3Q^2$$

$$TC = 100 + 60Q$$

- a. Compute the average variable cost, average cost, and marginal cost for each function. Plot them on a graph.
 - b. In each case, indicate the point at which diminishing returns occur. Also indicate the point of maximum cost efficiency (i.e., the point of minimum average cost).
 - c. For each function, discuss the relationship between marginal cost and average variable cost and between marginal cost and average cost. Also discuss the relationship between average variable cost and average cost.
5. Decide whether the following statements are true or false and explain why.
- a. A decision maker must always use the historical cost of raw materials in making an economic decision.
 - b. The marginal cost curve always intersects the average cost curve at the average cost’s lowest point.
 - c. The portion of the long-run cost curve that is horizontal indicates that the firm is experiencing neither economies nor diseconomies of scale.
 - d. Marginal cost is relevant only in the short-run analysis of the firm.
 - e. The rational firm will try to operate most efficiently by producing at the point where its average cost is minimized.
6. Indicate the effect that each of the following conditions will have on a firm’s average variable cost curve and its average cost curve.
- a. The movement of a brokerage firm’s administrative offices from New York City to New Jersey, where the average rental cost is lower
 - b. The use of two shifts instead of three shifts in a manufacturing facility
 - c. An agreement reached with the labor union in which wage increases are tied to productivity increases
 - d. The elimination of sugar quotas (as it pertains to those firms that use a lot of sugar, such as bakeries and soft drink bottlers)
 - e. Imposition of stricter environmental protection laws

7. You are given the following *long-run* cost function:

$$TC = 160Q - 20Q^2 + 1.2Q^3$$

- Calculate the long-run average cost and marginal cost. Plot these costs on a graph.
- Describe the nature of this function's scale economies. Over what range of output does economies of scale exist? Diseconomies of scale? Show this on the graph.

Refer to Appendix 7B for help in answering Problems 8 and 9.

8. During the last 50 years or so, many studies of cost curves have been published. The results of two of these are summarized here. In each case, interpret the equation and discuss the shapes of the total cost curve, the marginal cost curve, and the average cost curve.
- A study of a light plant over a 6-month period resulted in the following regression equation:

$$Y = 16.68 + 0.125X + 0.00439X^2$$

where Y = Total fuel cost

X = Output

- An early study of the steel industry indicated the following results, based on annual data for 12 years.

$$Y = 182,100,000 + 55.73X$$

where Y = Total cost

X = Weighted output, in tons

A time series analysis usually indicates a short-run cost study. Do you believe that a 12-year period is too long? Explain.

9. You have been presented with the following cost data and asked to fit a statistical cost function:

Quantity	Total Cost
10	104.0
20	107.0
30	109.0
40	111.5
50	114.5
60	118.0
70	123.0
80	128.5
90	137.0
100	150.0

- Plot the data on a graph, and draw a freehand curve that best fits the data.
 - Fit three possible statistical cost functions to the data. Use straight-line, quadratic, and cubic formulas. Do the results confirm the curve you drew in *a*?
 - Discuss the statistical results you obtained in *b*. Include in your discussion R^2 , the coefficients, and the statistical significance of the coefficients.
 - If the data represent 10 months of production for one plant of a specific company, would you consider this to be a short-run analysis?
 - How would your answer to part *d* change if you were told that the data represent 10 different plants during a particular month of the year?
10. The economist for the Grand Corporation has estimated the company's cost function, using time series data, to be

$$TC = 50 + 16Q - 2Q^2 + 0.2Q^3$$

where TC = Total cost

Q = Quantity produced per period

- a. Plot this curve for quantities 1 to 10.
 - b. Calculate the average total cost, average variable cost, and marginal cost for these quantities, and plot them on another graph.
 - c. Discuss your results in terms of decreasing, constant, and increasing marginal costs. Does Grand's cost function illustrate all these?
11. Discuss the following three cost functions:

$$TC = 20 + 4Q$$

$$TC = 20 + 2Q + 0.5Q^2$$

$$TC = 20 + 4Q - 0.1Q^2$$

- a. Calculate all cost curves:
 - Total cost
 - Total fixed cost
 - Total variable cost
 - Average total cost
 - Average fixed cost
 - Average variable cost
 - Marginal cost
 - b. Plot these curves on graphs.
 - c. Compare the shapes of these curves and discuss their characteristics. (Particularly interesting should be the last cost function, whose shape is often found in engineering cost studies.)
12. The Central Publishing Company is about to publish its first reference book in managerial economics. It is now in the process of estimating costs. It expects to produce 10,000 copies during its first year. The following costs have been estimated to correspond to the expected copies.

a. Paper stock	\$8,000
b. Typesetting	\$15,000
c. Printing	\$50,000
d. Art (including graphs)	\$9,000
e. Editing	\$20,000
f. Reviews	\$3,000
g. Promotion and advertising	\$12,000
h. Binding	\$22,000
i. Shipping	\$10,000

In addition to the preceding costs, it expects to pay the authors a 13 percent royalty and its salespeople a 3 percent commission. These percentages will be based on the publisher's price of \$48 per book.

Some of the preceding costs are fixed and others are variable. The average variable costs are expected to be constant. Although 10,000 copies is the projected volume, the book could sell anywhere between 0 and 20,000 copies.

Using the preceding data,

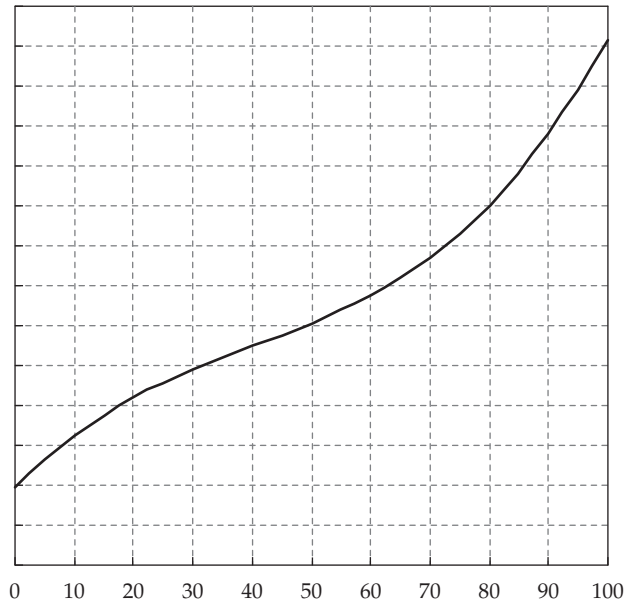
- a. Write equations for total cost, average total cost, average variable cost, and marginal cost.
 - b. Draw the cost curves for quantities from 0 to 20,000 (in intervals of 2,000).
13. The Big Horn Corporation commissioned an economic consultant to estimate the company's cost function. The consultant collected a large amount of data for a number of years from the books of the corporation and came up with the following equation:

$$TC = 170 + 22Q + 1.5Q^2$$

where TC = Total cost (in thousands)

Q = Quantity produced per period

- a. Plot this curve for quantities 1 through 15.
 - b. Calculate the average total cost, average variable cost, and marginal cost, and plot them on another graph.
 - c. Discuss your results in terms of decreasing, constant, and increasing marginal costs. Does Big Horn's cost function illustrate all these?
14. Use the total cost curve below to answer the following questions. (*Hint: Use a ruler.*)



- a. For what values of output is marginal product of labor decreasing? What is true about marginal cost over this range of output?
 - b. What is the approximate output level at which average variable cost is minimized? What is true about the average product of labor at this output level?
 - c. For what range of output does the average variable cost of production decrease? What can be said about marginal cost over this range of output?
 - d. For what range of output does average total cost of production decline?
- For e–g, suppose you know that minimum AVC is \$40.
- e. What is the approximate dollar amount of FC?
 - f. For what approximate Q is $AC(Q) = \$100$?
 - g. What is the approximate dollar amount of minimum AC?



Problems 15 and 16 are based on the Cobb Douglas Costs Excel App.

15. The Cobb Douglas Costs Excel App provides a nice link between the Cobb Douglas (CD) production analysis of Chapter 6 (see also Cobb Douglas Production Excel App) and the cost analysis in this chapter. The Cobb Douglas Costs Excel App depicts total and per-unit cost curves based on the CD production function, $Q = aL^bK^c$, given prices for labor and capital of w and r , respectively, for a variety of plant sizes. When initially opened, the graphs depict costs given $a = 1$, $b = 0.5$, $c = 0.5$, $w = 1$, and $r = 1$.
- a. What type of returns to scale exists in this instance?
 - b. Given equal weighted CD production, and equal factor prices, what does the long-run expansion path look like in terms of capital intensity, k (k is the number of units of capital per unit of labor)? In particular, what is k on the long-run expansion path $K = kL$.
 - c. Is $Q = 1$ produced at minimum cost in this instance by having $L = K = 1$? What is $LAC(1)$ in this instance?
 - d. Would LAC change if we considered producing $Q = 2$ by adjusting plant size accordingly?

16. Undertake a comparative statics analysis using the Cobb Douglas Costs Excel App file by filling in the following table. Increase each parameter by 50 percent holding the other parameters at their initial value (listed in Problem 15 above and in the table below) and describe what happens to each of the cost curves in both panels of the graphs on the Cobb Douglas Costs Excel App. Answer each question by describing whether the parameter change affects the *shape* of the function or its *placement*. (Placement means location on the graph.) In particular, answer the following two questions for each curve:
1. Does the power in the power function change as a result of this parameter change?

Hints:

- i. Look at the shapes of the curves to answer this question.
- ii. The LTC(1) ray, available using the click box in cell H27, is always a straight line and may help you see curvature.

If your answer to 1 is yes, Question 16.2 is:

2. Does the power become higher or lower? In this event, answer using the word *higher* or the word *lower*.

If your answer to 1 is no, Question 16.2 is:

2. Does the placement of the function change? In this event, you can describe the placement by referring to the points of the compass. For example, proportionally higher cost at every output level might be described as **N**, lower at every output would be **S**, and a stretching both upward and outward would be **NE**, and so on.

Note: The **N**, **S**, **E**, **W** descriptions are not unique, but they offer a ready way to describe the stretching of curves. For example, if a curve is a ray and remains a ray but becomes steeper, this could be viewed as either an **N** or a **W** movement.

Placement descriptions are perhaps most easily seen by focusing on a single LR/SR tangency and seeing how it changes as the parameter changes. For example, one might start from the LR/SR tangency for the second smallest plant size in the figure (the one with $K = 1$). The tangency in each panel occurs at (1,2) given the initial parameter values. What happens to this tangency in each panel as each parameter changes?

Parameter	Total Cost Curves		Per-unit Cost Curves					
	Set Value	Question:	LTC	STC	LAC	LMC	SATC	SMC
<i>a</i>	1	1						
		2						
<i>b</i>	0.5	1						
		2						
<i>c</i>	0.5	1						
		2						
<i>w</i>	1	1						
		2						
<i>r</i>	1	1						
		2						

A copy of this table is available on a separate worksheet on this file. You can fill it in electronically; then use Copy in Excel; then open Word and use Paste Special, Enhanced Metafile to transfer this to a Word document.

Provide short answers to the following two questions.

- a. Which of the five parameters altered returns to scale, and in which direction?
- b. Of the factors that did not alter returns to scale, what did they alter?



A Mathematical Restatement of the Short-Run Cost Function

The general form of the short-run cost function is

$$TC = f(Q) \quad (7A.1)$$

where TC = Total cost

Q = Output

As stated in this chapter, three specific forms of this function are used in economic analysis: the cubic, the quadratic, and the linear. Microeconomic theory relies primarily on the cubic equation because it encompasses the possibility of increasing returns to a factor as well as diminishing returns. The quadratic form of the total cost function implies that only the law of diminishing returns affects the short-run relationship between a firm's output and its variable input. The linear form indicates that neither increasing nor diminishing returns to a factor take place in the short run as the firm uses additional units of its variable input. In this appendix, we use calculus to analyze the cubic equation, the one most frequently used in economic theory, and show how Excel can help you graph per-unit cost curves in this instance.

For each cost curve algebraically described below, we provide the general form (i) as well as a specific example (ii). This example was the one used to generate the cost data in Table 7.2 and Figure 7.2. A cubic TC curve can be described by the four coefficients of a cubic function: a , b , c , and d .

$$TC = a + bQ - cQ^2 + dQ^3 \quad (7A.2i)$$

where a , b , c , and d are positive numbers and $c^2 < 3bd$.

Note the minus sign in front of the quadratic term c and plus signs in front of the linear and cubic terms. This is what produces U-shaped average and marginal cost curves. The second restriction, on the relative sizes of b , c , and d , is required to guarantee that $MC > 0$. A specific example of such a function is:

$$TC = 100 + 60Q - 5Q^2 + 0.7Q^3 \quad (7A.2ii)$$

The total fixed cost (TFC), a component of TC, is the constant term a , or 100. The balance of the

right-hand side of Equation (7A.2) is total variable cost (TVC). The various per-unit cost curves can be derived from TC using the definitions provided in the chapter.

Average fixed cost

$$AFC = a/Q = 100/Q \quad (7A.3)$$

Average variable cost

$$AVC = TVC/Q = b - cQ + dQ^2 \quad (7A.4i)$$

$$AVC = 60 - 5Q + 0.7Q^2 \quad (7A.4ii)$$

Average total cost

$$AC = AFC + AVC = a/Q + b - cQ + dQ^2 \quad (7A.5i)$$

$$AC = 100/Q + 60 - 5Q + 0.7Q^2 \quad (7A.5ii)$$

Marginal cost

$$MC = dTC/dQ = b - 2cQ + 3dQ^2 \quad (7A.6i)$$

$$MC = 60 - 10Q + 2.1Q^2 \quad (7A.6ii)$$

Even if you do not know calculus, it is easy to remember that MC is similar to AVC except that the linear and quadratic terms are multiplied by 2 and 3, respectively.

Economists often choose to graph AVC, AC, and MC but not AFC. This is because AFC is the difference between AC and AVC: $AFC = AC - AVC$. Therefore, we maintain our focus on MC, AVC, and AC.

An interesting rule relates minimum AVC and MC given cubic cost curves. This can be seen by taking the derivative of each curve, setting it equal to zero, and solving for Q . Let Q_M be the output level where MC is minimized, and Q_V is the output level where AVC is minimized.

$$dMC/dQ = -2c + 6dQ = 0 \quad \text{implies} \quad Q_M = c/(3d) \quad (7A.7)$$

$$dAVC/dQ = -c + 2dQ = 0 \quad \text{implies} \quad Q_V = c/(2d) \quad (7A.8)$$

Comparing Equations (7A.7) and (7A.8), we see that $Q_M = (2/3)Q_V$. One can readily see that this is true in the cubic per-unit cost figures in the chapter. One should note that finding minimum AC with derivatives requires solving a cubic equation. A faster solution for finding minimum AC is discussed below using Excel.

Graphing Per-Unit Cost Curves Using Excel



The per-unit cost curves are readily graphed and analyzed using Excel. The Per-Unit Costs Excel App provides a highlighted and annotated version of the steps described in this appendix and allows you to create graphs such as those in the text using Excel. A screenshot from that file is provided as Excel Sheet 7A.1 and the discussion below will reference specific cells within that worksheet (a second sheet, Excel Sheet 7A.2, shows the eight equations needed to create the graph and the three equations needed to find minimums for the curves in cells with a gray background).

If you are interested in graphing only one set of per-unit cost curves, you can simply type the three equations with numbers into Excel, referencing a runner for Q , then graphing the three curves using Scatter with Smooth Lines. A more flexible solution is to cell-reference the coefficients that form the data for the graph. That way, you can use the file for comparative statics analysis.

A useful starting point is to determine the value of Q_M because this provides an idea of how large the distance between each Q , ΔQ , should be in creating the graph. (The graphs used in the text, based on Equation (7A.2ii), suggest that $Q_M = 5/2.1 = 2.38$, a value that was chosen to allow easy description of tabular data.) For the present purposes, we will benefit from a more expansive TC function such as:

$$TC = 84700 + 1000Q - 15Q^2 + 0.1Q^3 \quad (7A.9)$$

The values of a to d are entered in A1:A4. In this instance, $Q_M = 15/0.3 = 50$ (as noted in A39). Given this, a Q runner with $\Delta Q = 10$ (cell A13) is used so that smooth curves are ensured given 21 values of Q from 0 to 200 in A15:A35. (If you find the ΔQ you initially have chosen does not provide satisfactory results, all you need to do is change that number and see if the graph looks better.) Create this runner by typing 0 in A15, then =A15+A\$13 in cell A16. Drag this down to A35, and the result will be 200 in A35.

The *coefficient bank* in B8:D11 minimizes errors in creating the graphs. This is the bank of coefficients associated with MC, AVC, and AC in Equations (7A.6i), (7A.4i), and (7A.5i). The first two are quadratic equations (parabolas), the third includes $AFC = a/Q$ which is why the equation in cell D8 is =A1 (notice that an equation is entered in this cell rather than typing in the number 84700). This equation can be dragged down to D11. The AVC coefficients are simply the VC coefficients reduced by a power of Q , so =A2 is entered in C9 and dragged to C11. The only tricky coefficients are for MC, as noted in Equation (7A.6i) above. The constant is =A2, the linear term is =2*A3 and the quadratic term is =3*A4.

By using Excel's mixed referencing method of cell referencing, you can create all three curves by typing in one equation in cell B15, then dragging it sideways and down. The only alteration that needs to be made is to add the AFC component of AC in D16. Both equations are described below. The main equation is MC(0) in cell B15:²⁷

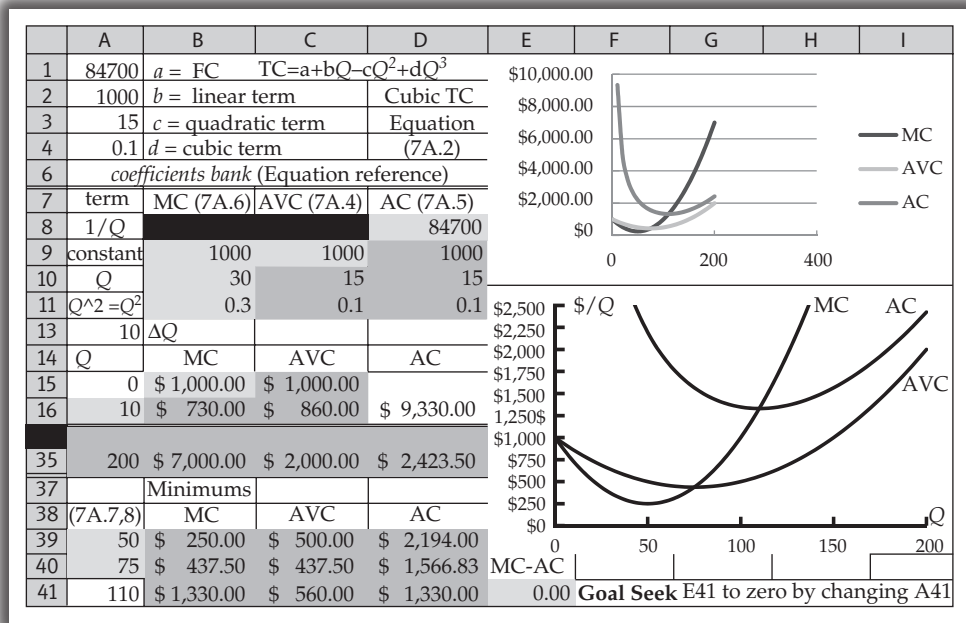
$$=B\$9-B\$10*\$A15+B\$11*\$A15^2 \quad (7A.10)$$

As described in the Introduction to Excel App, a \$ sign in front of a letter fixes the column reference (the letter), so the references to Q in column A are maintained as the equation is dragged sideways. A \$ sign in front of a number fixes the row reference (the number), so the references to coefficients in the coefficient bank are maintained as the equation is dragged down. Once Equation (7A.10) is entered in B15 (either manually or by clicking on cells and using the F4 key on the keyboard to fix the \$ signs in the appropriate location in each cell reference), drag B15 to C15, then drag both to B16:C16 (dragging multiple equations at once is shown by L-shaped arrows on Excel Sheet 7A.2). Next drag C16 to D16 and add AFC to AVC by adding +D\$8/A16 to the end of the equation (by clicking on the equation in the Formula Bar). Highlight B16:D16 and drag all three down to B35:D35.

Highlight A14:D35, click Insert, Scatter, Scatter with Smooth Curves, to obtain a preliminary version

²⁷It is worth noting that the $-$ sign in front of B\$10 would have been a $+$ sign had -15 been entered as the quadratic TC coefficient, c in the cell A3. This is worth keeping in mind because if you had obtained a , b , c and d from actual data via a regression model, your estimated coefficient for c would have a negative sign and the estimated TC function would be $TC = a + bQ + cQ^2 + dQ^3$. The TC function in Equation 7A.2 forced the minus sign for pedagogical purposes.

Excel Sheet 7A.1

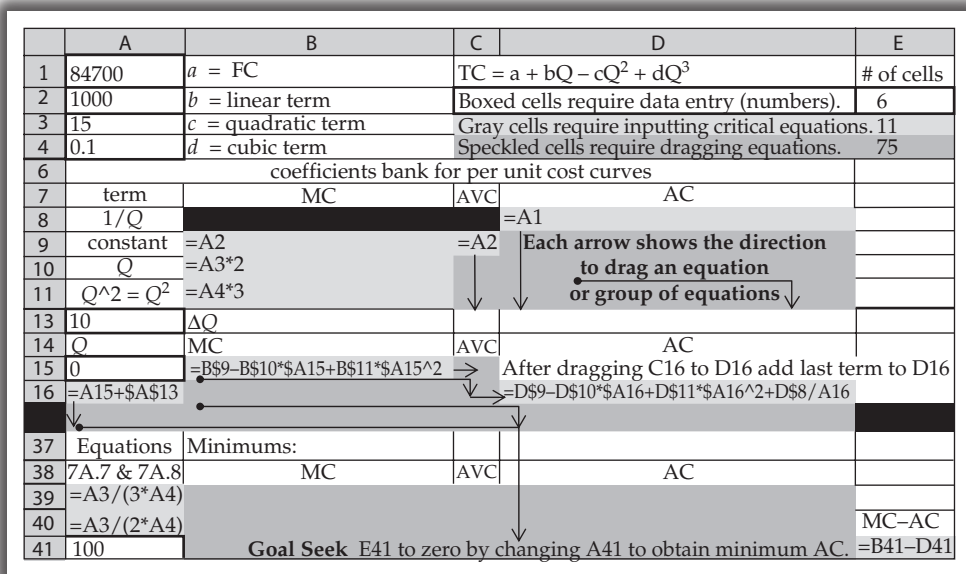


Source: Per-Unit Costs (Appendix 7A) Excel App, Main worksheet

of the graph similar to the top graph shown in Excel Sheet 7A.1. This can be cleaned up by formatting axes and adding titles to create the bottom graph shown in Excel Sheet 7A.1.

One task remains to be done: finding the minimums of each per-unit cost function. From above, we know in this instance that $Q_M = 50$ using Equation (7A.7). A useful strategy in order to

Excel Sheet 7A.2



Source: Per-Unit Costs (Appendix 7A) Excel App, Equations worksheet

have all three minimums in one location is to copy the three per-unit cost equations from the last row to start a new area for calculating purposes. To do this, highlight A35:D35 and click Copy. Then click on cell A39 and click Paste. Drag these equations down two times to A40:D40. The quantity values are wrong in A39:A41, but that is quickly remedied. Enter $=A3/(3*A4)$ in cell A39 [Equation (7A.7)] and $=A3/(2*A4)$ in cell A40 [Equation (7A.8)]. (By typing in the equation rather than typing 50 in A39, you will be able to see how Q_M changes as either c or d in A3:A4 change.) The results are as expected: $Q_M = 2/3Q_V$ (50 in A39 is $2/3$ of 75 in A40) and $AVC(Q_V) = MC(Q_V)$ (\$437.50 in B40 and C40).

The remaining task is to find minimum AC. As we have discussed extensively in this chapter, this occurs when $AC = MC$. The graph suggests that this occurs at about $Q = 100$, so type 100 in cell A40. (Make sure to type 100, do not type =100, because Goal

Seek, which is discussed below, works only on values, not equations.) The results show that $Q > 100$ is required because $MC(100) = \$1,000 < \$1,374 = AC(100)$ in B41 and D41. To find the intersection point, we suggest using a powerful search function built into Excel called Goal Seek. To use Goal Seek to find a point where two equations have equal values, type their difference in a new cell. In this instance, type $=B41-D41$ in E41. The difference, -374 , will show up in E41. If Q were appropriately chosen, then this value would equal zero. Click on cell E41, then click Data, What If Analysis, Goal Seek. The Goal Seek popup menu appears with E41 in the Set Cell area. Click on the To Value area and type 0, then click on the By Changing cell area and click on cell A41. Click OK twice, and 110 appears in A41 and \$1,330.00 appears in B41 and D41 (and 0 appears in E41). Minimum AC occurs at $Q_T = 110$, at which point $AC(Q_T) = \$1,330$.

The Estimation of Cost

Learning Objectives

Upon completion of this appendix, readers should be able to:

- Describe the use of time series regression analysis in measuring short-run costs, evaluating its advantages and disadvantages.
- Consider the question: are short-run marginal costs U-shaped or constant?
- Describe the use of cross-section regression analysis in measuring long-run costs, and evaluate its advantages and disadvantages.
- Consider the question: Are there economies of scale?
- Show how long-run costs can also be estimated by using engineering cost estimates and the survivorship technique.

The study of cost curves has its origins with Joel Dean, who wrote the first textbook on managerial economics and conducted many of the studies, dating back to the 1930s.

As in the case of production functions, we are interested in estimating cost functions both in the short run and in the long run. The short-run function helps define short-run marginal costs and thus assists the manager in determining output and prices. In the long run, the decision that a firm faces involves building the most efficient size of plant. That determination will depend on the existence of scale economies and diseconomies.

Short-run cost functions (like short-run production functions) assume at least one factor is fixed. Thus, cost is influenced by the quantity produced as changes occur in the variable factor. To estimate such a short-run function, we must find data in which quantity and costs change while certain factors change and others remain fixed.

In investigating short-run cost functions using regression analysis, researchers have most frequently employed the time series technique with data for a specific plant or firm over time. It is important that when time series data are collected, the period over which observations are taken is limited to a relatively short span because the size of the plant or firm, as well as technology, should not change significantly during the time interval used in a short-run cost function. To conduct a

meaningful analysis, there must be a sufficient number of observations, and there must be variations in production from observation period to observation period.²⁸ Thus, each observation period, where possible, should be limited to a month, and sometimes even a shorter period (1 week or 2 weeks).

Long-run cost functions—the planning functions—allow for changes in all factors, including plant size (or capital investment in general). Most studies of long-run cost functions have employed cross-sectional analysis. Observations are recorded in a specific time period (e.g., 1 year) for a number of different plants of different sizes with different amounts of inputs and outputs. We discuss short-term and long-term analyses separately.

The Estimation of Short-Run Cost Functions

Problems and Adjustments

Economic versus Accounting Costs Most empirical studies of cost functions have used accounting data that record the actual costs and expenses on a historical basis. However, decision-making data—economic data—should also include opportunity costs. No amount of adjustment will ever completely

²⁸Ideally, these changes in quantities produced should occur under relatively normal circumstances rather than due to some abnormal upheaval. Thus, production cuts because of plant damage or a strike, for instance, may not produce valid data.

reconcile these concepts, but certain corrections can be made.

- Changes in prices of labor, materials, and other inputs must be adjusted so current prices are used.²⁹
- Costs that are not a function of output should be excluded. Because we are dealing with short-run cost functions, fixed costs should not have an influence on pricing or output decisions.
- Closely allied to the previous point is the question of depreciation. Accountants usually record depreciation on a time-related basis. Depreciation is often not related to actual usage but follows accounting convention as adapted to tax rules. If “use” depreciation can be isolated from the accounting data, only that portion should be included in costs. But it must be kept in mind that recorded depreciation is based on the original cost of the equipment, whereas economic depreciation should be based on replacement value.

The problems that arise in the **estimation of short-run cost functions** because of differences between accounting and economic costs are the most difficult to solve. In most cases, some type of compromise is necessary. No definitive advice can be given here to prospective researchers. We can only point out the significant issues and suggest some possible appropriate adjustments. What is done in the final analysis depends on the data available and the ability of the investigator to make the corrections.

Rate Changes In addition to inflationary changes in the prices of various inputs, costs can also change due to variations in tax rates, Social Security contributions, labor insurance costs (unemployment insurance or worker’s compensation rates), and various benefit coverages that affect costs. Because most of these rate changes are not based on quantity produced, they should be excluded.

Output Homogeneity The problems encountered in cost estimation are similar to those discussed for the production function. The analysis is easiest when output is relatively homogeneous. If only one product is produced in the plant, the quantity produced (or shipped) can be handled in a rather uncomplicated manner. But if there are several products moving through the plant

simultaneously, some weighting apparatus must be employed to obtain the quantity produced.³⁰

Timing of Costs In many cases, costs and the service performed to create these costs do not occur at the same time. For instance, a machine that is in use continuously may be scheduled for maintenance periodically. In the airline industry, for example, major maintenance on engines is performed after a given number of flight hours. When such timing differences occur, care must be taken to spread the maintenance costs over the period of machine usage.

Accounting Changes When a time series analysis using accounting data is performed, it is important that the researcher ascertain whether changes in accounting methods, such as depreciation methods and recording of development expenses, have occurred during the period included in the study. Such changes must be adjusted to reflect uniformity in the measurements over time.

Given all the warnings about the problems that can be encountered in empirical cost estimation, you may have been persuaded that no useful conclusions can be obtained from such studies. Actually, such difficulties have not turned away many economists. Starting with Dean in the 1930s and continuing today, economic journals contain many articles investigating statistical cost functions.

The Shapes of Short-Run Cost Functions



Earlier in this chapter, three different specifications of cost functions were shown. Each represents a possible shape of the cost curves. The economist, after collecting and adjusting the data, will use one of these specifications to measure the relationship between cost and output. Other statistical functions could be employed (e.g., the Cobb-Douglas power function), but the three shapes are the ones most frequently encountered in statistical studies.

Figure 7.4*a* and *b* represent the normal theoretical function, which exhibits both decreasing and

²⁹Because different inputs may incur different relative price changes, some factor substitution may result over time. It can only be hoped that such effects will not be significant, because they are extremely difficult, if not impossible, to remove.

³⁰If a weighting scheme involving costs or direct inputs for each product is used, then, in a way, we are employing costs to determine output and then measuring costs as a function of this output. In other words, we are introducing a dependency into the relationship between costs and output, when we are really trying to determine the relationship between the two. This presents a serious problem. But again, as long as accounting and production records of a firm are employed, there is no easy way out of this dilemma.

increasing marginal and average costs. The cubic total cost function is³¹:

$$TC = a + bQ - cQ^2 + dQ^3$$

When estimating this equation, the analyst is not required to manually insert the $-$ sign in the model. The regression analysis will simply estimate $TC = a + bQ + cQ^2 + dQ^3$ and realize that the quadratic coefficient should have a negative sign.

If the data do not quite fit the cubic cost function, for example, if the estimated coefficient for the quadratic term in that function is positive, then a quadratic form can be tested:

$$TC = a + bQ + cQ^2$$

As can be seen in Figure 7.4c and d, the shapes of quadratic cost functions differ substantially from their cubic counterparts. The quadratic total cost curve only increases at an increasing rate. Thus, there are no decreasing marginal costs (increasing marginal product) in this construction. This can be seen in Figure 7.4d, where the MC curve is a straight line increasing at all points (i.e., not U-shaped).

A linear total cost function can also be fitted:³²

$$TC = a + bQ$$

Figure 7.4e and f show the curves and equations based on the linear form.³³ Note that the law of diminishing marginal returns has been eliminated. Each additional unit's cost is the constant, b .

³¹If only variable costs are estimated, the total curve would be the variable cost curve. Theoretically, it should begin at the origin because there are no variable costs when production is zero. However, when the curves are statistically estimated, even if only variable costs are included, the line will most likely intercept the Y-axis at a point other than zero. From the investigator's viewpoint, this is not terribly important because most of the observations included in the statistical estimate will not be anywhere near zero production. The intercept thus turns out to be meaningless. Alternatively, one can estimate variable cost by forcing the intercept to zero using any regression program. In Excel, this is accomplished by clicking Constant is Zero next to the Labels click box on the regression popup menu.

³²In Appendix B of Chapter 8, on break-even analysis, we use this function exclusively.

³³The average total cost curve in this case declines continuously and approaches the marginal cost curve asymptotically. The reason for this, in the formula, is that the first term a/Q decreases as Q increases and the second term, b , is a constant. If we were dealing exclusively with variable costs, where the a term does not exist (because costs are zero when production is zero), then $AVC = b$ and the marginal and average variable costs equal each other. This, of course, must be true, because if each additional (marginal) unit costs the same, then the average variable cost of each unit does not change and must equal the marginal cost.

Thus, this specification of the cost curve exhibits strictly constant marginal costs. This type of analysis does not appeal intuitively to an economist, who knows that if units of a variable factor are continually added to a fixed factor (plant), somewhere at higher production levels unit costs must rise.

These three specifications can fit various types of cost data. But even when a really good set of data is available to the investigator, chances are that the range of observations will be rather limited and will tend to cluster around a midpoint. Very seldom will we have data corresponding to near-zero production, and very seldom will data be given for production equaling or exceeding theoretical capacity. Thus, the statistical results economists obtain may not reflect the behavior of costs at the two extremes of the curve. In the following discussion of some of the empirical work that has been performed in the past, we find that this is precisely the case.

A Sampling of Short-Run Cost Studies

There is a large number of cost studies from which we could choose for purposes of illustration. We have selected three studies covering different industries and different time periods. Although some of the estimating procedures have become somewhat more sophisticated over time, the studies all used time series regression analysis.

The Hosiery Mill One of several studies conducted by Joel Dean in the 1930s and 1940s focused on a mill of a large hosiery manufacturer.³⁴ Its equipment was highly mechanized and its labor skilled.

Data for 54 months from 1935 to 1939 were employed. During this period, the size of the plant did not change, and the equipment remained approximately the same. Production during these months varied from zero to near physical capacity. Direct labor, indirect labor, and overhead costs were included. These costs were adjusted by factor price indexes. Output was a weighted index of individual products, and the weights were based on relative labor costs. The result, using a linear regression form, was

$$TC = 2,935.59 + 1,998Q$$

where TC = Total cost

Q = Output, in pairs of hosiery

³⁴Joel Dean, *Statistical Cost Functions of a Hosiery Mill*, Studies in Business Administration, 11, 4, Chicago: University of Chicago Press, 1941. This study and others have been reprinted in a volume of Dean's work, *Statistical Cost Estimation*, Bloomington: Indiana University Press, 1976.

The results were statistically significant, and the correlation coefficient was 0.973 (and, therefore, the coefficient of determination, R^2 , was 0.947).

Quadratic and cubic equations were also calculated but did not show a good fit to the data. Thus Dean's analysis points to the existence of a straight-line total cost curve, a decreasing average total cost curve, and constant marginal costs.

Road Passenger Transport In the United Kingdom, J. Johnston collected data from one of the larger transportation companies, with a fleet of some 1,300 vehicles, that operated some 45 million car miles per year.³⁵ The data were grouped into 4-week periods over 3 years, 1949 to 1952. Production exhibited a marked seasonal pattern, so varying production amounts per period could be observed.

The costs included all the vehicle operations (wages, clothing, gasoline, oil, and tires), maintenance, depreciation (based on mileage), variable overhead, and a minor amount of fixed overhead. Output was measured in car miles during the period. The data were adjusted by individual price indexes where price changes had occurred.

Again, a straight-line total cost function provided the best fit:

$$TC = 0.65558 + 0.4433Q$$

where Q represents car miles, in millions. The correlation coefficient was 0.95. So, as in Dean's study, the total cost curve was a straight line with constant marginal costs.

Plastic Containers In a study conducted for ten different products moving through a plant at the same time, twenty-one monthly observations were taken during the period of January 1966 to September 1967.³⁶ During this period plant capacity was fixed, and all input prices were also fixed by contract. Only direct costs of labor, machinery, and materials were included, so no allocation problems for overhead items had to be handled. The firm reported wide production fluctuations during the period, with several observations being somewhat above 90 percent of capacity level.

The straight-line function exhibited the best fit for each of the ten products. Tests to investigate

cost interrelationships among the products showed no influence.

The authors also tested for the firm's aggregate cost-output relationship. The total physical output was obtained by weighting each product by its base period price. Again, the best results were given by a straight-line equation,

$$TC = 56,393 + 3.368Q$$

with the t -test significant at the 1 percent level and the R^2 a respectable 0.89.

Marginal Cost: U-Shaped or Constant?

Each of the three studies just cited concluded that in the short run, the total cost curve is a linear function of production and that the marginal cost is constant. In addition to the hosiery mill, Dean also studied a furniture factory and a leather belt shop. Johnston conducted investigations into a food processing firm and coal mining, among others. Studies of the steel industry, cement industry, and electric power were also done. Most of these came to conclusions similar to those of the three studies we described briefly. Actually, some of the researchers found decreasing marginal costs to be the rule.

Does that mean that economists should revise their thinking about U-shaped average and marginal cost curves? Although these findings should cause economists to pause and do some additional thinking, the results of these studies—even over the relatively long period of 50 or more years—can be reconciled with the traditional shapes of cost curves:

1. The data employed in most of the studies concentrated on output levels that were limited in range. Thus, even though in some cases production may have taken place at about 90 percent of capacity, it is quite possible that plants are built and equipped in such a way that unit costs are relatively constant over a fairly long range of outputs. The cost curves that economists draw, as we have seen in this chapter, show very distinct minimum points to show their importance to students of elementary economics. However, the bottom of the curve may represent a fairly wide interval, and unit costs may rise (quite sharply) only when physical capacity is reached. Figure 7B.1 shows an average cost curve (A) that most students have seen drawn on blackboards. But curve B may be the one that is more true to life. The minimum point of this curve is not as low as that on curve A , but there is a relatively wide range of outputs on curve B where costs remain low—lower than for curve A . As production fluctuates from period to period, output is more likely to take place on the relatively constant

³⁵J. Johnston, *Statistical Cost Analysis*, New York: McGraw-Hill 1960, pp. 74–86.

³⁶Ronald S. Coot and David A. Walker, "Short-Run Cost Functions of a Multi-Product Firm," *Journal of Industrial Economics*, April 1970, pp. 118–28.

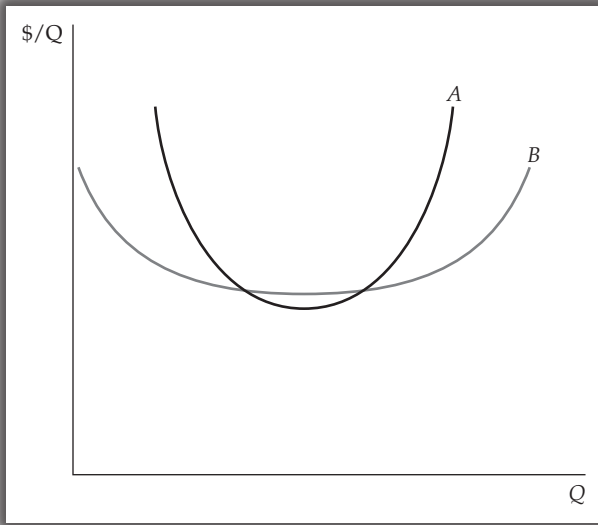


Figure 7B.1 Theoretical Versus Realistic Average Cost Curve

unit cost segment on curve *B* than at curve *A*'s narrowly defined minimum.

2. Many economists explain straight-line cost curves by pointing out that, although theory requires capital inputs to be fixed in the short run, they really are not. For instance, when production increases, a firm can easily set up an additional assembly line to keep the fixed/variable input ratio constant. Thus, because the fixed factor is used in a fixed proportion with the variable factor, increasing marginal cost will not necessarily occur.
3. Regression analysis is not a perfect tool. If most of the observations fall into an intermediate output range, there may just not be a sufficient number of observations at the extremes to convert a linear fit to a curvilinear one.

Thus, although economists would certainly be more gratified if empirical work were to confirm their theories, the evidence of the studies is not so overwhelming that an immediate reevaluation of microeconomic theory is necessary. Further, real-life data seem to be consistent with the existence of upward-sloping marginal cost curves. We learned earlier that the upward-sloping short-run supply curve of a firm or industry is based on the existence of an upward-sloping marginal cost curve. The fact that rising demand for a product raises the price and causes an increase in quantity supplied in the short run is evidence of the existence of an upward-sloping supply curve that is theoretically explained by the existence of an upward-sloping marginal cost curve. Thus, despite some of the empirical findings,

it appears that corporate managers act as if they were confronted by a marginal cost curve of the kind described by economic theory.

The Estimation of Long-Run Cost Functions

The estimation of long-run cost functions presents some new challenges. Remember, in the economic long run, all costs are variable. That means that capital (plant and equipment), which is ordinarily held fixed in the short run, is now permitted to change. As a matter of fact, it is precisely the goal of long-run cost analysis to trace unit costs for different sizes of plant and different amounts of equipment capacity. Our interest in long-run costs is to investigate the existence of returns to scale.

The question of costs in relation to size is important not only for planning decisions regarding plant or firm expansion. Firms must also consider costs when deciding on potential mergers. Will the synergy of the new units indicate a potential decrease in unit costs as a function of size?

Unlike the short-term cost function studies that use time series analysis, most long-run cost studies employ cross-sectional regression analysis. This method has several advantages:

1. Observations are recorded for different plants (or firms) at a given point of time. Because different plants generally come in different sizes, the independent variable, quantity of output, can vary over relatively large ranges.
2. Because all observations are taken at a given point in time (e.g., 1 year), the technology is known and does not change. Under ideal circumstances, each plant will use that level of technology, within the known state of the art, that is most efficient for that specific plant. But this ideal is usually not reached, as is explained in the following discussion.
3. Adjusting the various costs for inflation or other price changes is not necessary under cross-sectional analysis. If annual observations are used for the sample of plants, an average figure for costs (e.g., per labor hour or ton of material) will be satisfactory. Only under conditions of very severe inflation—hyperinflation—would the averaging of annual costs become a significant problem, but any studies undertaken under such conditions would surely be of dubious quality in any case.

The use of the cross-sectional technique also creates some difficulties of which the researcher must be well aware. Just a few of these are the following:

1. Inflationary problems may be avoided by abandoning the time series method, but a new problem is

introduced. Because the observations are taken from plants (or firms) in different geographic areas, we may encounter interregional cost differences in labor rates, utility bills, material costs, or transportation costs, for example. Adjustments to some common base must be made. But if relative prices of inputs differ, then the combination of inputs in a particular location may be a function of these relative differences. By adjusting prices to a specific benchmark, we may be obliterating the conscious choice made by management in a particular location based on relative prices of different inputs.

2. Although, as stated, the known state of the art for each firm is the same, it is not necessarily true that all plants are, at a given point of time, operating at the optimal level of technology. The assumption that each plant is operating most efficiently for its production level does not necessarily hold. This can be illustrated by drawing the familiar envelope curve. In Figure 7B.2 an envelope curve, ABC , is drawn. This long-run average cost curve represents all the optimal points of production. Point B is the minimum cost point, which represents the most efficient plant size. If, however, the plants in the sample do not operate at optimal points for any level of production, we may end up with observations along a curve such as DF , which will tend to be above the optimal curve, ABC , and which indicates not only a level, but also a slope quite different from the optimal cost curve. There really is no way to correct for this potential error completely.
3. If accounting data are used to make the estimates, it is important that the economist verify that there

are no significant differences in the way costs are recorded on the firms' books of account. Different depreciation procedures, varying inventory valuation methods, and amortization of other expenses can lead to substantial distortions. The best possible adjustments must always be made.

4. Different companies may pay their factors differently. This is most important in the case of labor costs. Total labor costs can comprise varying fractions of wages and benefits (e.g., vacation time, holidays, and medical care). In some cases, employees may receive part of their income in stock, or stock may be sold to them at a discounted price. The estimator must be certain to include all relevant costs.

A Sampling of Long-Run Cost Studies

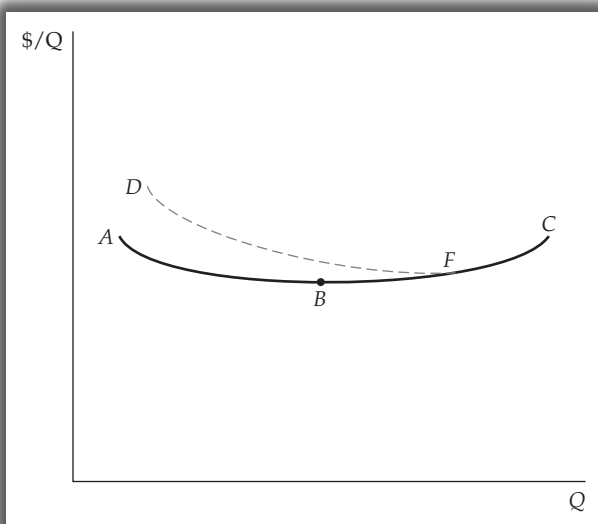
As in the case of short-run costs, a large number of inquiries into the properties of long-run costs have been conducted during the last 70 years or so. A small number of these are mentioned here to provide a flavor of method and results.

Shoe Store Chain In one of the early studies conducted by Joel Dean, the sample comprised 55 metropolitan area shoe stores owned by one firm.³⁷ Pairs of shoes were the output. Costs were mainly composed of selling expenses (including wages), handling expenses, and building expenses. Corporate allocations were omitted. The study was based on annual data for 1937 and 1938. Quadratic total cost curves gave the best fit for both years, indicating a U-shaped average cost curve and a rising marginal cost curve.

Building Societies In the United Kingdom, building societies are roughly equivalent to savings and loan associations in the United States. Johnston studied 217 such enterprises in the year 1953.³⁸ The companies were grouped into six size categories based on total annual revenue.

Johnston found that a U-shaped long-run cost curve existed for these enterprises. Cost in this study was measured as a ratio of management expenses to revenue. Middle-size societies were found to have the lowest ratio. However, when Johnston segregated the 217 observations into 149 societies with no branches and 68 with branches, somewhat

Figure 7B.2 Optimal and Suboptimal Cost Curves



³⁷Joel Dean and R. Warren James, *The Long-Run Behavior of Costs in a Chain of Shoe Stores: A Statistical Analysis*, Studies in Business Administration, 12, 3, University of Chicago Press, 1942, reprinted in Dean, *Statistical Cost Estimation*, pp. 324–60.

³⁸Johnston, *Statistical Cost Analysis*, pp. 103–5.

different results appeared. Most of the firms in the former group were concentrated in the four lower-size categories and most firms in the latter group were in the top four. When the nonbranch societies and the branch societies were analyzed separately, the former firms appeared to have a declining cost ratio, whereas the companies with branches had a constant ratio. Thus, the appearance of a U-shaped average cost curve, and therefore, economies of scale at the low end and diseconomies of scale at the high end, were the result of two separate underlying patterns.

Electric Utilities The electric utility industry has been a frequent target of economists studying cost curves. Output is fairly easily defined (electricity produced), and the United States has a large number of independently owned utilities of many different sizes. Because these companies must regularly deal with state regulatory commissions in rate cases, they collect large volumes of data pertaining to costs and revenues. Utilities have also been observed to have constant productivity improvements and technological progress. The study summarized here was performed using 1971 data for 74 utilities.³⁹ Because the data are for 1971, this study rules out any effects of pollution control equipment on costs.

The major costs were best described by a quadratic function, which led to a U-shaped average cost curve. The cost coefficients were statistically highly significant. The authors identified the intervals for which economies of scale existed and the appropriate minimum cost point. They found that diseconomies of scale appeared beyond moderate firm size. Many earlier studies of utilities found economies of scale throughout the range of observations. This study, as well as another one for the year 1970,⁴⁰ found the possibility of scale diseconomies in firms of larger size.

A more recent study analyzed economies of scale for thirty-one regional private steam-electric utilities located in five Midwestern states in 1987. The author of this study found that significant economies of scale existed for twenty-one of the smaller utilities producing 32 percent of the total regional output.

The larger firms showed no additional economies. The long-run cost curve for these firms was flat, and no diseconomies of scale were found.⁴¹

Financial Institutions A more recent article reviewed some thirteen studies that attempted to estimate economies of scale and economies of scope for credit unions, savings and loan associations, and commercial banks.⁴² Economies of scale are defined as those associated with firm size, whereas economies of scope relate to the joint production of two or more products. Economies of scale exist if per-unit or average production costs decline as output rises. Economies of scope arise if two or more products can be jointly produced at a lower cost than is incurred in their independent production. Each study used a logarithmic function and employed similar measures of economies. The author summarized the results of these thirteen studies as follows:

- Overall economies of scale appear to exist only at low levels of output, with diseconomies at larger output levels.
- There is no consistent evidence of global economies of scope.
- There is some evidence of cost complementarities (product-specific economies of scope).
- The results appear to exist generally across the three types of institutions studied, as well as across different data sets and product and cost definitions.

Another study looked at the cost functions of the largest 100 commercial banks in the United States, ranging from \$2.5 to \$120.6 billion in assets in 1984. The authors found the minimum efficient scale to range between \$15 billion and \$37 billion in assets. The results appeared to contradict some of the earlier studies that found economies of scale in banking to disappear at much smaller sizes.⁴³

A recent study appears to confirm previous findings. The authors find that the banking industry is characterized by economies of scale for smaller banks. This is followed by horizontal average costs

³⁹David A. Huettner and John H. Landon, "Electric Utilities: Scale Economies and Diseconomies," *Southern Economic Journal*, April 1978, pp. 883–912.

⁴⁰L. R. Christensen and W. H. Greene, "Economies of Scale in U.S. Electric Power Generation," *Journal of Political Economy*, 84, 4 (August 1976), pp. 655–76.

⁴¹Albert A. Ogunade, "Economies of Scale in Steam-Electric Power Generation in East-North-Central U.S.," *Journal of Economics and Finance*, 17, 1, (Spring 1993), pp. 149–56.

⁴²Jeffrey A. Clark, "Economies of Scale and Scope at Depository Financial Institutions: A Review of the Literature," *Economic Review*, Federal Reserve Bank of Kansas City, September/October 1988, pp. 16–33.

⁴³Sherrill Shaffer and Edmond David, "Economies of Superscale in Commercial Banking," *Applied Economics*, 23, 2 (February 1991), pp. 283–93.

over a fairly wide range. The largest banks begin to face diseconomies.⁴⁴

Dairy Farms A recent study, using data from 2000, estimates returns to scale across U.S. dairy farms of different sizes and over different regions. The findings show that there are substantial economies of scale across all sizes of farms, and they confirm the reason for increases in the average size of farms. The results of this study appear not to show an area of decreasing returns to scale, unlike those found in previous research.⁴⁵

Two Studies of Cost Curves in Switzerland One study analyzed economies of scale in the Swiss electricity distribution industry using a cross-sectional sample of 39 publicly owned companies serving 130 cities over the period 1988–1991. The results found economies of scale to be present for small, medium-sized, and large utilities. Using several models, the study estimated that the cost elasticity of output was less than one. The measure indicated that a 1 percent increase in the delivery of power increased variable costs by 0.86 percent. The author concludes that findings point to a possible decrease in costs of electricity distribution through a consolidation of local utilities.⁴⁶

A second study measured the economies of scale in postal services in Switzerland. The sample consisted of 47 small local post offices in 2001. Data on total annual operating costs, hourly wages, and the price of capital (approximated by rental price per square meter) were obtained. The results showed that cost elasticities with respect to output were positive, implying economies of scale. A 1 percent increase in production (represented by number of collected mail items) increased total costs by 0.2 percent, while a 1 percent increase in distributed mail increased total costs by 0.6 percent. Most of the post offices in the sample were relatively small and did not reach the minimum efficiency scale. Thus, in this case again, the authors suggest that efficiency could be increased by combining post offices serving the same or adjacent areas.⁴⁷

⁴⁴J. L. Stimpert and Judith A. Laux, “Does Size Matter? Economies of Scale in the Banking Industry.” *Journal of Business and Economic Research*, Vol. 9, Issue 3, March 2011, pp. 47–55.

⁴⁵Roberto Mosheim and C. A. Knox Lovell, “Scale Economies and Inefficiency of U.S. Dairy Farms,” *American Journal of Agricultural Economics*, Vol. 91, Issue 3, August 2009, pp. 777–794.

⁴⁶Massimo Filippini, “Economies of Scale and Utilization in the Swiss Electric Power Distribution Industry,” *Applied Economics*, Vol. 28, Issue 5, May 1996, pp. 543–550.

⁴⁷Massimo Filippini and Marika Zola, “Economies of Scale and Cost Efficiency in the Post Services: Empirical Evidence from Switzerland,” *Applied Economic Letters*, Vol. 12, Issue 7, June 2005, pp. 437–441.

Bottling Plants Some years ago, the authors received access to some confidential cost and production data for more than thirty bottling plants of a large company. The plant sizes and their production varied widely; thus the data for a specific time period lent themselves to a cross-sectional regression analysis. It was our aim to investigate the potential existence of an optimum plant size.

A straight-line total cost function obtained fairly good results. The R^2 equaled 0.62 (quite high for a cross-sectional analysis) and the slope coefficient was statistically significant. Surprisingly, when a cubic function was fitted, the results improved considerably. With the equation $TC = a + bQ + cQ^2 + dQ^3$, we obtained the following answers:

	$R^2 = 0.70$		$F\text{-statistic} = 24.87$	
	A	B	C	D
Estimate	0.078	0.891	-0.096	0.004
t -statistic		3.653	-2.240	1.910

The t -statistic for coefficients b and c are significant, for d , slightly low. The c coefficient is negative, indicating that the total cost curve would first increase at a decreasing rate and then at an increasing rate—resulting in a U-shaped unit cost curve.

Another piece of data was available—the percentages of canned and bottled sodas. Because the cost of cans was lower than the cost of bottles, we included the percent of canned soda production as another independent variable, with the following results:

	$R^2 = 0.83$		$F\text{-statistic} = 37.45$		
	A	B	C	D	E
Estimate	-0.177	1.316	-0.146	0.006	-1.575
t -statistics		6.358	-4.229	3.597	-4.823

All t -statistics are now significant and coefficient e , the percentage of can production, has the correct (negative) sign, showing that production costs decrease as the proportion of cans increases.

Granted that we used only data for one period and that we did not refine these data, but we still obtained some good results indicating a U-shaped average cost curve and an optimal plant size.

Empirical Long-Run Cost Studies: Summary of Findings Some of the studies just summarized show the possible existence of diseconomies of scale

for larger firm and plant size. However, a majority of the empirically estimated functions suggest the existence of scale economies up to a point. As output increases to a substantial size, these economies rapidly disappear and are replaced by constant returns for long intervals of output. Most studies have not found the existence of declining average costs at very large quantities. Findings of diseconomies of scale for high rates of production are also rather rare.

Two Other Methods of Long-Run Cost Estimating

At the beginning of the discussion of cost estimating, we issued a warning that accounting magnitudes may not correspond to economic measures. However, all the regression analyses just reviewed employed accounting data and, through various adjustment procedures, attempted to make the numbers correspond more closely to the cost concepts meaningful to the economist. In addition, economists have tried to use methods that are not dependent on numbers drawn from companies' accounting records. Two techniques of analysis are discussed briefly: the engineering cost method and the survivorship technique.

Engineering Cost Estimates Engineering costs are based on a thorough understanding of inputs and outputs and their relationships. Knowledgeable professionals will calculate the quantity of inputs needed to produce any quantity of outputs. These calculations are based on optimal assumptions (i.e., the largest output for a given combination of inputs). This is really a production function. From here it is a relatively easy step to apply monetary quantities to the inputs to arrive at costs.

The advantages of such a method are quite obvious. Technology is held constant. There are no problems of inflation. Problems of changing output mix are eliminated. As a matter of fact, such calculations are often made by corporations planning to introduce a new product. While the market research and product forecasting departments concentrate on preparing estimates of sales at different prices, cost estimators in a corporation (engineers and others) prepare estimates of costs for different levels of output. Then members of the pricing department take all the available data and calculate profitability at different prices and levels of output. It is through this process that product prices are obtained.

Although the **engineering cost estimating** technique avoids some of the pitfalls of regression

analysis, it suffers from some problems of its own. First, the estimates represent what engineers and cost estimators believe cost should be, not necessarily what they actually will be. Because they are really forecasts, calculations may omit certain components that contribute to costs. Further, most of the time, only direct output costs are estimated. Other costs that may be directly associated with the product (e.g., some portion of overhead and direct selling expenses) are not included, or if they are, it is through some rather arbitrary allocation. Often such estimates are made on the basis of pilot plant operations and do not consider actual production, which may be attended by bottlenecks and other problems, causing costs to differ from those estimated.

Generally, engineering cost estimates will show declining unit costs up to a point and substantially flat unit costs at higher production quantities. The possible existence of diseconomies of scale is usually ignored.

It is quite possible for a study to combine engineering cost estimates for some segments of costs with utilization of accounting data for others. Such an analysis was performed some years ago, when a group of economists at the Transportation Center at Northwestern University published a forecast of aircraft prices.⁴⁸ One important step in arriving at the forecast was the calculation of aircraft operating costs. These were computed for a large number of different types of aircraft and used both actual data and engineering estimates. Crew salaries were calculated from a number of union contracts and compared with historical data to arrive at an average. Maintenance costs were calculated from a formula supplied by the Air Transport Association. Fuel consumption was obtained from engineering curves produced by aircraft manufacturers and airlines, and reconciled with some actual published data. Other costs (employee benefits, landing fees, liability and property damage insurance) were estimated in similar ways. The resulting curves reflected direct operating costs per aircraft mile for a series of different stage length categories ranging from 0 to 200 miles to 2,500 miles and more. The resulting cost curves generally exhibited a downward slope up to the limit of the stage length that each aircraft could fly nonstop.

⁴⁸Stephen P. Sobotka, Paul G. Keat, Constance Schnabel, and Margaret Wiesenfelder, *Prices of Used Commercial Aircraft*, 1959–1965, Evanston, IL: Transportation Center at Northwestern University, 1959.

Survivorship Principle A prominent American economist and Nobel Prize winner, George J. Stigler, developed an intriguing method for estimating long-run costs.⁴⁹ Stigler believed that the use of accounting data, with all their distortions and subsequent need for adjustments, made the validity of cost estimation based on such data questionable. His method was to observe an industry over time, categorize the firms in the industry by size (measured as a percent of total industry capacity or output), and then arrive at a conclusion regarding cost efficiency based on the relative growth or decline of these size categories. His results for the steel industry (using data for 1930, 1938, and 1951) showed that medium-size firms (defined as between 2.5 percent and 25 percent of industry capacity) appeared to have gained in share of total industry output over the 21-year period, from 35 percent to 46 percent of total, whereas small firms (less than 2.5 percent of capacity) and large firms (actually just one firm, with more than 25 percent of capacity) lost market share. Stigler concluded the existence of a U-shaped long-run average cost curve whose path first showed net economies of scale, then constant returns, and finally diseconomies of scale.

Economists have continued to use the **survivorship technique** to investigate economies of scale. For example, a later study by R. P. Rogers examined economies of scale in steel manufacturing in the United States. The author measured the distribution of production in conventional steel mills in 1976 and 1987. He classified steel mills into four size categories by their annual capacity (1–1.49 million tons, 1.5–4.5 million tons, 4.5–7.5 million tons, and more than 7.5 million tons). He then examined these four categories for the proportion of total capacity in the 2 years. He found that the 4.5 to 7.5 MT category increased its market share considerably. The results were consistent with a previous study by D. G. Tarr, who had used the engineering cost approach to estimate the minimum optimal size (MOS) of a steel mill. Tarr had found the MOS plant to be of 6 MT capacity, which is the midpoint of Rogers' findings.⁵⁰

⁴⁹George J. Stigler, "The Economies of Scale," *Journal of Law and Economics*, 1, 1 (October 1958), pp. 54–81.

⁵⁰R. P. Rogers, "The Minimum Optimal Steel Plant and the Survivorship Technique of Cost Estimation," *Atlantic Economic Journal*, 21 (September 1993), pp. 30–37; D. G. Tarr, "The Minimum Optimal Scale Steel Plant in the Mid-1970's," *FTC Working Paper*, 3, March 1997; D. G. Tarr, "The Minimum Optimal Scale Steel Plant," *Atlantic Economic Journal*, 12, 2 (1984), p. 122.

A study of commercial health insurers found that the smallest-size firm category declined in numbers of firms, percentage of all firms, and percentage of total premium volume in the period 1958 to 1973. All but the smallest firms appeared to have held their own, and the largest did not appear to gain at the expense of their smaller rivals.⁵¹

A later study of the property-liability insurance industry appeared to show increases in the smallest and largest companies. The industry was divided into seven classes. All classes except the smallest and largest showed decreasing shares of output during the study period.⁵²

Another study looked at the concentration in the trucking industry since deregulation. The industry was divided into six classes based on ton-miles hauled. The year 1975 was chosen as a benchmark to compare to the years 1981–1993, which followed deregulation. The results show that the two smallest classes and the largest class lost share of the business. The three middle-size classes gained about 23 percent of the trucking business. The author concludes that "the competition created by deregulation has to some extent disturbed the quiet life of monopoly carriers operating outside the efficient range."⁵³

A more recent study looked at the market shares of U.S. banks using the survivorship technique. Much research has shown that scale and scope economies in the banking industry have been small. But during the period from 1984 to 2002, the largest banks appeared to have increased their share significantly, while the shares of banks in smaller categories have declined. The author ascribes this trend to economies of integration due to an important extent to deregulation in the banking industry. Thus, while this trend cannot be explained by economies of scale, "one possible explanation is that the benefits to bank size accrue from revenue economies rather than cost economies . . ." Further, "size may allow banks to benefit from an implicit

⁵¹Roger D. Blair and Ronald J. Vogel, "A Survivor Analysis of Commercial Health Insurers," *Journal of Business*, 51, 3 (July 1978), pp. 521–29.

⁵²David Appel, John D. Worrall and Richard J. Butler, "Survivorship in the Property/Casualty Insurance Industry," *Journal of Risk and Insurance*, Vol. 52, no. 3, September 1985, pp. 424–440.

⁵³James N. Giordano, "Returns to Scale and Market Concentration Among the Largest Survivors of Deregulation in the US Trucking Industry," *Applied Economics*, Vol. 29, Issue 1, January 1997, pp. 101–110.

too-big-to-fail subsidy or to reduce risk by achieving greater diversification.”⁵⁴

The survivorship principle is intuitively appealing due to its simplicity and avoidance of unreliable data. However, it also suffers some serious limitations. It is of no help in measuring costs for planning purposes. It merely tells us which company size appears to be more efficient; it says nothing about relative costs. Further, it implicitly assumes

the industry is highly competitive so survival and prosperity are solely a function of efficient use of resources and not of market power or the erection of barriers to entry. Changing technology and inflation over a long span can also cause distortions. As time passes, the structure of an industry can change in such a way that firms of a certain size are favored over others.

Moreover, although Stigler’s analysis of the steel industry showed a U-shaped cost curve, he could not obtain similar results for the automobile industry, where the survivor cost curve showed declining and constant portions, but there was no evidence of diseconomies of scale at high production quantities.

⁵⁴Timothy J. Yeager, “Economies of Integration in Banking: An Application of the Survivor Principle,” *Federal Reserve Bank of St. Louis, Supervisory Analysis Working Papers*, Working Paper 2004-04, March 2004.

SUMMARY

Accounting data have generally been used to investigate short-run and long-run cost functions. These data present the researcher with a host of problems because the economic and accounting definitions of costs can differ substantially. Also, depending on how the data are collected, adjustments for price changes, geographic differentials, and other variations must be made.

Time series analysis has largely been used to estimate short-run costs, whereas the cross-sectional regression technique appears more suited for long-run cost estimation.

A large majority of these studies have concluded that marginal cost in the short run is relatively constant. In the long run, economies of scale predominate at the low end of production, and at higher output, constant returns to scale appear to exist.

The upward-sloping—U-shaped—average and marginal cost curves postulated by economic theory tend to be the exception in empirical findings. Although such results should make economists pause and reexamine some of their theoretical conclusions, the studies have generally been conducted in such a way that the possibility of eventually rising marginal costs and diseconomies of scale cannot be discounted.

Some economists prefer not to use accounting data in their inquiries. Thus, two other methods of cost estimation were briefly described in this appendix. Engineering cost analysis is based on the expert knowledge of the relationship between inputs and outputs and standard costs. It avoids the use of accounting information and does not run into the problem of adjusting for changing technology and inflation. The survivorship method bases its findings on the change in the proportion of total industry output produced by firms of different size categories. It concludes that the more efficient firms will gain share of production at the expense of less efficient ones.

Interview with a Supply Chain Management Executive

Introductory Note: As already discussed in this chapter (see pp. 273–276), supply chain management is a critical function for today’s manufacturing companies. It is a function that they count on to reduce costs in today’s highly competitive global economy. In this section, Stephen C. Martson, shares his experiences and the insights he gained while working in supply chain management for some of the world’s leading information technology companies.⁵⁵

PHIL: Tell us about your professional background. How you get into supply chain management?

STEVE: I have a law degree and started out in a law firm after graduation. Shortly afterward, in the early eighties, I joined AMP Incorporated (now Tyco Electronics). One of my internal clients was the purchasing organization. I worked with them in negotiating and writing agreements with our suppliers.

PHIL: Isn’t it a bit unusual for a lawyer to go into supply chain management?

STEVE: Not really. First, way back in the eighties, purchasing was not considered as important a function as it is today. In fact, I think the term *supply chain management* was just starting to be used.⁵⁶ Also, there were no formal courses of study in the business and engineering curricula to prepare one for a career in this function.

⁵⁵The authors are very grateful to Steve for agreeing to be interviewed by one of us (Phil Young) for this text. Steve recently retired from the corporate world and is currently an adjunct professor of international business and management at Dickinson College.

⁵⁶An article in Wikipedia about supply chain management (see http://en.wikipedia.org/wiki/Supply_chain) notes the following as an early example of the use of this term: Oliver, R. K., Webber, M. D., 1982, “Supply-chain management: logistics catches up with strategy,” Outlook, Booz, Allen and Hamilton Inc. Reprinted 1992 in *Logistics: The Strategic Issues*, ed. M. Christopher, London: Chapman Hall, pp. 63–75.

People in the purchasing function came from all kinds of educational backgrounds and work experiences.

PHIL: I guess that was before “purchasing” became today’s “procurement.”

STEVE: You might say that! But in any case, I think law is a very good background to have in procurement and supply chain management. Legal training is useful because it teaches one to think and approach problems in a very logical and process-oriented way.

PHIL: But your résumé is a veritable showcase of technology companies that have helped to develop supply chain management into what it is today: Dell, IBM, Compaq [now part of HP], and Flextronics. How did you go from AMP to Dell?

STEVE: Well I believe I must have been doing a reasonably good job at AMP because a few years into the job, I was contacted by a headhunter who said that a company named “Dell Computer Corporation” was looking to fill the position of head of purchasing. I went through various rounds of interviews and finally ended up being interviewed by Michael Dell. I was told the interview went very well, but unfortunately he decided that I didn’t have enough of a technology background and an understanding of the PC business. But four months later, they came back and offered me the position of vice president of procurement. The year was 1993.

PHIL: By coincidence, I remember that 1993 was the year that Compaq launched a low-priced notebook called the “Presario” in order to counter the intense price competition from the no-name “PC clones.”⁵⁷

⁵⁷For an excellent history of the rise and eventual sale of Compaq (to HP), see <http://en.wikipedia.org/wiki/Compaq>.

STEVE: Yes, thanks to intense competition from companies who competed solely on price (rather than brand or differentiation), the day of the 40 to 50 percent gross profit margins for the PC notebook were rapidly starting to become a thing of the past. I remember one of the first decisions that Dell Computer made when I came onboard was to outsource the design of its notebook in order to save on costs and speed up the development and introduction of new notebooks with the latest technology.

PHIL: Yes, I think any student of managerial economics would be well served to read about the history of the PC business. In fact, I believe Compaq was the first company to offer a truly IBM-compatible PC on a large scale. So in a sense, one might say that Compaq itself was a clone, an “IBM PC clone.” But anyway, once Compaq had its lower priced Presario, didn’t it in fact force out of business well-known companies such as Packard Bell? Moreover, didn’t it also surpass IBM in total PC sales around this time?

STEVE: I believe that’s true. But Compaq used a combination of marketing and lower prices to increase sales and improve market share. And, of course, another big part of its competitive advantage and success was its supply chain management. In any case, let me get back to the situation at Dell since my particular stint at Compaq came several years after my work at Dell and IBM.

PHIL: Okay, Steve, sorry for getting us a bit off track.

STEVE: No problem, it’s all good. Anyway, you might say that I was there at the beginning of Dell’s strategic outsourcing and global supply chain initiatives [e.g., supplier-owned inventory warehouses near the final assembly factories], the key core competencies that helped make it into the company it is today. I had a very good relationship with Michael Dell. In 1996, I traveled extensively with him evaluating various locations for our first factory in Asia. We finally decided on setting up shop in Penang, Malaysia, mainly because

the infrastructure of PC suppliers was already there. We thought about China, but in those days it was far too risky and, of course, China was nowhere as developed as it is today.

PHIL: That made sense. I believe Intel had and still has manufacturing, research, and some shared services operations in Penang.

STEVE: That’s right. In fact, Penang is the place that Intel chose for its first offshore assembly operation way back in 1972. It also continues to be the location for factories making PCs for Dell and other companies.⁵⁸ Getting back to my work in the PC business, after about four years at Dell, I was contacted by the chief procurement officer at IBM, whom I had known for a number of years. He convinced me to go to IBM to help in its efforts to improve its supply chain management systems in its PC company. By the way, the head of the PC company at that time was Sam Palmisano, who eventually became IBM’s CEO.⁵⁹ As the head of procurement for IBM’s PC company, one of the most important things that I did was to gain the support of the business unit leaders to outsource our design and manufacturing for a number of our products to several Taiwanese companies. Up to that point, IBM Japan was doing most of the design work on the notebook products.

PHIL: This is the second time that you’ve mentioned the outsourcing of design work. I think most people think about outsourcing mainly in the area of manufacturing. For example, Apple is very careful to put on the back of all its products: “Assembled in China. Designed in California.”

STEVE: This may be the case with Apple, but in the PC market, as well as in many other consumer electronics products, both the design and manufacturing work are outsourced for a significant amount of the products.

⁵⁸See <http://www.intel.com/jobs/malaysia/sites/penang.htm>

⁵⁹Sam Palmisano was IBM’s CEO from 2003 to 2011. He was succeeded by IBM’s first woman CEO, Virginia Rometty starting on January 1, 2012.

Moreover, I'd ask your textbook readers to do some searching on their own on the Internet for discussions about the extent to which Apple itself designs its products versus contracting this work to outside parties [e.g., the original iPod was designed and manufactured by a Taiwanese company].

At this point, I should clarify the use of three basic acronyms used in supply chain management and strategic outsourcing.

OEM: This stands for "original equipment manufacturer." This is the company that puts its name brand on the final product and sells it to end users, either directly or through distributors. In the supply chain, it is primarily responsible for sourcing, cost reductions, supplier qualifications and quality, continuity of supply, and meeting the company's strategic objectives including providing a competitive advantage with the management of its supply chain.

ODM: This stands for "original design manufacturer." This is a company that is contracted by an OEM to design as well as manufacture the final product. Many OEMs do a combination of this and their own designs.

CM: This stands for "contract manufacturer." This is the company that is contracted by an OEM to assemble the final product using the OEM's design.

PHIL: Okay, if I understand you correctly, would you say that in the PC business, Dell, HP, Lenovo, and Apple are OEMs. Quanta is an ODM, and Foxconn and Flextronics are CMs?

STEVE: That sounds right.

PHIL: Life must be pretty easy when an OEM finds a good ODM. What is there for the OEM to do except tell the ODM what it wants and then go out and sell it?

STEVE: Yes and no. It's certainly true that when a company has a solid alliance with a supplier that designs as well as makes the final product, there can be certain advantages pertaining to cost, technology leadership, and speed to market. But don't think for a moment that this arrangement is a kind of "skunk works" where a bunch of people at the OEM sit

around dreaming up ideas and then passing off all the "heavy lifting" to the ODM. The OEM must have its own product strategists, design consultants, manufacturing engineers, marketing staff, and supply chain managers to make sure that the right products are developed for the right markets in the right way all over the world. Don't forget: The OEM bears all the responsibility for the product's performance. If something goes wrong with your Dell computer, Dell cannot say to you that it was the ODM's fault.

PHIL: Thanks for the clarification, Steve. By the way, I noticed you used the term *alliance* rather than the commonly used term *partnership* when referring to the relationship between the OEM and the ODM, and with suppliers.

STEVE: Yes, that's just one of my personal preferences and maybe that's a bias because of my legal background. I think the term *partnership* is a bit too strong. Also, I never liked the term *vendor*, particularly in the high-tech business. I think the term *supplier* is much more appropriate. The term *vendor* seems to conjure up images of wholesale produce companies selling fruit to supermarkets.

PHIL: I believe that after you worked for IBM for a couple of years, you then moved to Compaq?

STEVE: Yes, in 1998, Compaq decided to buy the Digital Equipment Company, better known as DEC. I was hired to help rationalize the supply chain management system of the combined companies. I'll be the first to admit that this was not a very pleasant assignment. As you know, a key reason for a merger or acquisition is to reduce costs through economies of scale. What this really means is that some people have to lose their jobs. Both Compaq and DEC had well-established supply chain management organizations, and one of my jobs was to slim down the combined operations of the two companies after they became one.

PHIL: Yes, we discuss mergers and acquisitions and economies of scale in various sections of our textbook. I guess you're reminding

us of the human dimension to all our graphs and theory.

STEVE: That's one way to put it. In any event, one of Compaq's major CMs was a company called Flextronics, based in California but with manufacturing operations in about 150 locations around the world. In working with the Flextronics people, it soon became evident to them that they themselves needed a better supply chain management operation. After all, much of what a CM does is assemble parts that are made by *their* suppliers using the OEM's design of the product.

PHIL: Okay Steve, don't tell me: Flextronics made you an offer you couldn't refuse?

STEVE: You could put it that way, Phil. I found myself on the other side of the fence, so to speak. Moreover, I couldn't resist the exciting challenge of helping to build up a supply chain management organization pretty much from scratch. Also around that time, the cell phone market really started to take off and there was increasing outsourcing work being done in this industry. One of the big contracts that Flextronics won was with Sony-Ericsson. I really enjoyed the challenge of helping to set up and manage that supply chain relationship.

PHIL: When a company decides to outsource some or all of its manufacturing, is it mainly a matter of cost savings?

STEVE: Yes, that may be the initial reason. But it becomes quickly apparent that supply chain management becomes just as critical as cost savings, sometimes even more. Let me explain. No matter what it costs to make a product, it doesn't matter if you can't sell it. Would you agree?

PHIL: Yes, that sounds like a reasonable statement. Putting it in a broader economic context, what good is supply if there is no demand?

STEVE: Well, that's a good point, but what I was thinking was more along the lines of this. Once a company has outsourced its manufacturing to an outside contractor, it is absolutely essential that the contractor be able to meet the demand for the company's products. *Continuity of supply* and *time to market* are absolutely critical.

PHIL: Oh, I get it. I guess one could say, What good is demand without supply? To that point, I've heard that OEMs definitely don't like to be "put on allocation." Is this part of the challenge of managing the supply chain with outside suppliers?

STEVE: Yes, indeed. It is typical for a contract manufacturer to be doing business with a number of OEMs who themselves are competing with each other to sell the same type of product in the market. When the demand for that product is really high, the contract manufacturer may have to allocate the supply among its competing OEM customers. Some of these OEMs may not get all the product that they ordered.

PHIL: So in summary, what would you say are the keys to successful supply chain management?

STEVE: First, good demand forecasting is important. But, as you know, forecasting is never perfect. So the second element is flexibility of supply. The supply chain manager must know how to speed up or slow down various stages of the production process to adjust to shifting demand [e.g., supplier-owned inventory warehouses can be critical to ensuring this flexibility]. Third, if some or all of the production is outsourced, the supply chain manager must have a good working relationship with the suppliers. Thus, fundamental elements such as good communications and trust are the keys. In other words, no matter how sophisticated the technology to monitor the supply chain, the human element is always important. Fourth, establish alliances with suppliers that will ensure technology leadership and cost reductions that provide a competitive advantage to your company.

PHIL: Thank you very much for sharing your experiences and insights with us, Steve. I believe you've given our readers some useful stories and examples to help them understand and appreciate the importance of supply chain management and managerial economics.

Pricing and Output Decisions: Perfect Competition and Monopoly

Learning Objectives

Upon completion of this chapter, readers should be able to:

- Describe the key characteristics of the four basic market types used in economic analysis.
- Compare and contrast the degree of price competition among the four market types.
- Provide specific actual examples of the four types of markets.
- Explain why the $P = MC$ rule leads firms to the optimal level of production in competitive markets.
- Describe what happens in the long run in markets where firms are either incurring economic losses or are making economic profits. Explain why this happens with particular attention to the key assumptions used in this analysis.
- Explain how and why the $MR = MC$ rule helps a monopoly to determine the optimal level of price and output.
- Explain the relationship between the $MR = MC$ rule and the $P = MC$ rule.
- (Appendix 8A) Use the tools of calculus to determine the optimal level of output for a firm in perfect competition and the optimal levels of price and output for a monopoly.
- (Appendix 8B) Explain how the fixed cost and variable cost help to determine a firm's break-even level of output.
- (Appendix 8B) Explain the concept of operating leverage and how it can be used to better understand the challenges faced by firms with a high fixed cost relative to their variable cost.

The Situation

For the job of product manager for the new energy drink, “Go-nuts,” Nicole Goodman, SVP of marketing, did not want a new MBA fresh out of the company’s management training program. There was simply too much at stake to entrust the job to a rookie. She firmly believed that the job of bringing the new product to market should be given to a seasoned manager with a proven record of accomplishments. Finding a person with the necessary qualifications from within the company was not going to be easy because all successful product managers were rapidly promoted to higher management positions. Nicole could have used a “headhunter” firm to find an outside person, but she preferred to give someone within the company a chance to make the project work.

“I’ve got just the person for you,” the executive vice president exclaimed at lunch one day. “There is a real sharp manager over in market research. I think he was the one who did the background study on the soft drink market. It would be nice to give him a chance to put his ideas into practice.”

The person recommended for this critical job was none other than Frank Robinson, the head of the forecasting department. (See Chapter 5.) In his first meeting with Nicole after being hired, Frank was briefed on the job. “Because this new product is so important to the growth strategy of our firm,” Nicole said, “and because of your experience and accomplishments, we decided that we wanted you rather than one of the outside people who were considered for the job.

“One of the first tasks you should pursue is an analysis of the optimal price of this beverage. Tell us what price we should charge to maximize our profit in this new venture. The CEO told us the other day that the Wall Street analysts were questioning our judgment in getting into such a crowded and highly competitive market. We need to prove as fast as possible that we made the right decision, so we want to maximize our profit in this venture in as short a time as possible.

“As always, the management committee has the final say on the price of the new beverage, just as for all our products and services. But don’t get discouraged. As I know from the other products that I’ve managed, pricing is a very useful exercise because it forces one to bring together all the different elements of the business. The market research that you have already done on this product will provide you with a quantitative estimate of the demand, as well as a general competitive analysis of the entire beverage industry. Our production people and cost accountants will give you the cost estimates. It will be up to you to put everything together to arrive at a suitable price for “Go-nuts.”

INTRODUCTION

Readers who love to cook may well appreciate the fact that in Chinese cuisine one of the most difficult tasks is the preparation of the ingredients. Once the meats and vegetables are properly sliced and all the spices are secured, everything is thrown into the wok in a systematic and timely fashion, and usually within a minute or two the dish is ready. In a sense, we have been “slicing and dicing” over the past seven chapters and at last we are ready to cook. We have taken you through the definition of economics and managerial economics, the goals of the firm, the market forces of supply and demand, the various types of demand elasticity (own price, cross price, and income), the estimation and forecasting of demand, and the key factors underlying supply (production and cost). We now combine the relevant knowledge from these previous chapters to address one of the most important questions in managerial economics: How do firms establish their prices and output levels in order to achieve their business objective of profit maximization?

The pricing and output decision will actually be answered within the framework of four basic types of markets: perfect competition, monopoly, monopolistic competition, and oligopoly. The distinguishing characteristics of each of the four market types are presented in Figure 8.1*a*. This chapter deals with pricing and output decisions in perfect competition and monopoly. The pricing and output decisions of firms operating in monopolistic competition and oligopoly markets are analyzed in Chapter 9. Perfect

Perfect Competition (no market power)

1. Large number of relatively small buyers and sellers
2. Standardized product
3. Very easy market entry and exit
4. Nonprice competition not possible

Monopoly (absolute market power subject to government regulation)

1. One firm, firm is the industry
2. Unique product or no close substitutes
3. Market entry and exit difficult or legally impossible
4. Nonprice competition not necessary

Monopolistic Competition (market power based on product differentiation)

1. Large number of relatively small firms acting independently
2. Differentiated product
3. Market entry and exit relatively easy
4. Nonprice competition very important

Oligopoly (market power based on product differentiation and/or the firm's dominance of the market)

1. Small number of relatively large firms that are mutually interdependent
2. Differentiated or standardized product
3. Market entry and exit difficult
4. Nonprice competition very important among firms selling differentiated products

Figure 8.1a The Four Basic Market Types

competition and monopoly can be considered the two extreme market environments in which a firm competes in terms of **market power**. The concept of market power was introduced in Chapter 3. But as a reminder, market power is simply the power of a firm to establish the price of its products. In **perfect competition**, there are so many sellers offering the same product that an individual firm has virtually no control over the price of its product. Moreover, there is no way for a particular firm to charge a higher price than its competitors because everyone sells a standardized product. Instead, the interaction of supply and demand decides the price for all participants in this type of **market structure**. A firm in this market has no market power and acts only as a **price taker**. All it can do is to decide whether to compete in the market and how much output to produce. In direct contrast to firms in perfect competition, the monopoly firm has a considerable amount of market power. Because it is the only seller in this type of market, it has the power to establish the price at whatever level it wants, subject to possible constraints such as government regulation. It is the consummate **price maker**.

In terms of market power, monopolistic competition and oligopoly are somewhere between the two extremes of perfect competition and monopoly. From a pedagogical standpoint, it is easier to understand and appreciate the particulars of monopolistic competition and oligopoly if there is first a thorough understanding of perfect competition and monopoly. This explains why we first cover perfect competition and monopoly in this chapter and the other two market types separately in the next one. Before proceeding to our first case of pricing and output decisions in perfect competition, let us elaborate further on market structure and the meaning of competition in economic analysis.

COMPETITION AND MARKET TYPES IN ECONOMIC ANALYSIS

The Meaning of Competition

In economic analysis, the most important indicator of the degree of competition is the ability of firms to control the price and use it as a competitive weapon. The extreme form of competition is “perfect” competition. In this market, the competition is so intense and the firms are so evenly divided that no one seller or group of sellers can exercise any control over the price. That is, they are all price takers. A second key measure of competition in economic analysis is the ability of a firm to earn an “above normal” or “economic” profit in the long run. This concept is described in greater detail in the next chapter. (Refer to Chapter 3 for a definition of the “long run.”)

Figure 8.1*b* is a reformatted version of Figure 8.1*a*. It shows the four market types according to the degree of competition as indicated by the extent of market power and the ability of firms to earn long-run economic profit. A firm in monopolistic competition may have some market power because its product can be differentiated from those sold by its competitors. A firm operating in an oligopoly derives its market power from its ability to differentiate its product, its relatively large size, or both.

Market entry and exit most directly affects the ability of a firm to earn economic profit in the long run. In perfect competition, entry into the market is easy. Therefore, if firms are observed to be earning economic profit, over time the entry of new firms eager to partake in these profits quickly reduces the ability of both incumbents and new entrants to earn economic profit. The same applies to monopolistic competition. This in fact helps us understand the basis for its name. This market is deemed to be *monopolistic* because product differentiation enables firms to exercise some market power (i.e., to act as price makers). In contrast, it is *competitive*

Figure 8.1*b* Comparison of Four Market Types by Characteristics Affecting the Degree of Competition

Market Characteristics	Market Type			
	Perfect Competition	Monopolistic Competition	Oligopoly	Monopoly
Number and size of firms	Very large number of relatively small firms	Large number of relatively small firms	Small number of relatively large firms	One
Type of product	Standardized	Differentiated	Standardized or differentiated	Unique
Market entry and exit	Very easy	Easy	Difficult	Very difficult or impossible
Nonprice competition	Impossible	Possible	Possible or difficult	Not necessary
Key Indicators of Competition				
Market power	None	Low to high	Low to high	High
Long-run economic profit	None	None	Low to high, subject to mutual interdependence	High, subject to regulation

because over time the entry of new firms reduces and ultimately eliminates any economic profit.

In the case of oligopoly, the size and/or the ability to differentiate a product provide firms with considerable market power. Furthermore, because it is relatively difficult to enter this market, there is more opportunity to earn economic profit for an extended period of time. It should be quite obvious why a monopoly is not competitive from the standpoint of economic analysis. Being the only firm, it has the ultimate power to set its price, and its ability to earn economic profit is only constrained by government regulation or perhaps by the introduction of breakthrough technologies or substitute products from completely different industries.

Nonprice competition plays a secondary role in determining the degree of competition in economic analysis. However, we recognize that nonprice factors often come to mind first when people think about how firms compete with one another. Nonprice competition involves firms trying to gain an advantage over one another by differentiating their products using such means as advertising, promotion, the development of new products and product features, and customer service. For example, when we consider the competition between Coke and Pepsi, we may think about the amount of money each spends to have well-known entertainers endorse their products. When we consider the battle between IBM and Sun Microsystems in the computer hardware market, we might consider the speed and reliability of each company's line of servers.

The extent to which buyers and sellers have information about the price of the product and the product itself (e.g., product quality, reliability, and integrity) can also be a factor in determining a firm's market power or competitive advantage. For example, have you ever bought an item at a store only to find out the next day that you could have bought it for less at another store? Your first reaction might have been anger at the first store for charging you more. But a store can do this if there are people like you who do not have complete information about the selling prices of the item wherever it is available.¹ Moreover, incomplete information about the product itself may well lead to the exercise of market power among sellers. A general overview of this situation is presented in Chapter 11.

Examples of Market Types

Perfect Competition

The markets for agricultural products (e.g., corn, wheat, coffee, pork bellies), financial instruments (e.g., stocks, bonds, foreign exchange), precious metals (e.g., gold, silver, platinum), and the global petroleum industry provide good examples of this type of market. In each market, the products are standardized commodities, and supply and demand are clearly the primary determinants of their market price.² Of course, it is precisely because of this that sellers sometimes form cartels in order to raise the price or to keep it from falling. OPEC and the International Coffee Growers Association are good examples of this. As explained in Chapter 3 and later in Chapter 10, these sellers band together to control price by restricting the supply of their product.

¹Online shopping has helped to make markets more competitive because it makes it easier for potential buyers to compare prices among the sellers.

²For an excellent review of market conditions in commodities, see "Oversupply and Slackened Demand Vex Commodity Markets," *New York Times*, January 2, 2002.

Monopoly

Examples of a pure **monopoly** are not easy to find. Some years ago, good examples could be found among government-sanctioned and regulated monopolies in the telecommunications and gas and electric industries. But these markets have been deregulated and opened up to competition over the past several decades, not just in the United States, but all over the world. However, local-government-sanctioned monopolies exist in many communities, and these monopolies provide services such as garbage collection and local electricity distribution. Patent laws sometimes provide companies with temporary monopolies. The pharmaceutical industry can definitely be said to earn economic profit during the time in which its products are protected by patents. There are certain industries in which a company is so dominant that it might be said to exhibit monopolistic status. The dominance of Microsoft in PC operating systems has led to its antitrust problems. In certain situations, a firm may enjoy monopoly status because of its particular surroundings. Simple examples come to mind, such as the only Brazilian barbecue restaurant (*churrascaria*) in town or a “last chance” gas station at the edge of the Mojave desert.

Monopolistic Competition

Small businesses, particularly retail and service establishments, provide the best examples of this kind of market. Among them are boutiques, luggage stores, shoe stores, stationery shops, restaurants, repair shops, laundries, and beauty parlors. There are many of them in any given city or area of the city. The start-up capital is relatively low, so it is fairly easy to start these types of businesses. Each one tries its best to stand out among its many competitors by differentiating its product. A Chinese restaurant may attempt to differentiate itself by offering a cuisine from a relatively unknown region of China. A dry cleaner may try to distinguish itself by keeping longer store hours or by having its service clerks greet every regular customer by name as he or she enters. If customers perceive these kinds of differences to be important enough, these retail establishments may be able to charge a higher price than their competitors.

Oligopoly

The oligopoly market is generally considered to be the playing field of big businesses. In the United States, a large segment of the manufacturing sector competes in oligopoly markets. For example, in the manufacturing sector, oil refining, certain types of computer hardware and software, chemicals and plastics, processed foods, tobacco, steel, automobile, copper, and soft drinks can all be considered oligopoly markets. Parts of the service sector also contain good examples of oligopoly markets. For example, airline travel, long-distance telecommunications service, and Internet access are dominated by a relatively small number of large companies. The names of these oligopoly firms are a prominent part of any listing of large companies based on the size of total revenue such as the *Fortune* 500 (American companies) and the *BusinessWeek* Global 100.

Market Types and Competition in Theory and Reality

The four different market structures discussed previously are intended by economists to be a theoretical framework in which to analyze the pricing and output decisions of the firms in these markets. As is the case with all theoretical constructs, the relationship among these four distinct market types and the actual market conditions may vary. Certain markets fit well into these market types. Others may not exhibit all four characteristics of a particular market type. Still others may evolve over time from one

type to another. These differences between theory and reality and the possible blurring of distinctions from one market type to another are discussed more thoroughly in Chapter 9. So now we proceed to examine pricing and output decisions using the four idealized types of markets as a conceptual framework, starting with perfect competition.

PRICING AND OUTPUT DECISIONS IN PERFECT COMPETITION

The Basic Business Decision

Imagine a firm that is considering entry into a market that is perfectly competitive. If it decides to compete in this market, it will have no control over the price of the product. Therefore, the firm's managers must make a business case for entering this market on the basis of the following questions:

1. How much should we produce?
2. If we produce such an amount, how much profit will we earn?
3. If a loss rather than a profit is incurred, will it be worthwhile to continue in this market in the long run (in hopes that we will eventually earn a profit) or should we exit?

Perhaps even the output decision may seem superfluous. After all, is not the firm so small that it can sell as much as it wants without affecting the market price? Yes, but although the market price does not vary with an individual firm's level of output, the *unit cost* of production most certainly does. Think back to our discussion in Chapter 7 about the cost of additional units of output. If we assume marginal cost rises as output increases (thanks to the law of diminishing returns), then it seems reasonable to expect that eventually the extra cost per unit will exceed the selling price of the product. At this point, it would no longer make sense for a profit-maximizing firm to produce because each additional unit sold would cost the firm more to produce than the price at which it could sell the product. Much more will be said about this shortly. But the point to emphasize here is that there is indeed a limit as to how much a perfectly competitive firm should produce in the short run. It is up to the firm to determine this limit.

Because the perfectly competitive firm must operate in a market in which it has no control over the selling price, there may be times when the price does not fully cover the unit cost of production (i.e., average cost). Thus, a firm must assess the extent of its losses in relation to the alternative of discontinuing production. In the long run, a firm that continues to incur losses must eventually leave the market. But in the short run, it may be economically justifiable to remain in the market, with the expectation of better times ahead. This is because in the short run certain costs must be borne regardless of whether the firm operates. These fixed costs must be weighed against the losses incurred by remaining in business. It is reasonable to expect that a firm will remain in business if its losses are less than its fixed costs—at least in the short run.

Key Assumptions of the Perfectly Competitive Market

As you are well aware, it is critical to know the assumptions made in the development of an economic model. Let us summarize the key assumptions made in analyzing the firm's output decision in perfect competition.

1. The firm operates in a perfectly competitive market and therefore is a price taker.
2. The firm makes the distinction between the short run and the long run.

3. The firm's objective is to maximize its profit in the short run. If it cannot earn a profit, then it seeks to minimize its loss. (See Chapter 2 for a review of the goals of a firm.)
4. The firm includes its opportunity cost of operating in a particular market as part of its total cost of production.

All four assumptions have been discussed earlier, some in greater detail than others. But it might be useful to review certain aspects of these assumptions before proceeding to numerical and graphical examples.

For the economic analysis of a firm's output and pricing decisions to have a unique solution, the firm must establish a single, clear-cut objective. This objective is the maximization of profit in the short run. If the firm has other objectives, such as the maximization of revenue in the short run, the output that it would select would differ from the one based on this model. (See Chapter 10 for a discussion of the Baumol revenue maximization model.)

The consideration of opportunity cost in the cost structure of the firm is vital to this decision-making model. The firm must check whether the going market price enables it to earn a revenue that covers not only its out-of-pocket costs, but also the costs incurred by forgoing alternative activities. A brief numerical example should help to convey this point.

Suppose the manager of a "stop and shop" convenience store wants to own and operate a store of her own. She knows she will have to leave her job and use \$50,000 of her savings (currently invested and yielding a 10 percent return). A statement of the projected cost of operating in the first year follows.

Cost of goods sold	\$300,000
General and administrative expenses	150,000
Total accounting cost	\$450,000
Forgone salary for being a store manager	45,000
Forgone returns from investments (10% return)	5,000
Total opportunity cost	\$50,000
Total economic cost (total accounting cost plus total opportunity cost)	\$500,000

To keep this example as simple as possible, we did not include depreciation and taxes.

Suppose this budding entrepreneur forecasts revenue to be \$500,000 in the first year of operation. From an accounting standpoint, her profit would be \$50,000 ($\$500,000 - \$450,000$). But from an economic standpoint, her profit would be zero because the revenue would just equal her total **economic cost**. Certainly, there would be nothing wrong with "breaking even" in the economic sense of the term because this indicates that the firm's revenue is sufficient to cover both its out-of-pocket expense and its opportunity cost. Another way to view this situation is to note that when a firm "breaks even" in the economic sense, it is actually earning an accounting profit equal to its opportunity cost. In other words, if this entrepreneur's annual revenue is \$500,000, she will earn an accounting profit that offsets the opportunity cost of going into business for herself. In economic terms, she would be earning a **normal profit**.

The reason for using the term *normal* can be seen in situations in which the entrepreneur's revenue is higher or lower than \$500,000. Suppose her revenue is \$550,000. In this case, she will earn a profit of \$50,000 ($\$550,000 - \$500,000$). We refer to this sum as "above normal," "pure," or **economic profit** because it represents an amount in excess of the out-of-pocket cost plus the opportunity cost of running the business.

Table 8.1 Normal Profit, Economic Profit, and Economic Loss

	Normal Profit	Economic Profit	Economic Loss
Revenue	\$500,000	\$550,000	\$480,000
Accounting cost	-450,000	-450,000	-450,000
Opportunity cost	-50,000	-50,000	-50,000
Profit	\$0	\$50,000	(\$20,000)
	<i>Note: Accounting profit of \$50,000 equals the opportunity cost of \$50,000.</i>	<i>Accounting profit of \$100,000 exceeds the opportunity cost of \$50,000.</i>	<i>Accounting profit of \$30,000 is less than the opportunity cost of \$50,000.</i>

In the case where revenue is less than economic cost, clearly a loss is incurred. However, this **economic loss** might well coincide with a firm earning an accounting profit. For example, suppose our entrepreneur's revenue is \$480,000. The economic loss would be \$20,000 ($\$480,000 - \$500,000$), but the accounting profit would be \$30,000 ($\$480,000 - \$450,000$). Table 8.1 summarizes the three scenarios discussed previously.

With these assumptions in mind, we are now ready to discuss the decision-making process. Suppose that, in determining whether to operate in a particular market at some level of output, the firm is faced with the short-run total cost structure presented in Table 8.2. (For convenience, the cost data are the same as those first presented in Table 7.2.)

Let us assume the market price is \$110. Given this price, the firm is free to produce as much as or as little as it desires. The demand, total revenue, marginal revenue, and average revenue schedules for this firm are shown in Table 8.3. Notice that, because the price to the firm remains unchanged regardless of its output level, the total, marginal, and average revenue schedules do not resemble the schedules analyzed in

Table 8.2 Total and Per-Unit Short-Run Cost

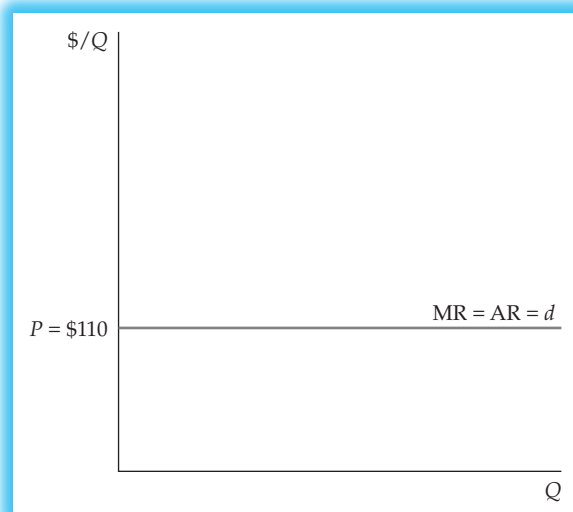
Quantity (Q)	Total Fixed Cost (TFC)	Total Variable Cost (TVC)	Total Cost (TC)	Average Fixed Cost (AFC)	Average Variable Cost (AVC)	Average Total Cost (AC)	Marginal Cost (MC)
0	\$100	\$ 0.00	\$ 100.00				
1	100	55.70	155.70	\$100.00	\$ 55.70	\$155.70	\$ 55.70
2	100	105.60	205.60	50.00	52.80	102.80	49.90
3	100	153.90	253.90	33.33	51.30	84.63	48.30
4	100	204.80	304.80	25.00	51.20	76.20	50.90
5	100	262.50	362.50	20.00	52.50	72.50	57.70
6	100	331.20	431.20	16.67	55.20	71.87	68.70
7	100	415.10	515.10	14.29	59.30	73.59	83.90
8	100	518.40	618.40	12.50	64.80	77.30	103.30
9	100	645.30	745.30	11.11	71.70	82.81	126.90
10	100	800.00	900.00	10.00	80.00	90.00	154.70
11	100	986.70	1,086.70	9.09	89.70	98.79	186.70
12	100	1,209.60	1,309.60	8.33	100.80	109.13	222.90

Table 8.3 Revenue Schedules

Quantity	Price (AR)	TR	MR
0	\$110	\$ 0	\$110
1	110	110	110
2	110	220	110
3	110	330	110
4	110	440	110
5	110	550	110
6	110	660	110
7	110	770	110
8	110	880	110
9	110	990	110
10	110	1,100	110
11	110	1,210	110
12	110	1,320	110

Chapter 4. As a price taker, the firm faces a demand curve that is “perfectly elastic.” That is, customers are willing to buy as much as the firm is willing to sell *at the going market price*. This special type of demand curve can be seen in Figure 8.2. Moreover, the firm receives the same marginal revenue from the sale of each additional unit of product. This marginal revenue is simply the price of the product. Recall that the price is tantamount to average or per-unit revenue. Hence, a perfectly competitive firm’s demand is also its marginal and its average revenue over the range of output being considered. Note in Figure 8.2 that the demand curve is also labeled “AR” and “MR.”

Figure 8.3 compares the perfectly elastic demand curve with the typical downward-sloping linear demand curve used in Chapter 4. It also shows the total revenue curves in relation to the two types of demand curves. As is the case with perfect elasticity, a downward-sloping demand curve is the same as the average revenue curve because P by definition is equal to AR. However, recall that a linear, downward-sloping demand curve is associated with a marginal revenue curve that is twice as steep. In addition, this

**Figure 8.2** Perfectly Elastic Demand Curve

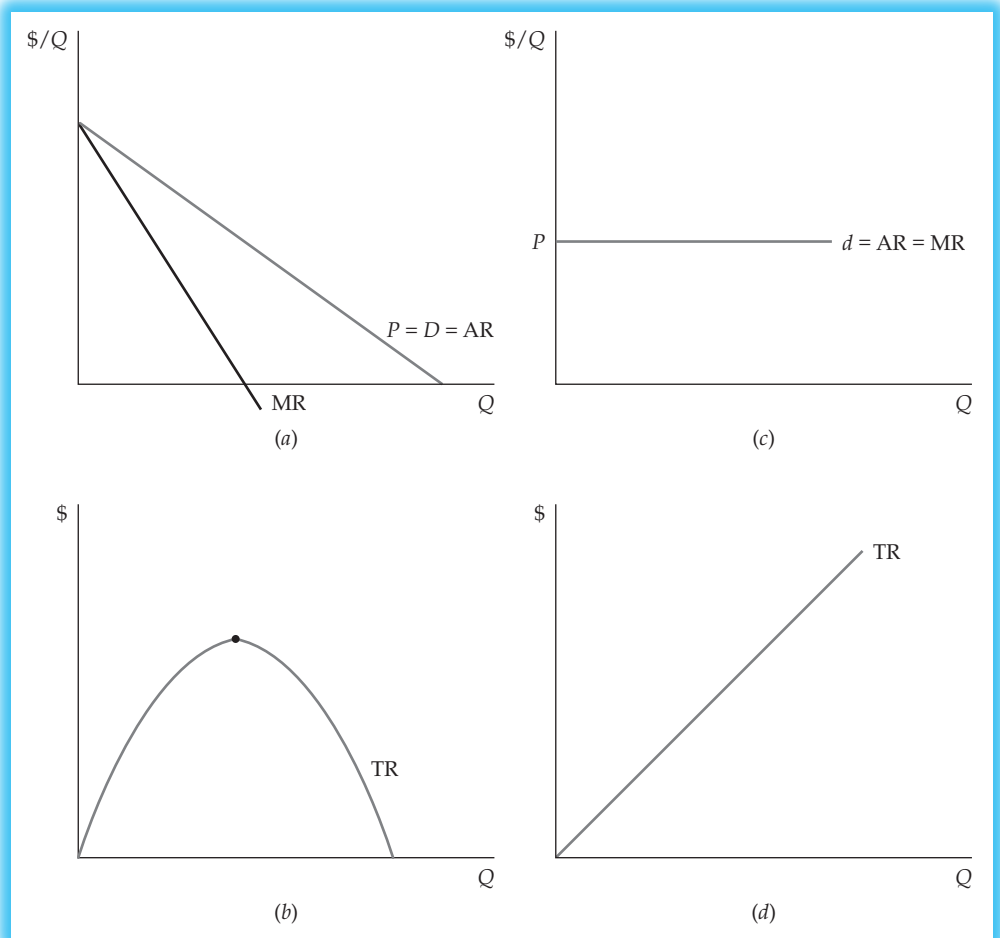


Figure 8.3 Different Types of Demand Curves and Associated Total Revenue Curves

type of demand results in a nonlinear total revenue curve that reaches a maximum at the point at which marginal revenue equals zero (see Figure 8.3*b*). In contrast, as shown in Figure 8.3*d*, there is no limit to the amount of total revenue that firms can garner in a perfectly competitive market. The more a firm produces, the more revenue it will obtain. The limit to its output is based on revenue in relation to the firm's cost of production, that is, the profit earned at various levels of output.

Armed with its cost and revenue schedules, all that a firm needs to do is to combine the sets of information to find the level of output that maximizes its profit (or minimizes its loss).



Module 8A

The Total Revenue–Total Cost Approach to Selecting the Optimal Output Level

The most logical approach to selecting the optimal level of output is to compare the total revenue with the total cost schedules and find that level of output that either maximizes the firm's profit or minimizes its loss. This is shown in Table 8.4 and Figure 8.4. As can be seen in the table and the figure, this output level is 8, at which the firm would be earning a maximum profit of \$261.60. Graphically, this output level can be seen as

Table 8.4 Cost and Revenue Schedules Used to Determine Optimal Level of Output

Quantity (Q)	Price (P)	Total Revenue (TR)	Total Fixed Cost (TFC)	Total Variable Cost (TVC)	Total Cost (TC)	Total Profit (π)
0	\$110	\$ 0	\$100	\$ 0.00	\$ 100.00	\$-100.00
1	110	110	100	55.70	155.70	-45.70
2	110	220	100	105.60	205.60	14.40
3	110	330	100	153.90	253.90	76.10
4	110	440	100	204.80	304.80	135.20
5	110	550	100	262.50	362.50	187.50
6	110	660	100	331.20	431.20	228.80
7	110	770	100	415.10	515.10	254.90
8	110	880	100	518.40	618.40	261.60
9	110	990	100	645.30	745.30	244.70
10	110	1,100	100	800.00	900.00	200.00
11	110	1,210	100	986.70	1,086.70	123.30
12	110	1,320	100	1,209.60	1,309.60	10.40

the one that maximizes the distance between the total revenue curve and the total cost curve. By convention, this point has been labeled Q^* .

The Marginal Revenue–Marginal Cost Approach to Finding the Optimal Output Level

Marginal analysis is at the heart of the economic analysis of the firm. Once we explain how marginal analysis is used by the firm to determine its optimal level of output, we rely primarily on this type of analysis throughout the rest of this chapter.

Figure 8.4 Determining Optimal Output from Cost and Revenue Curves—Perfect Competition

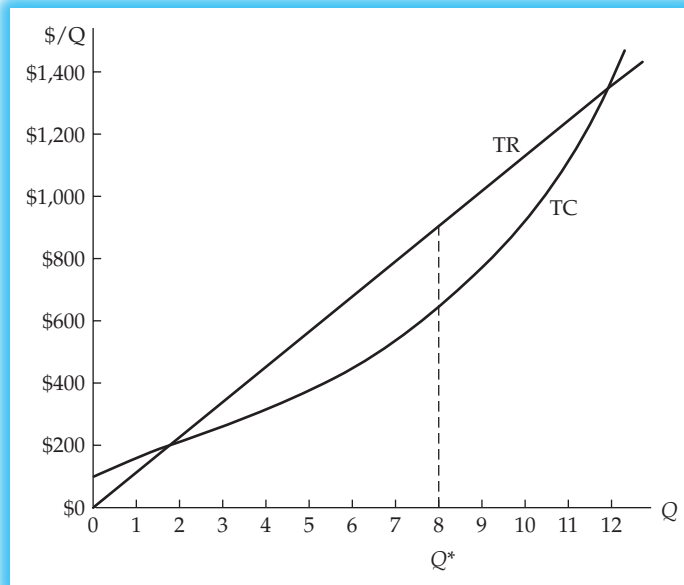


Table 8.5 Using Marginal Revenue (or Price) and Marginal Cost to Determine Optimal Output: The Case of Economic Profit

Quantity (Q)	Marginal Revenue (MR = P = AR)	Average Fixed Cost (AFC)	Average Variable Cost (AVC)	Average Total Cost (AC)	Marginal Cost (MC)	Marginal Profit (Mπ)
0	\$110				\$ 55.70	\$ 54.30
1	110	\$100.00	\$ 55.70	\$155.70	49.90	60.10
2	110	50.00	52.80	102.80	48.30	61.70
3	110	33.33	51.30	84.63	50.90	59.10
4	110	25.00	51.20	76.20	57.70	52.30
5	110	20.00	52.50	72.50	68.70	41.30
6	110	16.67	55.20	71.87	83.90	26.10
7	110	14.29	59.30	73.59	103.30	6.70
8	110	12.50	64.80	77.30	126.90	-16.90
9	110	11.11	71.70	82.81	154.70	-44.70
10	110	10.00	80.00	90.00	186.70	-76.70
11	110	9.09	89.70	98.79	222.90	-112.90
12	110	8.33	100.80	109.13		

Table 8.5 presents the cost and revenue data on a per-unit basis. The marginal revenue and marginal cost columns contain the key numbers the firm must use to decide on its optimal level of output. Let us examine the marginal revenue and the marginal cost associated with additional units of output, starting with zero units. As you can see in Table 8.5, the first unit would result in additional revenue of \$110 and cost the firm an additional \$55.70 to make. The second unit would add another \$110 to revenue and another \$49.90 to the firm's total cost. Continuing in this manner, we observe that it would be worthwhile for the firm to produce more as long as the added benefit of each unit produced and sold (i.e., the marginal revenue) exceeds the added cost (i.e., the marginal cost). Because the marginal revenue is equal to the existing market price, it does not change as output increases. However, because of the law of diminishing returns, the firm's marginal cost begins to increase with the fourth unit of output. From that point on, each additional unit of output costs *increasingly more* to produce. Between 0 and 8 units of output, we observe that marginal revenue exceeds marginal cost. However, production of the ninth unit of output will cost the firm more than the revenue that it would add (MC = \$126.90 and MR = \$110). In Table 8.5, MR = MC actually occurs between 8 and 9 units of output, but we use 8 as the approximate level of optimal output.

Using the relationship between marginal revenue and marginal cost to decide on the optimal level of output is referred to in economics as the **MR = MC rule**. The rule is stated as follows:

A firm that wants to maximize its profit (or minimize its loss) should produce a level of output at which the additional revenue received from the last unit is equal to the additional cost of producing that unit. In short, MR = MC.

The MR = MC rule applies to any firm that wants to maximize its profit, regardless of whether it has the power to set the price. However, in the particular case in which the firm has no power to set the price (i.e., it is a price taker), the MR = MC rule can be restated as the **P = MC rule**. This is simply because, when a firm is a price

taker, its marginal revenue is in fact the going market price. (Refer to Figure 8.3 for an illustration of this.)

Table 8.5 shows that by following the $MR = MC$ rule and producing 8 units of output, the firm would earn a profit of \$261.60 [$8(AR - AC)$], which is what we already learned by following the total revenue – total cost approach. Hence, the rule apparently works. Another way to think about MR and MC is in terms of marginal (i.e., additional) profit. If $TR - TC$ is equal to total profit, then $MR - MC$ must be equal to marginal profit ($M\pi = MR - MC$). The last column in Table 8.5 indicates the amount of additional profit that would be earned by the firm in producing additional units of output. As you can see, this column is merely the difference between the MR column and the MC column. When MR is equal to MC, marginal profit must be zero. When marginal profit is equal to zero, it indicates that the firm can make no more *additional* profit and, therefore, should not produce at a higher level of output.

Of course, there is nothing to prevent the firm from producing more or less than 8 units of output. As you can see in the last column of Table 8.4, it would still earn a profit if it produced at any of the output levels 2 to 12, but none of these amounts except 8 is the *maximum* that it could earn. Remember that we are now referring to *total revenue* and *total cost*. If the firm were to produce at the level where these two measures are equal, then clearly all it would be doing is earning a “normal” profit.

Although the optimal output level can be found just as easily by using the $TR - TC$ approach, economists rely much more on the $MR - MC$ approach in analyzing the firm’s output decision. Essentially, this approach is an extension of the basic analytical technique of “marginal analysis” first introduced in the chapters on demand, production, and cost. Furthermore, the practical implications of this approach are similar to those discussed in these earlier chapters. Often, firms do not have the benefit of complete columns of numbers depicting cost and revenue. Instead, they must rely on actual cost and revenue data at a particular level of output and then conduct sensitivity analysis involving relatively small incremental changes around that level. As is illustrated in the ensuing sections, “marginal analysis” is much better suited to this situation than “total analysis.”

The MR – MC Approach in Graphs

A graphical analysis using the $MR - MC$ approach employs the data in Table 8.5 and is shown in Figure 8.5. Also shown is the firm’s demand curve, a horizontal line intersecting the vertical axis at the level of the given market price of \$110. Thus, the demand curve of this price-taking firm is “perfectly elastic.” The optimal output level is clearly seen as the level at which the firm’s MC line and its MR line (demand line) intersect. The amount of profit earned is represented by the rectangle $ABCD$. Because these graphs are used in this way throughout the rest of the chapter, it is crucial that you clearly understand their interpretation.

Points on each of the unit cost curves indicate the dollar value of the cost at different levels of output. Therefore, at output level Q^* , the average cost is represented by the distance between point C and the horizontal axis (i.e., CQ^*). It follows that, because total cost is average cost multiplied by the quantity of output, it is shown as the area of the rectangle determined by OQ^* and CQ^* (rectangle $ODCQ^*$). In the same manner, we can show that total revenue can be displayed as the rectangle determined by OQ^* and BQ^* (rectangle $OABQ^*$). Therefore, profit (i.e., the rectangle $ABCD$) can be represented by the difference between the larger rectangle, depicting total revenue, and the smaller one, showing total cost.

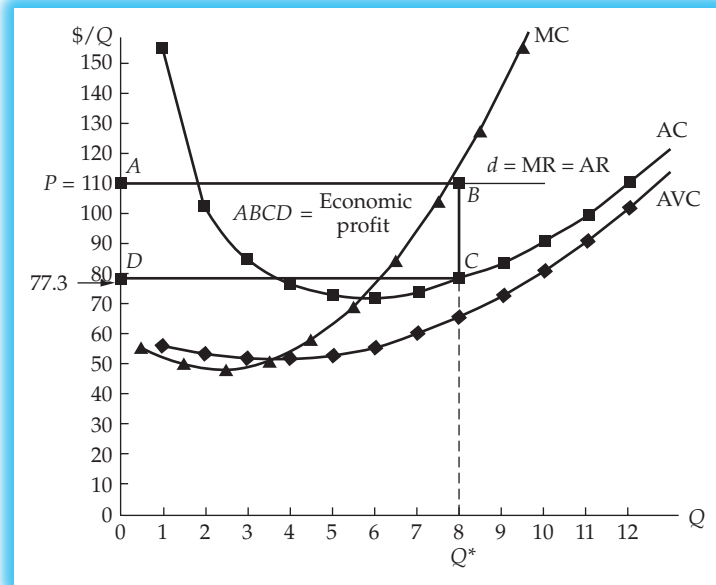


Figure 8.5 MR – MC Approach to Maximizing Economic Profit (using Table 8.5)

Economic Profit, Normal Profit, Loss, and Shutdown

The preceding example assumed the market price was high enough for the firm to earn an economic profit by following the $MR = MC$ rule. But because the firm is just one of many price-taking sellers in this market, there is no reason to expect that market price will always be this beneficial for the firm. Given the vagaries of supply and demand, it is just as likely that a firm will be faced with prices that result in only normal profit—or worse, in operating losses. Tables 8.6 and 8.7 and Figure 8.6 demonstrate these possibilities. To focus on the marginal revenue-marginal cost approach, we include only the per-unit cost data in the tables and figure.

The situation depicted in Table 8.7 and Figure 8.6*b* indicates a loss for the firm. Does this mean that the firm should not be in this market? As you know, in the short run, the firm must bear certain fixed costs regardless of the level of its output. Using the data in Table 8.7, if the firm were to shut down its operations (i.e., if $Q = 0$), it would still have a fixed cost of \$100. Given the market price of \$58, we know that the best that a firm can do is to follow the $MR = MC$ rule, produce 5 units of output, and lose \$72.50. But if the firm were to shut down, it would lose \$100 because this is the amount of fixed cost that it would incur, whether or not it operates in the short run. Therefore, with a market price of \$58, *it would be better for a firm to operate at a loss than to cease its activities in this market.* This is illustrated in Figure 8.7*a*.

Another way to understand this rationale is to compare the firm's total revenue at the \$58 price with its total variable cost, assuming a production level of 5 units. The total revenue is \$290 ($P \times Q$), and the total variable cost is \$262.50 ($Q \times AVC$). Clearly, this revenue is sufficient to cover the firm's total variable cost. Moreover, the amount left over (\$27.50) can be used to pay for a part of its fixed cost. Hence, we can also conclude that as long as a firm's total revenue is greater than its total variable cost (or on a per-unit basis, as long as the market price exceeds average variable cost), it is better to operate than to shut down because at least part of its fixed cost will be defrayed. We refer to the amount by which total revenue exceeds total variable cost as the **contribution margin** (see Figure 8.7*a*). You should also recognize that the portion of fixed cost that

Table 8.6 Using Marginal Revenue (or Price) and Marginal Cost to Determine Optimal Output: The Case of Normal Profit

Quantity (Q)	Marginal Revenue (MR = P = AR)	Average Fixed Cost (AFC)	Average Variable Cost (AVC)	Average Total Cost (AC)	Marginal Cost (MC)	Marginal Profit (M π)	Total Profit or Loss (Q[P - AC])
0							\$-100.00 ^a
1	\$71.87	\$100.00	\$ 55.70	\$155.70	\$ 55.70	\$ 16.17	-83.83
2	71.87	50.00	52.80	102.80	49.90	21.97	-61.86
3	71.87	33.33	51.30	84.63	48.30	23.57	-38.28
4	71.87	25.00	51.20	76.20	50.90	20.97	-17.32
5	71.87	20.00	52.50	72.50	57.70	14.17	-3.15
6	71.87	16.67	55.20	71.87	68.70	3.17	0
7	71.87	14.29	59.30	73.59	83.90	-12.03	-12.04
8	71.87	12.50	64.80	77.30	103.30	-31.43	-43.44
9	71.87	11.11	71.70	82.81	126.90	-55.03	-98.46
10	71.87	10.00	80.00	90.00	154.70	-82.83	-181.30
11	71.87	9.09	89.70	98.79	186.70	-114.83	-296.12
12	71.87	8.33	100.80	109.13	222.90	-151.03	-447.12

^aIf $Q = 0$, firm still incurs a total fixed cost of \$100 in the short run.

is *not* covered by the contribution margin is in fact the amount of the firm's loss (i.e., \$27.50 - \$100 = -\$72.50).

It is not always advisable to operate in the short run at a loss. Suppose the market price fell to \$50. In this case, even if the firm followed the $MR = MC$ rule, it would still incur a loss greater than it would have to bear by shutting down. This situation is not shown in a separate table but is discussed in relation to the AVC values in Table 8.7.

Table 8.7 Using Marginal Revenue (or Price) and Marginal Cost to Determine Optimal Output: The Case of Economic Loss

Quantity (Q)	Marginal Revenue (MR = P = AR)	Average Fixed Cost (AFC)	Average Variable Cost (AVC)	Average Total Cost (AC)	Marginal Cost (MC)	Marginal Profit (M π)	Total Profit or Loss (Q[P - AC])
0							\$-100.00
1	\$58	\$100.00	\$ 55.70	\$155.70	\$ 55.70	\$ 2.30	-97.70
2	58	50.00	52.80	102.80	49.90	8.10	-89.60
3	58	33.33	51.30	84.63	48.30	9.70	-79.89
4	58	25.00	51.20	76.20	50.90	7.10	-72.80
5	58	20.00	52.50	72.50	57.70	0.30	-72.50
6	58	16.67	55.20	71.87	68.70	-10.70	-83.22
7	58	14.29	59.30	73.59	83.90	-25.90	-109.13
8	58	12.50	64.80	77.30	103.30	-45.30	-154.44
9	58	11.11	71.70	82.81	126.90	-68.90	-223.29
10	58	10.00	80.00	90.00	154.70	-96.70	-320.00
11	58	9.09	89.70	98.79	186.70	-128.70	-448.69
12	58	8.33	100.80	109.13	222.90	-164.90	-613.56

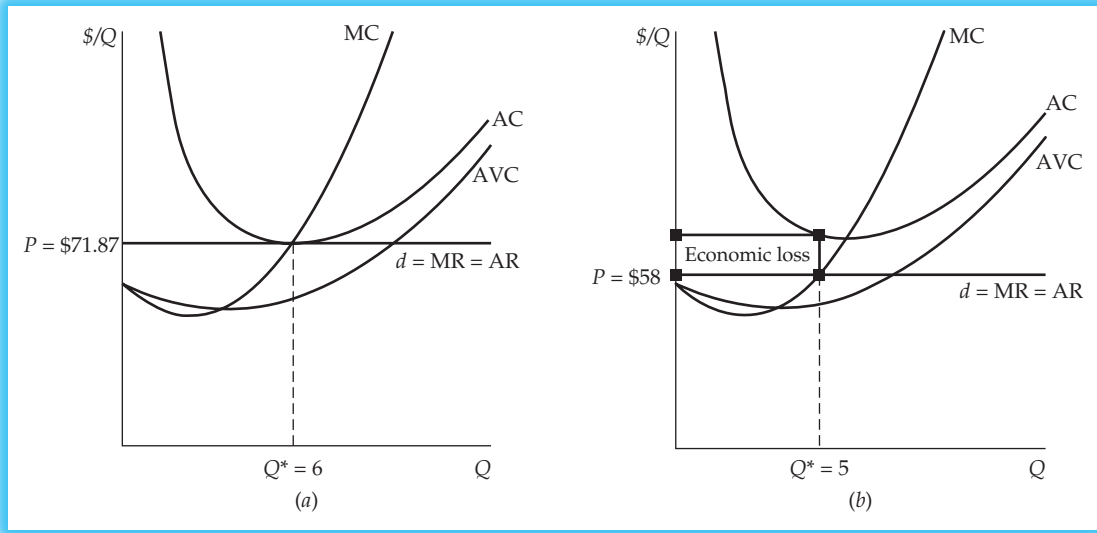
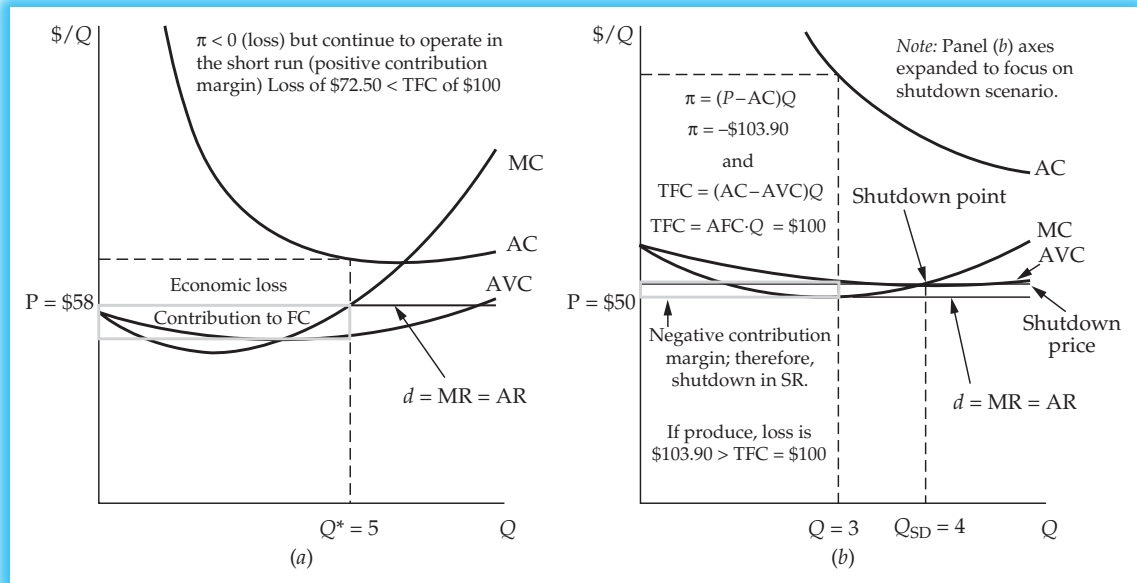


Figure 8.6 Normal Profit (a Using Table 8.6) and Economic Loss (b Using Table 8.7)

By literally following the $MR = MC$ rule, the firm would be led to produce 3 units of output. But we can see that at this level, the total revenue of \$150 ($\50×3) would not even be enough to cover the firm’s total variable cost of \$153.90 ($\51.30×3), resulting in a negative contribution margin of \$3.90. Looking at this situation in terms of the firm’s loss versus its fixed cost, we can see that its total loss of \$103.90 is clearly greater than the fixed cost of \$100 that it would incur if it decided to shut down its operations. (As you can see, the firm’s loss is the combination of its fixed cost and negative contribution margin.) Thus, given the market price of \$50, the firm would be better off by shutting down its operations. This is illustrated in Figure 8.7b.

Figure 8.7 Contribution Margin



Also shown in this figure is what economists refer to as the **shutdown point**. At this point, the market price is at a level in which a firm following the $MR = MC$ rule would lose an amount just equal to its fixed cost of production. Expressed in another way, this price would result in a zero contribution margin. At the shutdown point, we assume a firm would be indifferent about operating versus shutting down. However, it would certainly give strong consideration to ceasing to operate in the short run. As you can see, the shutdown point coincides with the point at which the firm's average variable cost is at its minimum. In conclusion:

In a competitive market, the firm's **short-run supply curve** is its MC curve subject to two caveats. The firm maximizes profits or minimizes losses by supplying at $P = MC$ along the positively sloped part of its MC curve if $P \geq$ minimum AVC. If $P <$ minimum AVC, then the firm should shut down (set $Q = 0$) in the short run.

The Competitive Market in the Long Run

Regardless of whether the market price in the short run results in economic profit, normal profit, or a loss for competing firms, economic theory states that in the **long run**, the market price will settle at the point where these firms earn a normal profit. This is because over a long period of time, prices that enable firms to earn above-normal profit would induce other firms to enter the market, and prices below the normal level would cause firms to leave the market. We just completed a discussion of the rationale for a firm's operating at a loss in the short run. However, in the long run, we assume firms that are losing money would have to seriously consider leaving the market even if they have positive contribution margins. Recall that in the long run, firms have the time to vary their fixed factors of production. This means that they would have sufficient time to liquidate the fixed assets that account for their fixed costs.

We discussed the long-run adjustment process of entering and exiting firms in Chapter 3. The entry of firms shifts the supply curve to the right, driving down market price. The exiting of firms shifts the supply curve to the left, placing upward pressure on market price. A firm's motivation to go into or get out of the market can now be examined in greater detail. There is only one price at which firms neither enter nor leave the market. This, of course, is the price that results in normal profits. The long-run process of entering and exiting firms is illustrated in Figure 8.8.

Figure 8.8*a* shows a hypothetical short-run situation in which the price (determined by supply and demand) is high enough to enable a typical firm competing in this market to earn economic profit. (Viewed in another way, given the market price, the firm's cost structure is low enough to enable it to earn economic profit.) Over time, new firms would enter the market, and the original firms would expand their fixed capacity in response to the incentive of economic profit. This would have the effect of increasing the market supply (shifting the supply curve to the right) and reducing the market price. At the point where firms earn only normal profit, this adjustment process would cease. Figure 8.8*b* shows the opposite case, in which a short-run loss incurred by firms in the market causes firms in the long run to leave the market. This causes price to rise toward the level in which the remaining firms would earn a normal profit.

The concept of a long-run "resting point" may seem a bit unrealistic. As much as we try to use real-world examples to support the theory of managerial economics, it is extremely difficult, if not impossible, to find examples of this principle in action. For one thing, in actual market situations, demand does not remain constant while supply adjusts toward the normal price. Tastes and preferences, as well as the number of buyers, incomes, and prices of related goods, are constantly changing. For another,

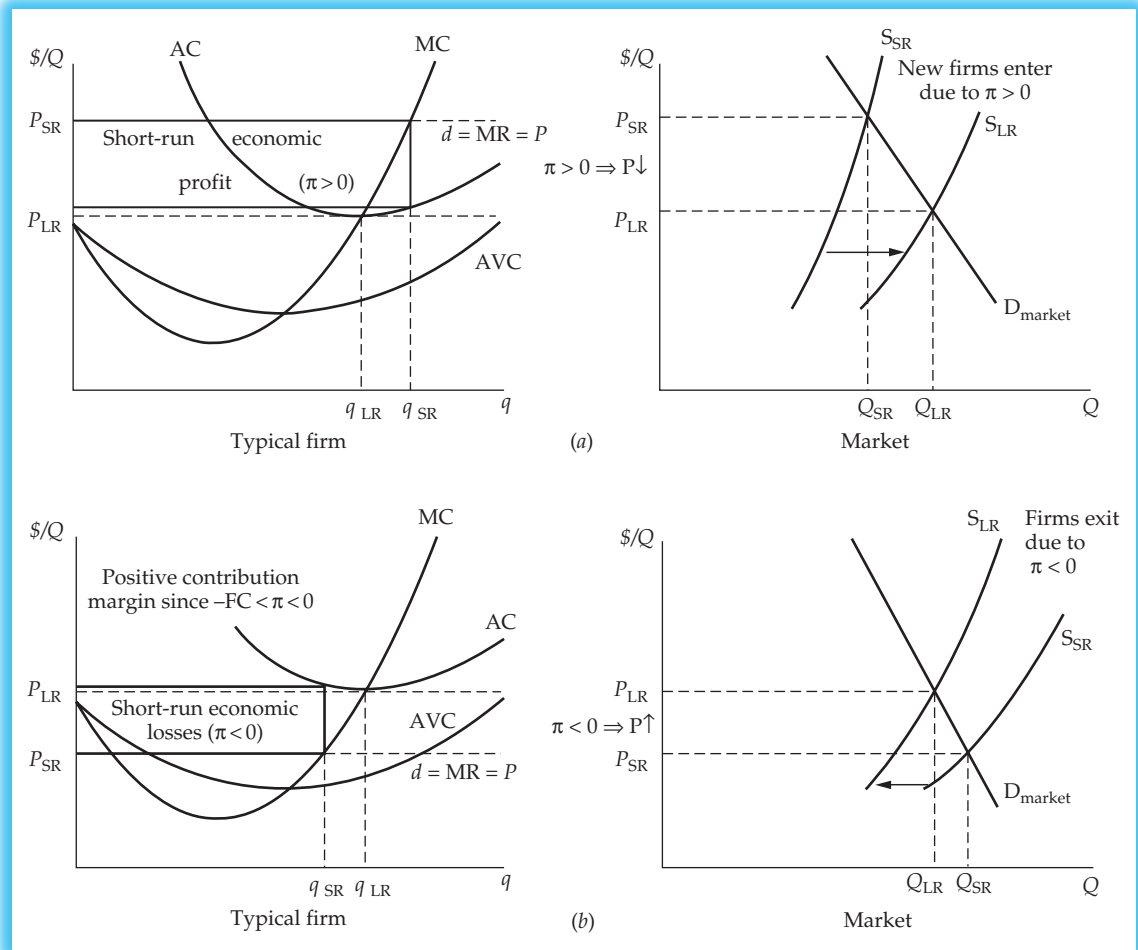


Figure 8.8 Long-Run Effect of Firms Entering (a) or Exiting (b) the Market

the economic notion of the long run is a theoretical construct, not a period that can be measured in calendar time. If the market price has not reached the normal level, economists can say that the market is still adjusting toward long-run equilibrium. But herein lies the principal relevance of this concept to the real world of business. For business decision makers, the process of *adjustment* toward equilibrium is far more important than the equilibrium price itself.

An understanding of the conditions motivating market entry or exit over the long run should lead the firms to consider the following points:

1. The earlier the firm enters a market, the better its chances of earning above-normal profit (assuming a strong demand in this market).
2. As new firms enter the market, firms that want to survive and perhaps thrive must find ways to produce at the lowest possible cost, or at least at cost levels below those of their competitors.
3. Firms that find themselves unable to compete on the basis of cost might want to try competing on the basis of product differentiation instead, although this is extremely difficult in this type of market.

PRICING AND OUTPUT DECISIONS IN MONOPOLY MARKETS

Module 8B
Module 8C

A monopoly market consists of one firm. The firm *is* the market. Examples are gas and electric utilities and firms selling products under protection of U.S. patent laws. Prior to its breakup in 1984, AT&T was considered one of the largest monopolies in the world. The regional companies that were formed after divestiture represented monopolies in many of the local calling areas within their respective regions for many years after 1984. Most monopolies cited previously are closely regulated by government or government-appointed agencies. (A notable exception is companies selling patented products.) Because this regulation severely constrains their ability to choose price and output levels, regulated monopolies are analyzed as a separate group of firms in Chapter 14.

In the absence of regulatory constraints, the monopoly stands in counterpoint to the perfectly competitive firm. Firms in perfectly competitive markets have no power to set their prices; the monopoly firm has the power to establish any price that it wants. If you were responsible for setting the price of a product that you alone were selling in the market, how much would you charge? The layperson's answer is usually "as much as I can" or "whatever the market will bear." On the surface, this answer seems reasonable enough. Unfortunately, it is too simplistic to be of much help to the monopolist. In 1948, when Polaroid first offered its camera, it could have charged any price that it wanted. The original price was \$85 (which was a considerable sum at that time), but it could just as well have been \$850 or \$8,500. The market could have borne those prices because some people probably would have been willing to buy the camera at higher levels. The question is *how many* people would have bought the cameras and *when*. As it turned out, Polaroid offered five cameras for sale on the first day and sold them all in several hours. Who knows how long it would have taken them to sell the 5 units at \$8,500?

The key point is that a monopoly firm's ability to set its price is limited by the demand curve for its product and, in particular, the price elasticity of demand for its product. (Recall that according to the law of demand, people will buy more as price falls, and vice versa.) The price elasticity of demand indicates how much more or less people are willing to buy in relation to price decreases or increases. If we assume the firm's downward-sloping demand curve is linear, we know that as the price of the product falls, the marginal revenue from the sale of additional units falls, reaches zero, and then becomes negative. For purposes of illustration, let us also assume the firm's marginal cost is constant in the short run. The linear, downward-sloping demand curve, the marginal revenue curve, and the constant marginal cost curve for such a firm are shown in Figure 8.9. Notice that if the firm charges too high a price (e.g., P_1), its marginal revenue will exceed its marginal cost; hence, it will be forgoing some amount of marginal profit. If the firm sets its price at too low a level (e.g., P_2) its marginal cost will exceed its marginal revenue, and the firm will experience a marginal loss.

The ability of a monopoly to set its price is further limited by the possibility of rising marginal costs of production. If this is the case, then surely at some point the increasing cost of producing additional units of output will exceed the decreasing marginal revenue received from the sale of additional units. This begins at Q^* shown in Figure 8.10.

In conclusion, the firm that exercises a monopoly power over its price should not set its price at the highest possible level. Instead, it should set it at the *right* level. And what is this "right" level? It is the level that results in $MR = MC$.

To see how the $MR = MC$ rule applies to the monopolist, see Table 8.8. Note that the table presents only the cost data relevant to this example. For purposes of comparison, the same cost figures used in the previous section for the perfectly competitive

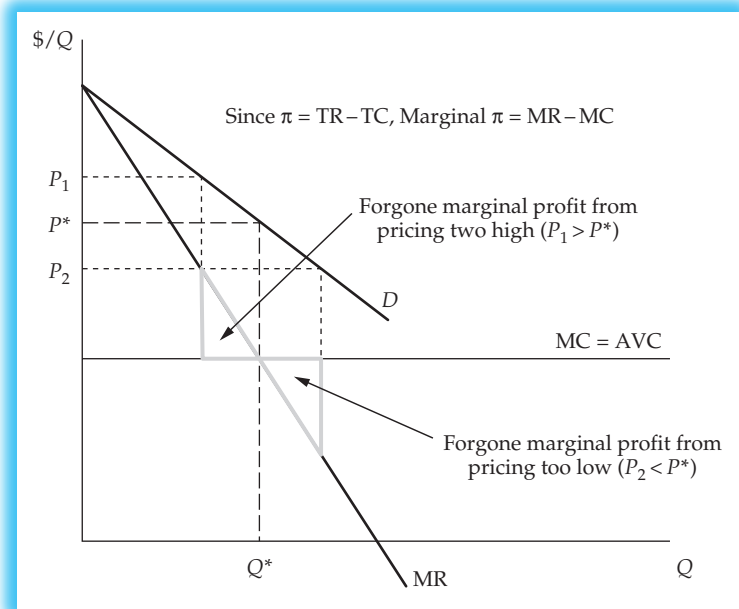
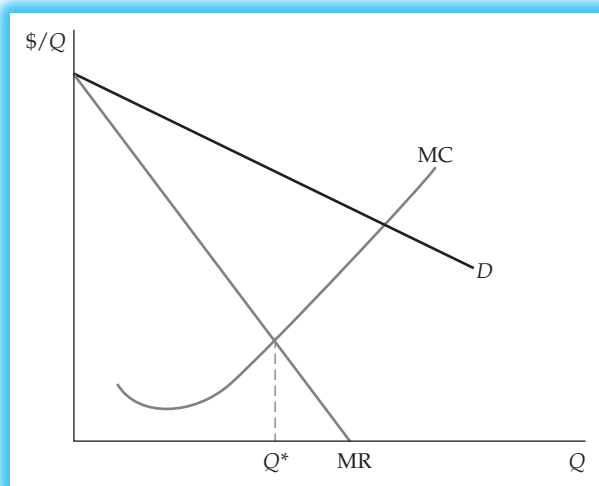


Figure 8.9 Marginal Loss in Profit from Failing to Produce at $MR = MC$

firm have been selected.³ But in this case, we assume the firm is the “only game in town.” Note that the price is not equal to the marginal revenue because the firm is a price setter and not a price taker. Its demand schedule consists of columns 1 and 2, and the total revenue and marginal revenue schedules are those that normally accompany a downward-sloping demand curve.

Figure 8.10 Increasing Marginal Costs in Relation to Decreasing Marginal Revenue



³By maintaining consistency in the cost data, we realize that we are sacrificing some realism because a monopoly would obviously produce more than a perfectly competitive firm. However, this shortcoming can be rectified by simply assuming each unit of output is the equivalent of a larger number of units (e.g., $Q = 1 = 1,000$).

Table 8.8 Using Marginal Revenue and Marginal Cost to Determine Optimal Price and Output: The Case of Monopoly

Quantity (Q)	Price (P)	Total Revenue (TR)	Marginal Revenue (MR)	Average Total Cost (AC)	Total Cost (TC)	Marginal Cost (MC)	Total Profit (π)
0	\$180	\$ 0			\$ 100.00		\$-100.00
1	170	170	\$170	\$155.70	155.70	\$ 55.70	14.30
2	160	320	150	102.80	205.60	49.90	114.40
3	150	450	130	84.63	253.90	48.30	196.10
4	140	560	110	76.20	304.80	50.90	255.20
5	130	650	90	72.50	362.50	57.70	287.50
6	120	720	70	71.87	431.20	68.70	288.80
7	110	770	50	73.59	515.10	83.90	254.90
8	100	800	30	77.30	618.40	103.30	181.60
9	90	810	10	82.81	745.30	126.90	64.70
10	80	800	-10	90.00	900.00	154.70	-100.00
11	70	770	-30	98.79	1,086.70	186.70	-316.70
12	60	720	-50	109.13	1,309.60	222.90	-589.60

Starting from the zero output level, let us consider the price, output, marginal revenue, marginal cost, and marginal profit as additional units of output are produced. You can see that as output increases, the marginal revenue associated with each unit exceeds the marginal cost up to 6 units. Beyond this level, the firm actually incurs a marginal loss. As the firm moves beyond this level, total profit is still positive, but it is not at its maximum. In other words, by following the $MR = MC$ rule, a profit-maximizing firm would want to produce 6 units of output per time period. To do so, it would have to set a price of \$120.

The way in which the $MR = MC$ rule underlies the monopoly price can perhaps be seen more clearly in a graph. In Figure 8.11, we see that the firm would select P^* because, given the particular demand for the product, this is the price that would prompt customers to buy Q^* . Q^* is the quantity that the firm would want to produce per time period because this is the amount at which the revenue received from the last unit produced is just equal to its cost (i.e., marginal revenue = marginal cost). With the same graphical references to total revenue and total cost as used in the analysis of perfect competition, we arrive at the measure of total profit as the area of the $(P - AC) \times Q$ rectangle, $ABCD$.

In a perfectly competitive market, the short-run economic profit enjoyed by the monopoly firm in this example would be vulnerable in the long run to the entry of other firms wanting to earn similar amounts of profit. But because we assume it is a monopoly, this firm would not be subject to such threats in the long run. Nonetheless, the preceding illustration is not meant to give the impression that a monopoly automatically earns economic profit in either the short or the long run. Whether or not it does depends on the demand for its product. For example, a company may have a monopoly on a toy for children that is in great demand and consequently enables it to earn the kind of economic profit illustrated in Figure 8.11. But as the market demand is filled or as children begin to tire of the product, the demand could decline (e.g., the demand curve shifts to the left) to the extent that the firm earns only a normal profit or perhaps even incurs a loss.

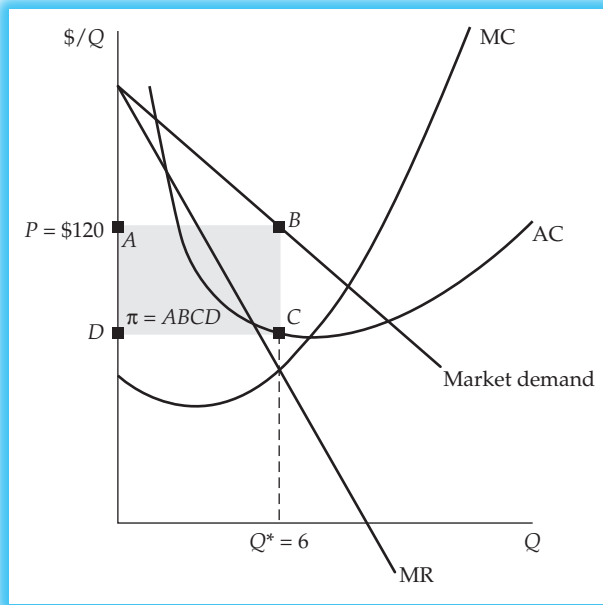


Figure 8.11 Graphical Depiction of $MR = MC$ Rule for a Monopoly Using Table 8.8

Suppose a price-setting firm does not want to maximize its short-run profit but instead wants to maximize its revenue. Let us explore this possibility using the data in Table 8.8. The price that maximizes the total revenue can be determined simply by observation. As you can see, by charging \$90 for its product, the firm will receive the maximum total revenue of \$810. You can also see that this revenue-maximizing price is lower than the one that maximizes the firm's total profit (i.e., $\$90 < \120). This relationship is illustrated in Figure 8.12. This dichotomy between **pricing for profit** and **pricing for revenue** is discussed in greater detail in Chapter 10.

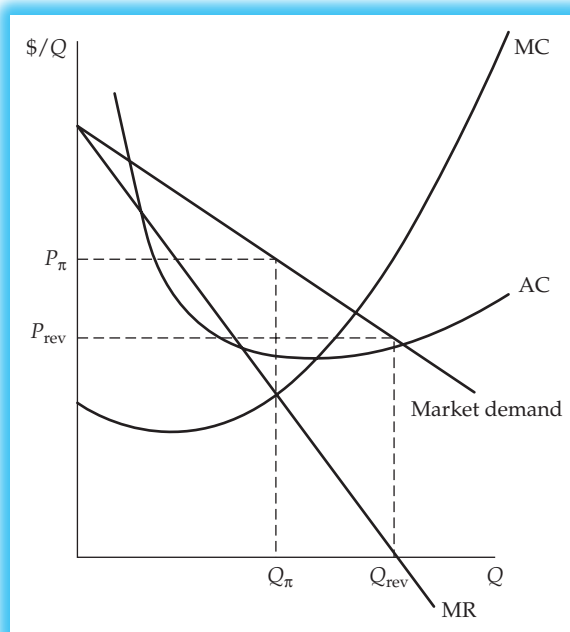


Figure 8.12 Relationship Between Profit-Maximizing and Revenue-Maximizing Price and Quantity

THE IMPLICATIONS OF PERFECT COMPETITION AND MONOPOLY FOR MANAGERIAL DECISION MAKING

After studying this chapter, it might appear to the reader that managers in a perfectly competitive or a monopoly market do not have much of a challenge in deciding on the price and output levels of their firm. In the case of perfect competition, the market price is determined for the managers by the forces of supply and demand. All that they have to do is to decide whether their cost structure will enable their firm to at least earn a normal amount of profit. In the case of monopoly, the fact that the firm has no competition enables its managers simply to follow the $MR = MC$ rule to maximize its profit. We would agree with this view, particularly when the pricing and output challenges of these firms are compared with those in monopolistic competition and oligopoly. This will be evident when reading the next chapter. Nonetheless, as explained earlier, we have presented a detailed analysis of perfect competition and monopoly because those models serve as the basis from which pricing and output decisions in monopolistic competition and oligopoly can be better understood and appreciated. Besides this, we believe that the analysis of firms in perfect competition and monopoly offers lessons to managers that go beyond the routine application of the $P = MC$ or $MR = MC$ rule.

The most important lesson that managers can learn by studying the perfectly competitive market is that it is extremely difficult to make money in a highly competitive market. Indeed, the only way for firms to survive in perfect competition is to be as cost efficient as possible because there is absolutely no way to control the price. Another lesson offered by the perfectly competitive model is that it might pay for a firm to move into a market before others start to enter. This might mean entering a market even before the demand is high enough to support an above-normal price. Spotting these market opportunities and taking the risk of going into these markets are key tasks of a good manager. Of course, the demand may never materialize or the long-run increase in supply might be so great that no one makes any money in this market. But that is all part of the risk that a manager must sometimes take. We see more of the making of pricing and output decisions in highly competitive markets in Chapter 9 when we examine the case of monopolistic competition.

In monopoly markets not sanctioned by the government via regulations or patent laws, a monopoly presents a manager with somewhat of a paradox. What happens if the managers of a firm are so successful in beating the competition that the firm in fact becomes a monopoly, or at least one that exercises monopolistic power? An excellent historical example of this was the case of the Department of Justice antitrust suit against IBM in 1969, when IBM dominated the computer hardware business. The suit was dropped in 1982. Some years later, the Department of Justice turned its attention to possible antitrust violations by the software giant Microsoft. Microsoft settled with the Department of Justice in 2001, but the case continued on with separate lawsuits by nine states. In early 2002, AOL Time Warner instigated a new suit against Microsoft on behalf of its subsidiary, Netscape. The suit was settled in May of the following year. As part of the settlement, the two companies agreed to a new royalty-free, seven-year license of Microsoft's browsing technology. In addition, Microsoft agreed to pay \$750 million to AOL Time Warner.⁴ The assault against Microsoft has continued to the present time. The Justice Department maintains a

⁴<http://www.microsoft.com/presspass/press/2003/may03/05-29msaolsettlementpr.msp>.

website with listings of Department of Justice Antitrust case filings against Microsoft with links to settlement information as well as Supreme Court, Appeals Court, and District Court.⁵

Microsoft was also a subject of antitrust actions by the European Union. In 2004, European Competition Commissioner Mario Monti ruled that Microsoft had failed to provide to rivals information that they needed to compete fairly in the market for server software, and that the company had been offering Windows on the condition that it come bundled with Windows Media Player, stifling competition. Microsoft was fined €497 million. At first, Microsoft appealed the ruling in 2007, but later decided to drop the appeal.⁶

Our personal view on this matter is that changing market conditions can often be more effective than government regulation in dampening the price-setting power of a monopoly. Shortly after the Justice Department dropped its case against IBM, the company began to lose a considerable amount of business as personal computers, workstations, and client server networks become more predominant than mainframes in many aspects of computing. Polaroid, the company that had a near-monopoly on its instant developing camera, was hurt by one-hour photo developing and was eventually overcome by digital photography. The Microsoft operating system has dominated the PC market since its inception in the mid-1980s, but it is not the dominant player in the fast-growing smartphone and tablet PC markets. Even Google's dominance in Internet search is being challenged by Microsoft's Bing and by Facebook.

Perhaps one of the best examples of the vulnerability of a monopoly can be found in the pharmaceutical industry. Until the early 1990s, firms in this industry enjoyed among the highest profit margins and returns on equity of all the companies in the *Fortune* 500. However, a number of recent events have started to erode their profitability. To start with, companies that do not have the patent on a drug that treats a particular illness are coming up with what the industry calls "me-too" products. These are drugs that offer a therapy for the same illness but whose chemical compositions are different enough to come under different patents. Furthermore, patents are beginning to run out on a number of highly profitable drugs and generic drugs are entering the market at far lower prices. Finally, those who make the purchasing decisions are exerting much more market power than ever before. It used to be that the majority of purchasers were fee-based private physicians who wrote prescriptions regardless of price, partly because they knew that their patients would be reimbursed by third-party payers such as insurance companies. Now much more of the purchasing decisions are being made by cost-conscious health maintenance organizations, hospital associations, and networks of retail pharmacies. At the same time, copay rates have increased, making end consumers (i.e., the patients) more aware of the price of the drugs they are being prescribed.

The key lesson for managers to learn from the many examples of once-powerful monopolies or near-monopolies that have eventually been affected by changing economics is not to be complacent or arrogant and assume their ability to earn economic profit can never be diminished. This is certainly the case for monopolistic competition and oligopoly, as seen in Chapter 9.

⁵http://www.justice.gov/atr/cases/ms_index.htm

⁶For example, see: http://www.news.com/EU-slaps-record-fine-on-Microsoft/2100-1001_3-5178281.html and <http://www.msnbc.msn.com/id/20818452>.

GLOBAL APPLICATION: THE \$736,000 BLUEFIN TUNA

In January 2012, a 593-pound bluefin tuna at the Tsukiji Fish Market in Tokyo sold for a price that comes to over \$1,200 a pound. Of course, this made global headline news.⁷ It happened to be some sort of publicity stunt by the owner of a chain of sushi restaurants in Japan. Nonetheless, it does not negate the fact that supply and demand conditions for this coveted species continues to put upward pressures on its “normal” market price. In the previous edition of this text, we talked about the factors behind the supply and demand for bluefin tuna. It starts—and ends—with the worldwide demand for sushi.

The global sushi phenomenon has been written about extensively in the popular press. The simple keyword “sushi” on any search engine will produce thousands of such sources of information. One freelance writer of travel magazine articles, Sasha Issenberg, became so interested in sushi that she wrote an entire book on this subject.⁸ In her book, Issenberg describes the challenges of maintaining adequate fish stocks because of rising global demand and overfishing by international fleets. Governments of the world have established quotas, but these quotas are difficult to enforce. In an interview with Devin T. Steward, director of the Global Policy Innovations program at the Carnegie Council for Ethics in International Affairs, Issenberg explained that the new members of the Chinese middle class were starting to demand more sushi, in a fashion similar to their economic counterparts in places like Russia and Dubai. Moreover, she pointed out that the reason why China has had such a noticeable impact on the market for seafood served in sushi restaurants is because so many Chinese have experienced rising incomes so quickly.⁹

Suppose you are an owner-manager of a sushi restaurant. Would current market conditions be good or bad for your business? What key managerial decisions would you have to make under these circumstances? Consider the plight of those who started this whole phenomenon: the owners and operators of sushi restaurants in Japan. Outside Japan, we may consider “sushi” to be an assortment of all types of fanciful dishes such as the “California Roll”¹⁰ (whose name derives from the fact that avocado is part of its mix) and “Rock and Roll” (considered by one of the authors as simply a mix of whatever fish was not sold the day before!). However, in Japan, a restaurant cannot be considered a true sushi restaurant unless bluefin tuna (*maguro* in Japanese, generally called simply “tuna” in the United States) is readily available to customers. One part of the bluefin tuna, the belly (*toro* in Japanese), is considered to be the premium cut, demanded by discriminating Japanese customers and commanding the highest price.¹¹ So here is the critical challenge for sushi restaurant owners in Japan: The market price for bluefin tuna is skyrocketing for all the reasons cited above. And the problem is that wholesale prices are now so high that Japanese restaurant owners have a difficult time passing on the full cost to their customers. Nonetheless, they continue to offer bluefin tuna on their menus because, as Izumi Niitsu, manager of the Kihachi

⁷See, for example, the article at the Huffington Post: http://www.huffingtonpost.com/2012/01/05/bluefin-tuna-record_n_1185770.html. This is not a new phenomenon. In 2009, several owners of sushi bars in Tokyo paid over \$100,000 for a Bluefin tuna in the Tokyo fish auction. This was several times more than the average price in 1999. (http://www.msnbc.msn.com/id/28503702/ns/world_news-weird_news/t/premium-tuna-fetches-auction/#.UDdk0Rw3U-0)

⁸See Sasha Issenberg, *The Sushi Economy: Globalization and the Making of a Modern Delicacy*, New York: Gotham, 2007, pp. 323.

⁹From Sasha Issenberg, “How the World Grew a Taste for Sushi,” Asia Times Online (<http://www.atimes.com/atimes/Japan/IH22Dh01.html>), August 22, 2007.

¹⁰Issenberg’s “Chengdu roll” in the above quote was made in satirical reference to our California roll.

¹¹In some Japanese sushi restaurants in the United States, we have noticed that there are sometimes two or three grades of *toro*, with the top grade sometimes selling for up to \$12 for a sliver about the surface size of a matchbox. This is probably the case in Japan as well.

restaurant in Tokyo, says: “If you have good tuna, you have a reputation of being a proper restaurant.”¹²

In order to offset the low or even negative profit margin of bluefin tuna, sushi restaurant owners in Japan must try to make more profit from other sushi items on the menu. Besides other kinds of raw and prepared seafood, there is now talk among restaurant owners in Japan about offering such foods as raw horse and deer meat. These items were offered in the early 1970s when few people wanted to eat tuna in Japan because of a mercury scare.¹³ But some traditional sushi restaurant owners do not place too much hope in such exotic alternatives. Tadashi Yamagata, vice chairman of Japan’s national union of sushi chefs, declared, “It’s like America running out of steak. Sushi without tuna just would not be sushi.”

Yet Mr. Yamagata is also a realist and knows that he cannot just stand by while tuna prices continue to rise and the long-run prospects for survival of this valued fish are uncertain. At his family restaurant, Miyakozushi, which has been in business for four generations, Mr. Yamagata, has been experimenting with more creative tuna alternatives. According to the just cited article in the *New York Times* (see footnote 13),

His [Yamagata’s] most successful substitutes were ideas he “reverse imported” from the United States, like smoked duck with mayonnaise and crushed daikon with sea urchin. He said he now made annual visits to sushi restaurants in New York and Washington for inspiration.

As Yamagata said, “We can learn from American sushi chefs. Sushi has to evolve to keep up with the times.”¹⁴

The Solution

Armed with all the available figures on the estimated cost and demand for “Go-nuts,” Frank spent the next week trying to come up with an optimal price for the product. The weekly demand for the firm’s energy drink was estimated to be

$$Q_D = 2,000 - 1,000P \quad (8.1)$$

where Q_D = Quantity of 12-ounce plastic bottles (in thousands)
 P = Price per container

Based on estimates provided by the bottling plant, Frank expressed the cost function as

$$TC = 150 + 0.25Q \quad (8.2)$$

where TC = Total cost per week (in thousands of dollars)
 Q = Output of 12-ounce plastic bottles (in thousands)

To find the optimal price on the basis of the $MR = MC$ rule, Frank first found the total revenue and marginal revenue functions based on the data in Equation (8.1). Expressing this equation in terms of price,

$$P = 2 - 0.001Q \quad (8.3)$$

and substituting this into the equation for total revenue (i.e., $TR = P \times Q$), he found total revenue to be

$$TR = 2Q - 0.001Q^2 \quad (8.4)$$

(continued)

¹²Blaine Harden, “Japan’s Sacred Bluefin, Loved Too Much,” *Washington Post Foreign Service*, November 11, 2007.

¹³Martine Fackler, “Waiter, There’s Deer in My Sushi,” *New York Times*, June 25, 2007.

¹⁴Ibid.

(continued)

To find marginal revenue, he took the first derivative of this equation and set it equal to the firm's marginal cost. [Based on Equation (8.2), he knew that the firm incurred a constant marginal cost of \$0.25 per unit of the product.] He then solved for the quantity (Q^*) that satisfied the equality. Then he found the optimal price by substituting the value of this optimal quantity into Equation (8.3).

$$\begin{aligned} MR &= \frac{dTR}{dQ} = 2 - 0.002Q \\ 2 - 0.002Q &= 0.25 \\ Q^* &= 875 \text{ (875,000 units per week)} \\ P &= 2 - 0.001(875) \\ P^* &= 1.125 \\ &= \$1.10 \text{ (rounding to the nearest 10 cents)} \end{aligned} \tag{8.5}$$

Figure 8.13 shows Frank's solution on a per-unit basis.

Frank assumed the main distributors of "Go-nuts" would be small retail food establishments, which generally mark up the wholesale price of energy drinks by about 100 percent. Therefore, Frank determined the wholesale price that Global Foods could charge the retail stores by simply taking 50 percent of the optimal price of \$1.10. For every unit sold, the company would receive \$0.55.

Frank then estimated the company's weekly profit from the production and sale of energy drinks.

$$\begin{aligned} TR &= \$0.55 \times 875 \\ &= 481.25, \text{ or } \$481,250 \\ TC &= 150 + 0.25(875) \\ &= 368.75, \text{ or } \$368,750 \\ \text{Total profit} &= \$481,250 - \$368,750 \\ &= \$112,500 \end{aligned}$$

Although Frank realized that he had the optimal price for "Go-nuts," he also knew that there were already companies with established energy drinks in the market that were selling for about \$1.25 a bottle. As the new entrant into the market, should Global Foods go along with the price established by the market leaders or should it try to sell it at a lower price? He was going to have to take up this issue with Nicole Goodman.

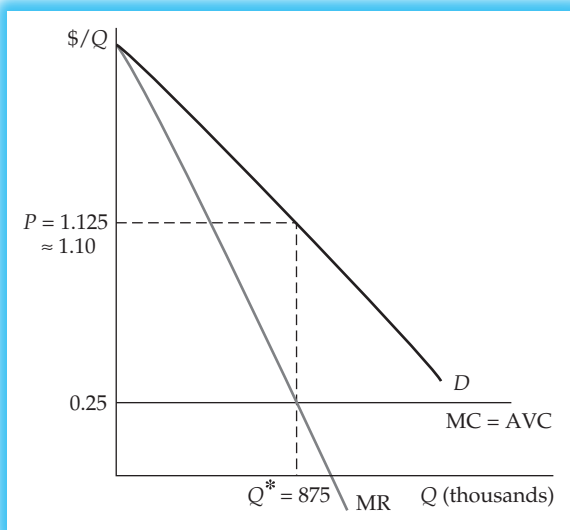


Figure 8.13 Frank's Result for the Optimal Pricing Problem

SUMMARY

This chapter presents a view of the pricing and output decisions facing firms in two extreme situations. In the case of perfect competition, the firm has virtually no power to set the price and is only able to decide to what extent (if at all) it wants to produce in this market, given the going market price. In the case of a monopoly, the firm is the entire market supply. This monopoly of supply gives the firm the power to set any price that it desires. In certain cases, this monopoly power is regulated by the government.

We demonstrate that firms wanting to maximize their short-run profit (or minimize their short-run loss) should establish their price and output levels according to the $MR = MC$ rule. For those firms in perfectly competitive markets, MR is in fact equal to the price that has already been established for them by the forces of supply and demand. For these price-taking firms, the only task is to decide what output quantity results in the matching of the market price (i.e., marginal revenue) and the marginal cost of producing the last unit of its output. For the monopoly firm, following the $MR = MC$ rule involves pricing the product at the level whereby the quantity that people purchase is the amount needed to bring MR in line with MC . We now turn to the cases between the two extremes of perfect competition and monopoly. For these “imperfect competitors,” the $MR = MC$ rule is an important part of their pricing decision. However, as we show, the actions or reactions of their competitors also play a major role in the pricing of their products.

IMPORTANT CONCEPTS

Contribution margin. The amount of revenue that a firm earns above its total variable cost. According to economic analysis, a firm experiencing a loss may continue operating in the short run if it has a positive contribution margin. A firm experiencing a negative contribution margin must shut down its operations because its revenue cannot even cover its variable costs of operations. (p. 324)

Economic cost. All cost incurred to attract resources into a company’s employ. Such cost includes explicit cost usually recognized on accounting records and opportunity cost. (p. 317)

Economic loss. A situation that exists when a firm’s revenues cannot cover its accounting cost and its opportunity cost of production. (p. 318)

Economic profit. Total revenue minus total economic cost. An amount of profit earned in a particular endeavor above the amount of profit that the firm could be earning in its next-best alternative activity. Also referred to as *abnormal profit* or *above-normal profit*. (p. 317)

Long run (market analysis). Firms are expected to enter a market in which sellers are earning economic profit. They are expected to leave a market in which sellers are incurring economic losses. (p. 327)

Market power. The power to establish the market price. (p. 312)

Market structure. The number and relative sizes of the buyers and sellers in a particular market. A

“competitive” market structure implies that the number of buyers and sellers in a market is large enough that it is difficult, if not impossible, for any one buyer or seller to determine the market price. (p. 312)

Monopoly. A market in which there is only one seller for a particular good or service. There may be legal barriers to entry into this type of market (e.g., regulated utilities, patent protection). (p. 315)

$MR = MC$ rule. A rule stating that if a firm desires to maximize its economic profit, it must produce an amount of output whereby the marginal revenue received at this particular level is equal to its marginal cost. This implies that those firms with market power must set a price that prompts buyers to purchase this particular level of output. (p. 322)

Normal profit. An amount of profit earned in a particular endeavor that is just equal to the profit that could be earned in a firm’s next-best alternative activity. When a firm earns normal profit, its revenue is just enough to cover both its accounting cost and its opportunity cost. It can also be considered as the return to capital and management necessary to keep resources engaged in a particular activity. (p. 317)

$P = MC$ rule. A variation of the $MR = MC$ rule for those firms operating in perfectly competitive markets. In such markets, firms are price takers. Thus, the price they must deal with (which has been determined by the forces of supply and

demand) is in fact the same as a firm's marginal revenue. When $P = MC$ at multiple outputs, the firm should choose the one on the positively sloped part of the marginal cost curve (the output on the negatively sloped part is a profit minimum). Firms using this rule must also be careful that the price is greater than average variable cost as well as equal to marginal cost (i.e., $AVC < P = MC$). If a firm cannot operate at the production level where this condition holds, it should shut down its operations. (p. 322)

Perfect competition. A market with four main characteristics: (1) a large number of relatively small buyers and sellers, (2) a standardized product, (3) easy entry and exit, and (4) complete information by all market participants about the market price. Firms in this type of market have absolutely no control over the price and must compete on the basis of the market price established by the forces of supply and demand. (p. 312)

Price makers. Firms that exercise market power through product differentiation or by being dominant players in their markets. (p. 312)

Price takers. Firms that operate in perfectly competitive markets. (p. 312)

Pricing for profit. The method of pricing that follows the $MR = MC$ rule. (p. 332)

Pricing for revenue. The pricing of a product to maximize a firm's revenue. In this case, the firm would try to price its product to sell an amount of output whereby the revenue earned from the last unit sold would be equal to zero (i.e., $MR = 0$). Assuming the firm faces a linear demand curve, the price it establishes to maximize revenue would be lower than the price that would maximize its profit. (p. 332)

Short-run supply curve. The supply curve for a firm in a competitive market is the positively sloped part of its marginal cost curve above minimum AVC. For prices below this level, the firm minimizes losses by shutting down in the short run. (p. 327)

Shutdown point. The point at which the firm must consider ceasing its production activity because the short-run loss suffered by operating would be equal to the short-run loss suffered by not operating (i.e., the operating loss = total fixed cost). In a perfectly competitive situation, this point is found at the lowest point of a firm's average variable cost curve. If the market price falls to this point, the firm should consider shutting down its operations. Any price lower than this would dictate that the firm should cease its operations in the short run. (p. 327)

QUESTIONS

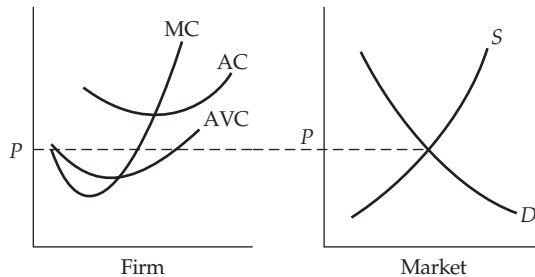
1. What are the main characteristics of a perfectly competitive market that cause buyers and sellers to be price takers? Explain.
2. Explain the importance of free entry and exit in the perfectly competitive market. That is, if free entry and exit did not exist, what impact would this have on the allocation of resources and on the ability of firms to earn above-normal profits over time?
3. "The perfectly competitive model is not very useful for managers because very few markets in the U.S. economy are perfectly competitive." Do you agree with this statement? Explain. Regardless of whether you agree, what lessons can managers learn by studying perfectly competitive markets?
4. Explain why the demand curve facing a perfectly competitive firm is assumed to be perfectly elastic (i.e., horizontal at the going market price).
5. Explain why the demand curve facing a monopolist is less elastic than one facing a firm that operates in a monopolistically competitive market (all other factors held constant).
6. Use the model of perfect competition described in this chapter to explain, illustrate, or elaborate on the following statements.
 - a. "Increasing competition from new firms entering the market is good because it means one is in a good business."
 - b. "One important difference between an entrepreneur and a manager is that the former gets into a market before demand increases, while the latter gets into the market after the shift."
7. Explain the relationship between $P > AVC$ and a firm's contribution margin.
8. Why do economists consider zero economic profit to be "normal"?
9. "Economic profit" is a theoretical concept used to help explain the behavior of firms in competitive markets. Suggest ways in which this concept can actually be measured.

10. Explain why the $P = MC$ rule is the same as the $MR = MC$ rule for perfectly competitive firms.
11. Explain why a price-setting firm will always set its revenue-maximizing price below the price that would maximize its profit.
12. Provide some examples of business cases that a typical firm must consider. If possible, use current examples reported in the business press.
13. How “perfectly” competitive do you think are the following markets: (1) stock market, (2) bond market, (3) foreign exchange market, (4) world sugar market, and (5) world oil market? Explain.
14. Explain how the concept of “economic profit” might help explain the rationale for the government’s granting of monopolies to those firms that protect their product with a patent.

PROBLEMS

You may wish to consult Appendix 8A if you need help answering the more algebraic problems.

1. Following is the graphical representation of a short-run situation faced by a perfectly competitive firm. Is this a good market for this firm to be in? Explain. What do you expect will happen in the long run? Explain.



2. Indicate whether each of the following statements is true or false and explain why.
 - a. A competitive firm that is incurring a loss should immediately cease operations.
 - b. A pure monopoly does not have to worry about suffering losses because it has the power to set its prices at any level it desires.
 - c. In the long run, firms operating in perfect competition and monopolistic competition will tend to earn normal profits.
 - d. Assuming a linear demand curve, a firm that wants to maximize its revenue will charge a lower price than a firm that wants to maximize its profits.
 - e. If $P > AVC$, a firm’s total fixed cost will be greater than its loss.
 - f. When a firm is able to set its price, its price will always be less than its MR.
 - g. A monopoly will always earn economic profit because it is able to set any price that it wants to.
3. Kelson Electronics, a manufacturer of DVRs, estimates the following relation between its marginal cost of production and monthly output:

$$MC = \$150 + 0.005Q$$

- a. What does this function imply about the effect of the law of diminishing returns on Kelson’s short-run cost function?
- b. Calculate the marginal cost of production at 1,500, 2,000, and 3,500 units of output.
- c. Assume Kelson operates as a price taker in a competitive market. What is this firm’s profit-maximizing level of output if the market price is \$175?
- d. Compute Kelson’s short-run supply curve for its product.

4. A manufacturer of electronics products is considering entering the telephone equipment business. It estimates that if it were to begin making wireless telephones, its short-run cost function would be as follows:

Q (Thousands)	AVC	AC	MC
9	\$41.10	\$52.21	\$30.70
10	40.00	50.00	30.10
11	39.10	48.19	30.10
12	38.40	46.73	30.70
13	37.90	45.59	31.90
14	37.60	44.74	33.70
15	37.50	44.17	36.10
16	37.60	43.85	39.10
17	37.90	43.78	42.70
18	38.40	43.96	46.90
19	39.10	44.36	51.70
20	40.00	45.00	57.10

- Plot the average cost, average variable cost, marginal cost, and price on a graph.
 - Suppose the average wholesale price of a wireless phone is currently \$50. Do you think this company should enter the market? Explain. Indicate on the graph the amount of profit (or loss) earned by the firm at the optimal level of production.
 - Suppose the firm does enter the market and that over time increasing competition causes the price of telephones to fall to \$35. What impact will this have on the firm's production levels and profit? Explain. What would you advise this firm to do?
5. This same manufacturer of electronics products has just developed a handheld computer. Following is the cost schedule for producing these computers on a monthly basis. Also included is a schedule of prices and quantities that the firm believes it will be able to sell (based on previous market research).

Q (Thousands)	Price	MR	AVC	AC	MC
0	\$1,650				
1	1,570	\$1,570	\$1,281	\$2,281	\$1,281
2	1,490	1,410	1,134	1,634	987
3	1,410	1,250	1,009	1,342.33	759
4	1,330	1,090	906	1,156	597
5	1,250	930	825	1,025	501
6	1,170	770	766	932.67	471
7	1,090	610	729	871.86	507
8	1,010	450	714	839	609
9	930	290	721	832.11	777
10	850	130	750	850	1,011

- What price should the firm charge if it wants to maximize its profits in the short run?
- What arguments can be made for charging a price *higher* than this price? If a higher price is indeed established, what amount would you recommend? Explain.
- What arguments can be made for charging a *lower* price than the profit-maximizing level? If a lower price is indeed established, what amount would you recommend? Explain.

6. The manufacturer of high-quality flatbed scanners is trying to decide what price to set for its product. The costs of production and the demand for the product are assumed to be as follows:

$$\begin{aligned}TC &= 500,000 + 0.85Q + 0.015Q^2 \\Q &= 14,166 - 16.6P\end{aligned}$$

- a. Determine the short-run profit-maximizing price.
 - b. Plot this information on a graph showing AC, AVC, MC, P , and MR.
7. The demand and cost function for a company are estimated to be as follows:

$$\begin{aligned}P &= 100 - 8Q \\TC &= 50 + 80Q - 10Q^2 + 0.6Q^3\end{aligned}$$

- a. What price should the company charge if it wants to maximize its profit in the short run?
 - b. What price should it charge if it wants to maximize its revenue in the short run?
 - c. Suppose the company lacks confidence in the accuracy of cost estimates expressed in a cubic equation and simply wants to use a linear approximation. Suggest a linear representation of this cubic equation. What difference would it make on the recommended profit-maximizing and revenue-maximizing prices?
8. Overheard at the water cooler: “The demand and cost estimates that were provided at the meeting are very useful [$Q = 90 - 6.5P$ and $TC = 150 + 3.5Q$]. Unfortunately, what we didn’t realize at the time was that our fixed costs were underestimated by at least 30 percent. This means that we’ll have to adjust our price upward by at least 30 percent to cover the added fixed cost. In any case, there is no way in the world that we can survive by charging less than \$9 for our product.”
- a. Comment on this statement. Do you agree with the speaker? Explain. Illustrate your answer with the use of a graph indicating the firm’s short-run cost structure.
 - b. What price do you think this firm should charge if it wants to maximize its short-run profit?
9. Use the following equation to demonstrate why a firm producing at the output level where $MR = MC$ will also be able to maximize its total profit (i.e., be at the point where marginal profit is equal to zero).

$$\begin{aligned}P &= 170 - 5Q \\TC &= 40 + 50Q + 5Q^2\end{aligned}$$

10. “In a perfectly competitive market, a firm has to be either *good* or *lucky*.” Explain what is meant by this statement. Illustrate your answer with the use of the diagrams shown in Figures 8.5, 8.6, and 8.7.

The Use of Calculus in Pricing and Output Decisions



Thus far, we have discussed the firm's pricing and output decisions with the use of tabular and graphical examples. Using both the "total" approach and the "marginal" approach, we arrived at the $MR = MC$ rule for determining the optimal level of output and price for those firms able to exercise market power. As a supplement, we now explain the $MR = MC$ rule with the use of calculus.

To simplify our illustrations, we assume the firm has a quadratic total cost function, rather than the cubic function used throughout the examples in the previous sections of this chapter.

Perfect Competition

Suppose you are the owner and operator of a perfectly competitive firm with the following total cost function:

$$TC = 2,000 + 10Q + 0.02Q^2 \quad (8A.1)$$

Further, suppose the current market price is \$25. By definition, $TR = P \times Q$, so your total revenue function can be stated as:

$$TR = 25Q \quad (8A.2)$$

Profit (π) is defined as $TR - TC$. Therefore, using Equations (8A.1) and (8A.2), your firm's profit function can be expressed as:

$$\begin{aligned} \pi &= 25Q - (2,000 + 10Q + 0.02Q^2) \quad (8A.3) \\ &= 25Q - 2,000 - 10Q - 0.02Q^2 \\ &= -2,000 + 15Q - 0.02Q^2 \end{aligned}$$

The optimal output level (Q^*) can be found at the point where your firm's marginal profit is equal to zero. In other words, additional units of output should be produced as long as your firm earns additional profit from their sale. Using calculus, the marginal profit can be expressed as the first derivative of the profit function:

$$\frac{d\pi}{dQ} = 15 - 0.04Q \quad (8A.4)$$

Setting Equation (8A.4) equal to zero and solving for the optimal level of output (Q^*),

$$\begin{aligned} 15 - 0.04Q &= 0 \quad (8A.5) \\ Q^* &= 375 \end{aligned}$$

Returning to the total profit function presented in Equation (8A.3) and substituting Q^* for Q results in the following profit:

$$\begin{aligned} \pi &= -2,000 + 15(375) - 0.02(375)^2 \quad (8A.6) \\ &= 812.50 \end{aligned}$$

We conclude that at the price of \$25, the firm will earn maximum economic profit by producing 375 units of output per time period.

An alternative way of finding P^* and Q^* is to set the firm's marginal revenue function equal to its marginal cost function and then solve for Q^* . We already know that $MR = P$. The marginal cost function is the first derivative of the total cost function:

$$MC = \frac{dTC}{dQ} = 10 + 0.04Q \quad (8A.7)$$

Setting MR equal to Equation (8A.7) and solving for Q^* gives us

$$\begin{aligned} 25 &= 10 + 0.04Q \quad (8A.8) \\ 15 &= 0.04Q \\ Q^* &= 375 \end{aligned}$$

Comparison of Equations (8A.8) and (8A.5) provides a useful and concise explanation of the $MR = MC$ rule. As you can see, using this rule is the mathematical equivalent of finding the level of output that maximizes the total profit function. The equivalence of the two rules is because $\pi = TR - TC$. Given this, $d\pi/dQ = 0$ is the same as saying $d(TR - TC)/dQ = dTR/dQ - dTC/dQ = MR - MC = 0$. Regrouping we see $MR = MC$. The twin conditions for profit maximization are simply an application of the sum/difference rule for derivatives.

Had the total cost function been cubic in Equation (8A.1), finding the profit maximizing Q would have required the use of the quadratic formula;

however, the same rules would apply. To maximize profits, set $d\pi/dQ = 0$ or set $MR = MC$ and solve for Q .

Monopoly

As the manager of a product that only your company sells (e.g., a patent-protected product), suppose you are given the following information:

$$TC = 10,000 + 100Q + 0.02Q^2 \quad (8A.9)$$

$$Q_D = 20,000 - 100P \quad (8A.10)$$

You can use the same procedure employed in the case of perfect competition to find Q^* and P^* .

First, determine your marginal revenue function. Because you are a price setter and not a price taker, you cannot assume that $MR = P$. Instead, you must derive the marginal revenue function from your firm's demand function, shown in Equation (8A.10). Because your objective is to find the level of output that will maximize your profit (i.e., Q^*), you must rearrange the terms in the equation so price depends on the level of output:

$$P = 200 - 0.01Q \quad (8A.11)$$

By definition, $TR = P \times Q$. So by substitution,

$$\begin{aligned} TR &= (200 - 0.01Q)Q & (8A.12) \\ &= 200Q - 0.01Q^2 \end{aligned}$$

The marginal revenue function is the first derivative of the total revenue function:

$$MR = \frac{dTR}{dQ} = 200 - 0.02Q \quad (8A.13)$$

Notice that this is a specific example of the *MR has the same intercept and twice the slope as demand* rule that you learned in introductory microeconomics (and depicted in Figures 8.9 and 8.10). This is easy to prove: If inverse demand is $P = b - mQ$, then

$TR(Q) = bQ - mQ^2$ and $MR = dTR/dQ = b - 2mQ$. The derivative of the total cost function is the marginal cost function:

$$MC = \frac{dTC}{dQ} = 100 + 0.04Q \quad (8A.14)$$

Thus the $MR = MC$ rule is adhered to by setting Equation (8A.13) equal to Equation (8A.14) and solving for Q^* :

$$200 - 0.02Q = 100 + 0.04Q \quad (8A.15)$$

$$0.06Q = 100$$

$$Q^* = 1,667 \text{ (rounded to the nearest whole number)}$$

To find P^* we return to Equation (8A.11) and substitute Q^* for Q .

$$\begin{aligned} P &= 200 - 0.01(1,667) & (8A.16) \\ P^* &= \$183.33, \text{ or } \$183 \end{aligned}$$

At the rounded price of \$183, your firm can expect to sell 1,667 units of output per time period and earn an economic profit of \$73,333 (rounded to the nearest dollar). From the example on perfect competition, you should be aware of how the profit figure was determined.

As you can see from the preceding examples, the use of calculus offers a concise way of explaining the output decision for price-taking firms in perfectly competitive markets and the pricing/output decision for monopoly firms. The same procedures could be applied for those firms in monopolistic competition and even for oligopolistic firms that have clear-cut roles in their markets. However, tables and graphs similar to those used in this chapter provide the same answers as the calculus method. Our intention is for this appendix to serve as a supplement rather than an advanced treatment of the pricing/output decision.

Break-Even Analysis (Volume-Cost-Profit)

Learning Objectives

Upon completion of the appendix, readers should be able to:

- Understand the concept of break-even analysis.
- Calculate the break-even point and break-even revenue.
- Show how break-even analysis can be combined with demand analysis.
- Explain and apply the concept of degree of operating leverage.
- Describe the benefits and limitations of break-even analysis.

The Situation

October and November of each year are extremely busy months for the department of financial planning at Global Foods, Inc. It is during this period that the financial plan for the next 2 years is prepared. As is customary in business, greater emphasis is always placed on the first of the 2 years. Planning data are collected from all departments—covering projected sales, costs, and expenses. After checking as much as possible to ensure reasonability and accuracy, the department consolidates the numbers to obtain a planned income and expense statement. The plans are prepared along profit center lines, usually by specific flavors of the products.

Suzanne Prescott is the senior analyst responsible for the company's new energy drink product, "Go-nuts." She and her assistant have worked on this project for 2 weeks and have completed the profit plan, which she will present to the manager of the financial planning department. The first page of the long and detailed presentation shows the summary income statement for the "Go-nuts" profit center for the year 2013 (Table 8B.1).

Suzanne has kept her manager, Dorothy Simon, informed regarding the progress of the plan. As is quite common during a corporate planning cycle, their final discussion has been delayed several times due to late data, changed numbers, and missed schedules. Thus Suzanne and Dorothy are meeting just 1 day before the results are to be presented to the company's controller. Dorothy agrees with the method with which the plan has been put together and with the results Suzanne has presented. But she expects that the controller will require additional information. She asks Suzanne whether she has performed a sensitivity analysis calculating profit results if sales were to be 10 percent lower or 10 percent higher than planned. She is also interested in the level of sales at which profit would be zero to establish the "worst case." Suzanne admits that this analysis is incomplete due to lack of time. Because the presentation must be ready the next day, there is not enough time to rework the complete plan to obtain the alternative results. Suzanne will therefore have to devise a method by which she can obtain some good estimates for the "what-if" cases, estimates sufficiently reliable to show the controller. She remembers that in graduate school, she learned a method called break-even or volume-cost-profit analysis. Fortunately, she happens to have a few old textbooks in her office.

(continued)

(continued)

Table 8B.1 Income Statement for “Go-nuts” Profit Center

	Plan for 2013 (\$000)
Sales	\$5,000
Cost of sales:	
Materials	800
Labor	1,000
Overhead	950
Total cost of sales	<u>2,750</u>
Gross profit	<u>2,250</u>
Selling and administrative expenses	1,150
Research and development expenses	300
Total expenses	<u>1,450</u>
Net earnings before taxes	\$800

INTRODUCTION

The analysis described in this appendix is derived from price/output decision making in the short run, which is discussed in the body of Chapter 8. **Break-even analysis** (or volume-cost-profit analysis), which is used a great deal in actual business situations, is basically a simplification of the usual short-run analysis in economics. The following are the important elements that go into this analysis:

1. It retains the distinction between fixed and variable costs.
2. Typically it uses a straight-line total revenue curve. Thus, it implicitly assumes the existence of perfect competition because the price is considered to be the same regardless of quantity. (This procedure is followed, however, more for convenience than for theoretical reasons and can be changed.)¹⁵

¹⁵It also must be understood that the horizontal (quantity) axis on graphs used in economics tends to measure large quantity intervals. To claim that exact relationships hold over these long intervals is rather unrealistic. It is done for convenience and ease of exposition. But a firm will generally not consider such wide-ranging alternatives. It is much more likely that a company, given a particular level of production, will try to analyze some limited deviations from that level to, say, 5 to 10 percent higher or 5 to 10 percent lower. In such a limited interval, it is quite possible that a significant price change may not be necessary. Thus, even if the particular firm is not in a perfectly competitive market, a straight-line total revenue curve (i.e., no price change) in the relevant range may be close to reality.

3. Probably the most important deviation from the analysis in Chapter 8 is that break-even analysis employs a straight-line total variable cost curve. This means that marginal cost and average variable cost are both constants and that marginal and average variable costs are equal to one another. Although such an assumption may not seem realistic over large quantity intervals, it may not be unreasonable when a relatively limited range of quantities is considered. The possible existence of constant unit variable costs over a range of quantities has already been mentioned in the discussion of empirical cost estimates.

Figure 8B.1 presents a graph depicting the typical break-even analysis chart.

There are, however, some important differences between customary economic analysis and the break-even method that should not be overlooked:

1. The short-run economic chart shows two points where economic profits are zero, and maximum profit is identified somewhere between these two points. In break-even analysis, there is only one no-profit (break-even) point. As quantities rise beyond this point, profit increases continuously until, presumably, capacity is reached, and no additional quantities of product can be achieved.¹⁶ At this point, costs

¹⁶Because break-even analysis deals with the short run only, capacity does not change.

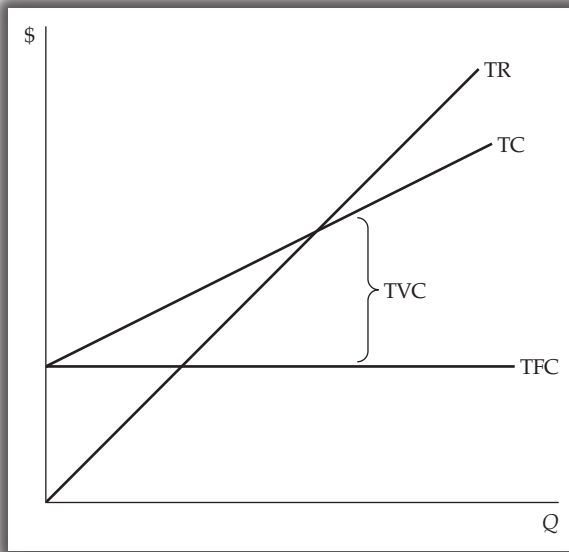


Figure 8B.1 Total Revenue and Cost Curves (Constant Average Variable Cost)

become infinitely high, and the total cost line would thus cross the total revenue line.

2. There is also a major difference between the objectives of the two analyses. In Chapter 8, interest focused on the question of resource allocation—the effect of a price change or a cost change on the quantity produced. In break-even analysis, the question is what impact does a change in quantity have on variable costs and profit?
3. The third difference lies in the use of the cost concept. Economic costs, as previously discussed, are based on replacement costs and include imputed costs and normal profit. Break-even analysis, as practiced in business, usually relies on accounting costs (often standard costs as used in cost accounting), which include explicit costs only and represent historical data. However, in this case, careful treatment of data could convert accounting data into economic cost data. For instance, as is shown later, a “required” minimum profit, which can represent the normal profit, can be included in the calculation.

Another point to consider is that break-even analysis does not consider the time dimension of business activity. We discussed this dimension in Chapter 2 in connection with the maximization of shareholder wealth. In Chapter 12, we deal with the question of time and the time value of money as well as risk. We have included a brief discussion in Chapter 12 of the calculation of the time value of money and present values in solving for break-evens.

The Break-Even Point



Module 8BA
Module 8BB

The first of the calculations we examine identifies the quantity at which the company will just break even—no profit, no loss. But this is not the point where company executives wish the firm to be. The break-even point merely sets the stage to investigate the relationship among quantity of the product, the cost to produce this quantity, and the profit—hence the name *volume-cost-profit* analysis.

The same abbreviations used in previous chapters are used here:

$P =$ Price	$TC =$ Total cost
$TVC =$ Total variable cost	$Q =$ Quantity produced
$AVC =$ Average variable cost	$TR =$ Total revenue
$TFC =$ Total fixed cost	$\pi =$ Profit

The very simple equation for profit is:

$$\begin{aligned} \pi &= TR - TC \\ &= TR - TVC - TFC \\ &= (P \times Q) - (AVC \times Q) - TFC \\ &= Q(P - AVC) - TFC \end{aligned}$$

To obtain the **break-even point**, total revenue is set equal to total cost:

$$\begin{aligned} TR &= TVC + TFC \\ (P \times Q) &= (AVC \times Q) + TFC \\ (P \times Q) - (AVC \times Q) &= TFC \\ Q(P - AVC) &= TFC \end{aligned}$$

Table 8B.2 Break-Even Analysis

Variables	
Price per unit	\$5.00
Variable cost per unit	\$3.00
Total fixed cost	\$20,000
Results	
Break-even quantity	10,000
Break-even revenue	\$50,000

Units	Fixed Cost	Variable Cost	Total Cost	Revenue	Profit
0	\$20,000	\$ 0	\$ 20,000	\$ 0	\$-20,000
5,000	20,000	15,000	35,000	25,000	-10,000
10,000	20,000	30,000	50,000	50,000	0
15,000	20,000	45,000	65,000	75,000	10,000
20,000	20,000	60,000	80,000	100,000	20,000
25,000	20,000	75,000	95,000	125,000	30,000
30,000	20,000	90,000	110,000	150,000	40,000
35,000	20,000	105,000	125,000	175,000	50,000
40,000	20,000	120,000	140,000	200,000	60,000

Thus the break-even quantity is

$$Q = \text{TFC} / (P - \text{AVC})$$

For example, if $P = 5$, $\text{AVC} = 3$, and $\text{TFC} = 20,000$,

$$Q = 20,000 / (5 - 3) = 20,000 / 2 = 10,000$$

This result can be checked as follows:

Total revenue (10,000 × \$5)	\$50,000
Total variable cost (10,000 × \$3)	30,000
Total fixed cost	20,000
Total cost	50,000
Profit	\$0

If the quantity produced is larger than 10,000 units, a profit will result. If quantity drops below 10,000, the company will incur a loss. Table 8B.2 illustrates the revenue, costs, and profits resulting from changes in quantity, and Figure 8B.2 graphs the results.

What happens to the break-even point when one or more of the variables change? An increase in the average variable cost will increase the slope

of the total cost curve, and increase the break-even point (a decrease will cause the opposite). A change in the unit price will change the slope of the total revenue curve; a price increase (decrease) will decrease (increase) the break-even point.

An increase (decrease) in fixed costs will cause a parallel shift up (down) in the costs curve and an increase (decrease) in the break-even point.¹⁷

Break-Even Revenue



Module 8BC

Under certain circumstances, the product price and the unit variable costs may not be available. This will happen—frequently—when more than one product is produced in a plant. Because each of the different products being manufactured side by side has a different price and a different unit variable cost, it is difficult to use the formula of the previous section to establish the break-even point of the plant.

¹⁷The reader should substitute new numbers for those shown in Table 8B.2 and Figure 8B.2, and observe the results. Modules 8BA and 8BB will perform this analysis quickly.

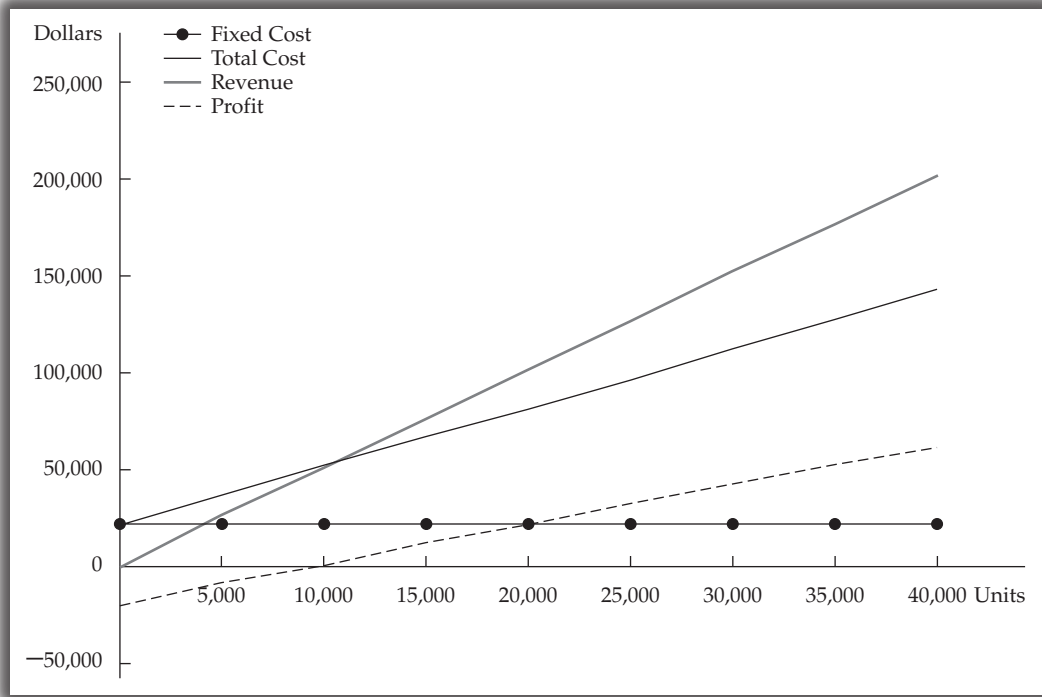


Figure 8B.2 Break-Even Analysis

A weighted price and unit cost can be calculated for the products if the various products are produced in constant proportions. If we assume variable costs are a constant percentage of total revenue, then we can calculate **break-even revenue** directly. Again, we are assuming the relationship between average variable cost and price for each product remains the same and the quantities of the various products are produced in constant proportions. Such an assumption does not appear to be unrealistic for relatively small changes in total revenue.

Starting with the equation that shows revenue at break-even,

$$TR = TVC + TFC$$

we now convert TVC into a constant fraction of total revenue, $TVC = a \times TR$, where a is a constant less than 1. The break-even total revenue then becomes

$$TR = TFC / (1 - a)$$

For example, if $TFC = 20,000$ and $a = 0.6$, then break-even TR equals 50,000, which is the same result as obtained previously, because if $P = 5$, and $AVC = 3$, $a = 0.6$ expresses the same relationship.

Required Profit

If the only objective of this analysis were to find the point at which a plant or a company breaks even, not much would be accomplished. For a company to prosper, it must earn profits, not just break even. If a firm has a particular dollar profit objective per period, a small adjustment of the break-even equation will provide the appropriate output measure. A specific, fixed dollar amount of **required profit** can be handled as an addition to fixed cost.

Continuing the illustration that has been used throughout this chapter, if the owners of the firm require a \$10,000 profit, the equation is altered as follows:

$$\begin{aligned} Q_{\pi}^f &= (TFC + \text{Profit requirement}) / (P - AVC) \\ &= (20,000 + 10,000) / (5 - 3) \\ &= 30,000 / 2 = 15,000 \end{aligned}$$

where Q_{π}^f stands for break-even with a fixed profit requirement.

If a specific profit per unit of product is required rather than a fixed amount of profit as above, this unit profit must be added to the average variable cost. For instance, suppose the company's

objective is 40 cents profit per unit. Then AVC is changed to \$3.40, and

$$Q_{\pi}^u = \text{TFC} / (P - (\text{AVC} + \pi \text{ per unit})) \\ = 20,000 / 1.60 = 12,500$$

Earlier in this chapter it was stated that the calculation of cost in volume-cost-profit analysis usually involves the use of accounting data. However, the “required profit” concept can easily be interpreted to represent the implied or opportunity costs that economists find crucial to the analysis of the firm. Thus, the profit measure, whether total or per unit, can be the normal profit, which is the minimum amount necessary to cause the owner to continue operating this business.

Combining Break-Even Analysis with Demand¹⁸

Until now, we have assumed a specific sales price to obtain our results. Going one step further, we can calculate break-even points at different prices by substituting those prices into our basic formula. We use the same numbers we used before (i.e., TFC = \$20,000 and AVC = \$3). Table 8B.3 gives us the resulting break-even prices and quantities when we set prices from \$3.75 to \$8.

The results of Table 8B.3 can now be plotted on a graph (Figure 8B.3). We obtain what we can call a zero profit “iso-profit” curve. Note that the curve approaches the level of average variable cost asymptotically. The curve also approaches the vertical axis at extremely high prices. We have now established the break-even point at a large range of prices.

The next step is to combine the iso-profit curve with the demand curve. The company’s economist has estimated the demand curve for this product to be $Q = (8 - P)4,000$. In Table 8B.4, we show quantities and prices. In Figure 8B.4, we combine the iso-profit curve with the demand curve D . The demand curve intersects the iso-profit curve at two points, at somewhat less than \$7 (point A) and at about \$4.50 (point B). At these points, there is no profit. At prices between points A and B , there will be a positive profit, whereas above and below the intersections a loss is incurred. The highest profit

Table 8B.3 Break-Even Prices and Quantities

Price	Quantity
\$8	4,000
7	5,000
6	6,667
5	10,000
4.50	13,333
4	20,000
3.75	26,667

is obtained where the iso-profit curve is the furthest horizontally to the left of the demand curve.¹⁹

Degree of Operating Leverage



Module 8BD

Because volume-cost-profit analysis is concerned with the effect of a change in quantity of product on the profits of a firm, we must develop a method to quantify this effect. Such a method, called **degree of operating leverage (DOL)** is, in fact, a type of elasticity formula. The calculation result is a coefficient that measures the effect a percentage change in quantity has on the percentage change in profit.

$$\text{DOL} = \frac{\% \Delta \pi}{\% \Delta Q}$$

The percent change in profit can be written as follows:

$$\% \Delta \pi = \frac{\Delta \pi}{\pi} = \frac{\Delta Q(P - \text{AVC})}{Q(P - \text{AVC}) - \text{TFC}}$$

The percent change in quantity equals $\Delta Q/Q$. Thus, we obtain the result,

$$\text{DOL} = \frac{Q(P - \text{AVC})}{Q(P - \text{AVC}) - \text{TFC}}$$

To better explain the meaning of DOL, let us use the example of the preceding section. At a production of 15,000 units, profit was \$10,000. DOL can now be measured at the 15,000 quantity:

$$\text{DOL} = \frac{15,000(5 - 3)}{15,000(5 - 3) - 20,000} = \frac{30,000}{10,000} = 3$$

¹⁸This section is based on Jon M. Hawes, Michael F. D’Amico and Thomas L. Baker, “Simultaneous Use of Break-Even and Demand Analysis for Pricing Decisions: A Teaching Method,” *Journal of Education for Business*, May/June 1995, pp. 285–89.

¹⁹We could also plot additional iso-cost curves to the right of the break-even iso-cost curve. Where the demand curve is just tangent to the highest possible iso-cost curve, profit is maximized.

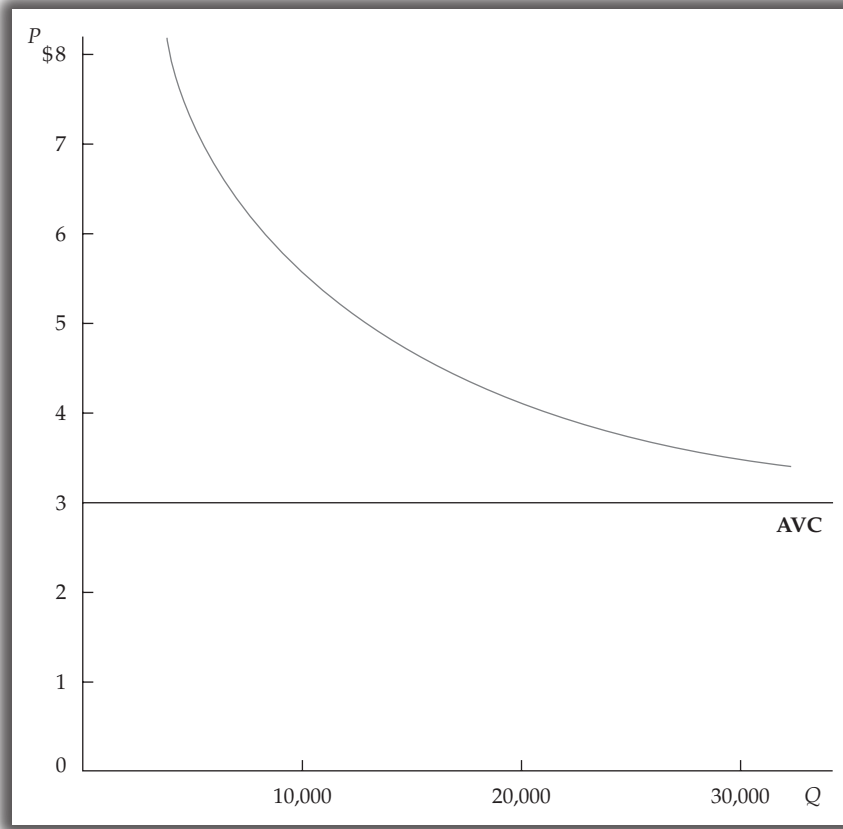


Figure 8B.3 The Zero Profit Iso-Profit Curve

DOL = 3 means that, at $Q = 15,000$, a 1 percent change in quantity will result in a 3 percent change in profit (and a 10 percent change in Q will lead to a 30 percent change in profit). The DOL effect can

be seen in terms of an income statement at quantities of 13,500 and 16,500 (a 10 percent decrease and a 10 percent increase):

Table 8B.4 Demand Curve, Prices, and Quantities

Price	Quantity
\$8	0
7	4,000
6	8,000
5	12,000
4	16,000
3	20,000
2	24,000
1	28,000
0	32,000

	$Q = 13,500$	$Q = 15,000$	$Q = 16,500$
Total revenue	\$67,500	\$75,000	\$82,500
Total fixed cost	20,000	20,000	20,000
Total variable cost	40,500	45,000	49,500
Total cost	60,500	65,000	69,500
Profit	\$7,000	\$10,000	\$13,000

At a quantity of 13,500, profit is \$7,000, or 30 percent less than at a quantity of 15,000 units. Conversely, at a quantity of 16,500 units, the profit is \$13,000, or 30 percent larger.²⁰

²⁰The degree of operating leverage can be calculated at any point in the profit or loss area. However, it cannot be calculated at the break-even quantity because a percentage change in profit from zero does not make sense. (The denominator of the DOL formula is zero at break-even quantity.)

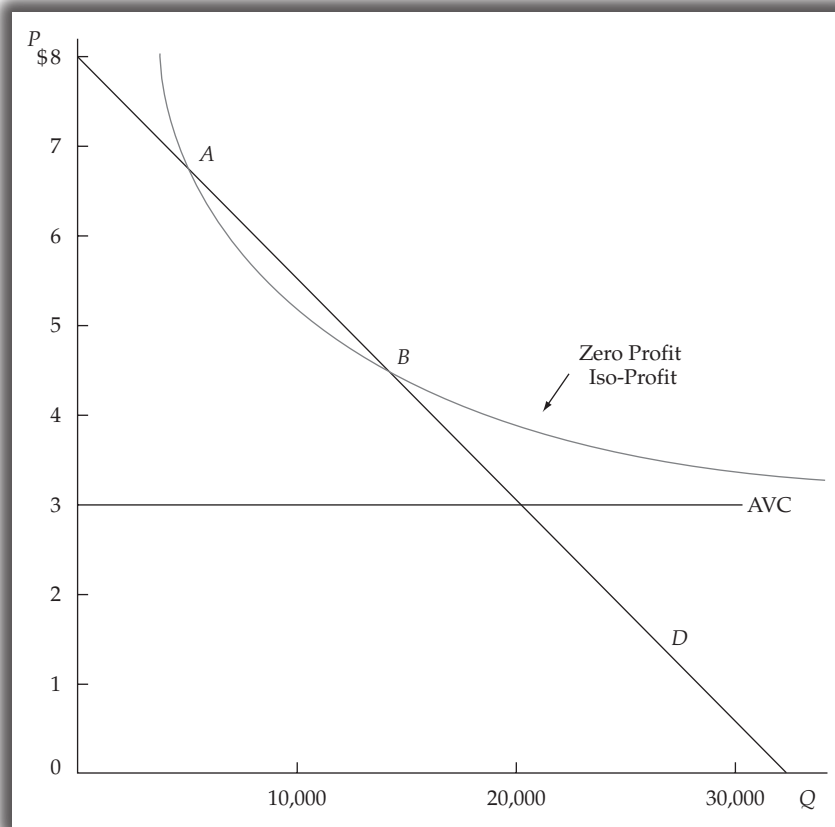


Figure 8B.4 The Iso-Profit and Demand Curves

The importance of the degree of operating leverage is that it reveals to management the effect on profits of a small change in quantity. This construction will hold only, of course, as long as all variables remain the same (i.e., price, average variable cost, and total fixed cost).

The relative sizes of fixed and variable costs influence the level of the DOL coefficient. A plant with high fixed costs and low variable costs will have a higher DOL than a plant with lower fixed costs and higher variable costs. The former plant will also have a higher break-even point. The significance of this relationship is that a firm with high fixed costs—a capital-intensive firm—will usually achieve break-even at a higher quantity, but because it has a higher DOL, its profits will grow at a relatively high rate when production rises above break-even. Its profits will also decline more quickly during downturns in economic activity, and the firm will become unprofitable at a relatively high point of production (since its break-even quantity will be

high). In contrast, a plant with lower fixed costs and higher variable costs—a labor-intensive plant, perhaps somewhat obsolete—will break even at lower quantities, and its profits will tend to rise or fall less quickly when quantity produced moves up or down.

Thus, the break-even quantity and the DOL can have a very significant influence on a firm deciding whether to convert from an old—labor-intensive—manufacturing facility to a more modern, automated (i.e., capital-intensive) plant. For example, using the income statement presented at the beginning of this appendix, let us assume the average variable costs per unit are \$2.75 and total fixed costs are \$1,450.²¹ The break-even point for a \$5 price is easily calculated as 644 units (rounded to the nearest integer).

Now suppose additional, more up-to-date machinery were installed in the plant, increasing annual fixed costs to \$2,000 and driving average

²¹All dollar and unit numbers (except price and variable cost per unit) are in thousands.

Table 8B.5 Break-Even and DOL Data for Old Versus Modernized Plants

	Old Plant	New Plant
Price per unit	\$5.00	\$5.00
Variable cost per unit	\$2.75	\$2.25
Total fixed cost	\$1,450	\$2,000
Break-even quantity	644	727
Break-even revenue	\$3,222	\$3,636
<i>Equal Profit</i>		
Quantity	1,100	1,100
Profit	\$1,025	\$1,025
Degree of operating leverage at equal-profit quantity	2.41	2.95

variable costs down to \$2.25. The break-even point for the modernized plant will have increased to 727 units. Why, then, should the company invest in the new machinery? Because the newly equipped plant enjoys a higher degree of operating leverage, its profits will rise more quickly with increases in production. At some quantity, the modernized plant and the old one will achieve equal profits. In this case, the point of equality is 1,100 units, at which both plants will show a profit of \$1,025.²²

At the production level of 1,100 units, the DOL values for the old and modernized plants are 2.41 and 2.95, respectively. Thus, if more than 1,100 units are to be produced annually in the future, the profit from the modernized plant will be greater. If, however, the quantity is expected to remain at 1,000 units per year as in the company's plan, modernization would not be advisable at this time. The data for the present example are shown in Table 8B.5.

The Uses and Limitations of Volume-Cost-Profit Analysis

Volume-cost-profit analysis is a useful tool under certain circumstances, but its limitations must be

²²This point is calculated as follows:

$$Q(P - AVC_a) - TFC_a = Q(P - AVC_b) - TFC_b$$

where the subscripts *a* and *b* denote the old plant and the modernized plant, respectively.

$$\begin{aligned} Q(2.25) - 1,450 &= Q(2.75) - 2,000 \\ 0.5Q &= 550 \\ Q &= 1,100 \end{aligned}$$

The profit for a quantity of 1,100 units is \$1,025 for each plant.

understood. When a corporation prepares its financial plan for the next year or even the next 2 years, it usually engages in what is referred to as "bottom-up" planning, a process that is not only time consuming but extremely detailed. Many parts of the corporate organization contribute data on sales forecasts, prices, manufacturing costs, administrative and marketing expenses, and other measures. Data may be generated from every department in the corporation. Consolidating these data and making various changes in them before bringing a final plan up to top management for approval is a mammoth undertaking. A corporation would not use volume-cost-profit analysis for this type of planning.

The main use of this analysis lies in calculating alternative cases in a restricted period of time. It can also be used to make small, relatively quick corrections. In addition, during early stages of the plan, when detailed data are not yet available, estimates using variable and fixed costs can be used to establish some rough benchmarks for the eventual detailed plan.

However, despite its usefulness, break-even analysis has important limitations, some of which have already been mentioned:

1. It assumes the existence of linear relationships, constant prices, and constant average variable costs. However, when the effects of relatively small changes in quantity are measured, linear revenues and variable costs are certainly good approximations of reality.
2. It is assumed costs (and expenses) are either variable or fixed. The existence of fixed costs limits this analysis to the short run. Changes in capacity are ordinarily not considered.

- For break-even analysis to be used, only a single product must be produced in a plant or, if there are several products, their mix must remain constant.
- The analysis does not result in identification of an optimal point; it focuses on evaluating the effect of changes in quantity on cost and profits.

An Application: The Restaurant Industry

The following is an illustration of the application of break-even analysis to the restaurant industry. Although the numbers that follow are hypothetical, they are based on a survey conducted in conjunction with the National Restaurant Association. The study compared a full-menu restaurant with a fast-food restaurant. Costs and expenses were classified into fixed and variable. It is interesting to note that a large percentage of payroll costs was considered fixed (for skeleton staff necessary to run a restaurant).

	Full-Menu Restaurant	Fast-Food Restaurant
Revenue	\$950,500	\$622,100
Fixed cost	445,700	260,700
Variable cost	459,500	280,900
Profit	\$ 45,300	\$ 80,500

From the preceding numbers, several inferences result:

- The fast-food restaurant has a higher profit margin than the full-menu restaurant, 12.9 percent against 4.8 percent.
- The break-even points are at revenues of \$862,800 and \$475,300, respectively.
- The degree of operating leverage for the full-menu restaurant is 10.8 against the fast-food restaurant's 4.2 at the revenue levels given previously.²³ An increase in the full-menu restaurant's revenue would increase its profits at a much faster rate than a similar revenue increase in the fast-food restaurant's revenue.

Thus, the study concludes that volume-cost-profit analysis "can help you understand your establishment's cost structure, and help improve your decision making by quantifying the effect of specific policy decisions on the bottom line."²⁴

²³The formula for the degree of operating leverage is $TR(1 - a) / [TR(1 - a) - TFC]$, where a is the fraction TVC/TR .

²⁴Carol Greenberg, "Analyzing Restaurant Performance," *The Cornell H.R.A. Quarterly*, May 1986, pp. 9-11.

The Solution

To calculate results for "Go-nuts" sales of 10 percent more and 10 percent less, Suzanne Prescott needs to make estimates of price per unit, average variable cost, and total fixed cost. She had the production figure and price per case of \$5. Breaking up the cost and expense numbers into fixed and variable components was a much more difficult task. Working with her assistant, she arrived at the following estimated breakdown:

Cost of sales (\$000)	
Variable materials and labor	\$1,800
Variable overhead	500
Fixed overhead	450
Expenses (\$000)	
Variable selling and administrative	450
Fixed selling and administrative	700
Fixed research and development	300

Thus, total fixed costs were found to be \$1,450, and total variable cost \$2,750 or \$2.75 per unit.

She calculated the break-even point as follows:

$$\frac{1,450}{5 - 2.75} = 644$$

For the +10 percent and -10 percent, she used the DOL equation:

$$\frac{1,000(5 - 2.75)}{1,000(5 - 2.75) - 1,450} = \frac{2,250}{800} = 2.8125$$

For every 1 percent change in quantity, profit will change by 2.8125 percent. Thus, if quantity changes by 100 units (i.e., 10 percent), profit will change by 28.125 percent, from its \$800 level down to \$575 or up to \$1,025.

Suzanne proceeded to prepare the presentation chart showing the planned figures, the two 10 percent variations, and the worst case. Table 8B.6 illustrates the results. She is now ready for the next day's meeting with the controller.

(continued)

(continued)

Table 8B.6 Alternative Plans for Year 2013 (\$000)

	Best Estimate	+10%	-10%	Worst Case
Quantity	1,000	1,100	900	644
Sales	\$5,000	\$5,500	\$ 4,500	\$ 3,222
Cost of Sales				
Variable mat. and labor	\$1,800	\$1,980	\$ 1,620	\$ 1,160
Variable overhead	500	550	450	322
Fixed overhead	450	450	450	450
Total cost of sales	<u>2,750</u>	<u>2,980</u>	<u>2,520</u>	<u>1,932</u>
Gross Profit	\$2,250	\$2,520	\$ 1,980	\$ 1,290
Expenses				
Variable selling and admin.	\$ 450	\$ 495	\$ 405	\$ 290
Fixed selling and admin.	700	700	700	700
Fixed research and devel.	300	300	300	300
Total expenses	<u>\$1,450</u>	<u>\$1,495</u>	<u>\$ 1,405</u>	<u>\$ 1,290</u>
Net earnings before taxes	<u>\$ 800</u>	<u>\$1,025</u>	<u>\$ 575</u>	<u>\$ (0)</u>

A Marketing Application: The Wine Industry

It may surprise you to learn that wineries do not need to own vineyards to make wine; there is an active market for wine grapes and bulk wine. The wine industry is product differentiation on steroids; as a result, it does not readily fit into the perfectly competitive model. Nonetheless, the break-even model is a common tool used to provide back-of-the-envelope estimates of how many bottles would need to be sold in order to cover brand introduction costs. Although the analysis can be done on the back of an envelope, it is better to use Excel because comparative statics analysis is easily accomplished in this instance.

Three variations using the same basic data will be presented here, in increasing degrees of realism. The first uses three pieces of data, cells A1:A3 of Excel Sheet 8B.1; the second adds two more, A4 and A5; and the third adds one final element, A6, to the analysis. Excel Sheet 8B.1 is a screenshot of the Breakeven Analysis Excel App.

Version I

Suppose that you own an existing winery (with plenty of excess bottling and storage capacity) and

that you are considering introducing a new brand of wine. Your estimated AVC of production is constant at \$5.00 per bottle. You intend to introduce the wine at \$10.00 per bottle (to the wholesaler). You estimate that advertising and promotion costs to launch a new brand should come to \$100,000 in the first year. How many bottles must you sell in order to break even by the end of one year?

Solution I. The answer here is simple: $FC / (P - AVC) = 100,000 / (10 - 5) = 20,000$ bottles. Excel is really not required for this version, but it is done in $B10 = A1 / (A3 - A2)$.

Version II

You realize that, in order to sweeten the deal in the first year and to build brand loyalty, you will have to offer a \$3.00 per bottle discount during 4 of the first 12 months (i.e., the discounted price will be \$7.00 to the wholesaler during these months; marketers in the wine industry would say that the wine goes “on allocation” for 4 months). How many bottles must you sell in order to break even by the end of 1 year?

Solution II. The key is to figure out what the average price for a bottle will be in the first year; from there, the problem is easy. One-third of the time,

Excel Sheet 8B.1



	A	B	C	D	E	F	G
1	\$ 100,000	FC		Version III: Incorporating demand elasticity			
2	\$ 5.00	AVC			-30%	%ΔP	
3	\$ 10.00	P_{full}			150%	%ΔQ at discount	
4	\$ 7.00	$P_{discount}$		$N_{mo.}$	Sales/mo.	Sales = $N_{mo.}$ ·Sales/mo.	
5	4	$N_{mo.}$ at discount		8	100%	8	At full
6	-5.00	$\epsilon = \% \Delta Q / \% \Delta P$		4	250%	10	At discount
7	Data Area					18	Total sales in year
8							
9		Version I		Version II	Version III		
10	Q break-even	20,000		25,000	30,000	$Q_{BE} = FC / (P - AVC)$	
11	Price, P	\$ 10.00		\$ 9.00	\$ 8.33	$P = f \cdot P_{full} + d \cdot P_{discount}$	
12							
13		Fraction of months		of sales			
14		at full, f		2/3	4/9	= At full/Total sales	
15		at discount, d		1/3	5/9	= At discount/Total sales	

Source: Breakeven Analysis Excel App, worksheet 8B.1

the wine is at discount; the rest of the time, it is at full price. The weighted average price in D11 is therefore $1/3 \cdot 7 + 2/3 \cdot 10$, or \$9.00. Given this, the break-even quantity is 25,000 in D10.

Version III

Historically, new wine brands have little consumer loyalty and face substantial competition—both factors lead to a high price elasticity of demand for new wine brands. You estimate the price elasticity of demand for your new wine brands to be -5.0 . (Assume for simplicity that there is no seasonality involved in this product entry; 1 month of the year is as good as another.) How many bottles must you sell in order to break even by the end of one year?

Solution III. The key here is to incorporate demand elasticity information into your calculation of weighted average price. The law of demand says that you expect to sell more wine in discounted months than in full-price months; the question is, How much more? Elasticity of demand answers that question. Seven dollars is a 30 percent discount off \$10 (E2). Given an elasticity of -5 , a 30 percent decrease in price will increase unit sales by 150 percent (E3). In a year, you will have 8 months

of sales at full price (D5:G5) and 4 months at 2.5 times full price (D6:G6). Total sales will be the size of 18 full-price months of sales in the 12-month period (F7).

Even though two-thirds of the months are at full price, less than one half of sales are at full price. In particular, $8/18 = 4/9$ is a full price (E14) and $10/18 = 5/9$ is at discount (E15). The weighted average price is therefore \$8.33 (E11), and the break-even quantity is 30,000 in E10.

The data for these problems have been chosen to highlight the series of calculations needed to answer the question. Allocations often occur on a weekly rather than a monthly basis, and there is a danger in having a new brand on allocation too often during the first year. The last thing the marketing director wants is to signal that the new brand is not really at the price point. On the other hand, if they do not provide allocations at all, then there is a good chance that wholesalers, and more important, retailers, will not be as likely to provide shelf space on a very crowded field of available wine choices. The Breakeven Analysis Excel App allows you to adjust the assumptions of the model and track what happens to the break-even quantity as a result.

SUMMARY

Break-even (volume-cost-profit) analysis is a simplification of the economic analysis of the firm. It involves several limiting assumptions, such as constant prices and constant average variable costs. Because fixed costs are an essential component of this technique, it is strictly a short-run tool. Despite these simplifications—and possibly because of them—break-even analysis is a very useful aid to an economic or financial analyst. It is necessary, however, to be aware of the method's limitations.

Several specific tools are discussed. The first is the break-even formula itself stated in terms of quantities of production units. If a single product cannot be identified, then a break-even formula for total revenue can be used.

Because the firm's objective is certainly not to break even but to achieve profitability, an equation was developed to identify the required quantity of production, given a lump-sum profit requirement or a profit-per-unit requirement.

Break-even analysis can also be combined with demand analysis to find the most profitable price and quantity.

To measure the effect of change in quantity on profits, the concept of degree of operating leverage is introduced. This elasticity-like formula measures the relation between a percentage change in quantity sold and a percentage change in profit. This equation is also shown to be useful in comparing two plants employing differing technologies (and therefore having different relationships between fixed and variable costs) or in making decisions on modernizing a plant.

The usefulness of break-even analysis in evaluating alternatives and in making quick corrections is discussed. The limitations in the application of this technique are also pointed out.

IMPORTANT CONCEPTS

Break-even analysis. Also called *volume-cost-profit analysis*, a simplification of the economic analysis of the firm that measures the effect of a change in quantity of a product on the profits of the firm. (p. 346)

Break-even point. The level of output at which the firm realizes no profit and no loss. (p. 347)

Break-even revenue. The amount of revenue at which the firm realizes no profit and no loss. (p. 349)

Degree of operating leverage (DOL). An elasticity-like formula that measures the percentage change in profit resulting from a percentage change in quantity produced or revenue. (p. 350)

Required profit. Profit that can represent the opportunity cost or the normal profit and that can be

incorporated in the break-even formula. A fixed dollar amount of required profit can be handled as an addition to fixed cost; a specific profit per unit of product can be added to the average variable cost. (p. 349)

Total fixed cost (TFC). A cost that remains constant as the level of output varies. In a short-run analysis, fixed cost is incurred even if the firm produces no output. Also referred to simply as *fixed cost*. (p. 347)

Total variable cost (TVC). The total cost associated with the level of output. This can also be considered the total cost to a firm of using its variable inputs. Also referred to simply as *variable cost*. (p. 347)

QUESTIONS

1. Although volume-cost-profit analysis uses graphs similar to those used by economists, the analysis differs in content. Discuss these differences.
2. Does the volume-cost-profit method analyze short-run or long-run situations? Why?
3. What is the difference between fixed costs and constant costs?
4. How realistic is the assumption of constant variable unit costs in volume-cost-profit analysis? Does it detract a great deal from the value of this analysis? Explain briefly.
5. What is the effect on break-even quantity of
 - a. A decrease in unit price?
 - b. A decrease in average variable cost?
 - c. A decrease in fixed cost?
 Assume some numbers and illustrate the effect by drawing graphs showing the break-even point.

6. Business risk is usually defined in terms of variations of return (or profit) to a firm due to changes in activity resulting from changes in general economic activity. Can the degree of operating leverage therefore be described as a measure of business risk? Why?
7. Would you expect a company whose production is rather stable from period to period and growing slowly from year to year to have relatively high fixed costs?
8. How would you account for required profit in the break-even formula when
 - a. Profit is set as a requirement for a time period (e.g., a year)?
 - b. Profit is set as a specific monetary amount per unit?
9. Can the degree of operating leverage be measured at the break-even quantity point? Why or why not?
10. Is volume-cost-profit analysis a good planning tool? Discuss briefly.
11. What are some useful applications of volume-cost-profit analysis?

PROBLEMS

1. The Automotive Supply Company has a small plant that produces speedometers exclusively. Its annual fixed costs are \$30,000, and its variable costs are \$10 per unit. It can sell a speedometer for \$25.
 - a. How many speedometers must the company sell to break even?
 - b. What is the break-even revenue?
 - c. The company sold 3,000 units last year. What was its profit?
 - d. Next year's fixed costs are expected to rise to \$37,500. What will be the break-even quantity?
 - e. If the company will sell the number of units obtained in part *d* and wants to maintain the same profit as last year, what will its new price have to be?
2. Writers' Pleasure, Inc., produces gold-plated pen and pencil sets. Its plant has a fixed annual cost of \$50,000, and the variable unit cost is \$20. It expects to sell 5,000 sets next year.
 - a. To just break even, how much will the company have to charge for each set?
 - b. Based on its plant investment, the company requires an annual profit of \$30,000. How much will it have to charge per set to obtain this profit? (Quantity sold will still be 5,000 sets.)
 - c. If the company wants to earn a markup of 50 percent on its variable costs, how many sets will it have to sell at the price obtained in part *b*?
3. Bikes-for-Two, Inc., produces tandem bicycles. Its costs have been analyzed as follows:

<i>Variable Cost</i>	
Materials	\$30/unit
Manufacturing labor	3 hours/unit (\$8/hour)
Assembly labor	1 hour/unit (\$8/hour)
Packing materials	\$3/unit
Packing labor	20 minutes/unit (\$6/hour)
Shipping cost	\$10/unit
<i>Fixed Costs</i>	
Overhead labor	\$50,000/year
Utilities	\$5,000/year
Plant operation	\$65,000/year
<i>Selling price</i>	\$100/unit

- a. Calculate the break-even quantity.
- b. Calculate the break-even revenue.
- c. Develop a chart to show profits at quantities of 2,000, 4,000, 6,000, 8,000, and 10,000.

4. Music Makers Company, a wholesale distributor, is considering discontinuance of its line of CDs due to stiff competition from MP3 downloads and other new, technologically advanced recordings. The variable cost of its CDs last year was about 40 percent of its CD revenue, and the allocated fixed cost equaled \$100,000 per year. Last year's sales were \$250,000, but it is expected that in the future, annual revenue will drop by 20 percent and variable costs will rise to 50 percent of revenue (because of price reductions). Will CDs still be profitable for the company?
5. The ABC Company sells widgets at \$9 each; variable unit cost is \$6, and fixed cost is \$60,000 per year.
 - a. What is the break-even quantity point?
 - b. How many units must the company sell per year to achieve a profit of \$15,000?
 - c. What will be the degree of operating leverage at the quantity sold in part *a*? In part *b*?
 - d. What will be the degree of operating leverage if 30,000 units are sold per year?
6. Two companies, Perfect Lawn Co. and Ideal Grass Co., are competing in the manufacture and sale of lawn mowers. Perfect has a somewhat older plant and requires a variable cost of \$150 per lawn mower; its fixed costs are \$200,000 per year. Ideal's plant is more automated and thus has lower unit variable costs of \$100; its fixed cost is \$400,000. Because the two companies are close competitors, they both sell their product at \$250 per unit.
 - a. What is the break-even quantity for each?
 - b. At which quantity would the two companies have equal profits?
 - c. At the quantity obtained in part *b*, what is each company's degree of operating leverage?
 - d. If sales of each company were to reach 4,500 units per year, which company would be more profitable? Why?
7. Elgar Toaster Co. is contemplating a modernization of its antiquated plant. It now sells its toasters for \$20 each; the variable cost per unit is \$8, and fixed costs are \$840,000 per year.
 - a. Calculate the break-even quantity.
 - b. If the proposed modernization is carried out, the new plant would have fixed costs of \$1,200,000 per year, but its variable costs would decrease to \$5 per unit.
 1. What will be the break-even point now?
 2. If the company wanted to break even at the same quantity as with the old plant, what price would it have to charge for a toaster?
 - c. If the new plant is built, the company would want to decrease its price to \$19 to improve its competitive position.
 1. At which quantity would profits of the old and the new plants be equal (assuming the price of a toaster is \$20 for the old plant but \$19 for the new)? How much would the profit be at this quantity?
 2. Calculate the degree of operating leverage for each plant at the quantity obtained in part (1).
 3. If sales are projected to reach 150,000 units per year in the near future, would you recommend construction of the new plant? Why or why not? (Assume that both plants have the capacity to produce this quantity.)
8. The Saline Company produces and sells rock salt. Its annual fixed cost was \$10,000. During the past year, the company sold 8,000 bags of its product. It estimates that at this level of sales its degree of operating leverage is 1.5.
 - a. How much was Saline's profit last year?
 - b. At which level of production would the company just break even?
9. The Amazing Book Co. sells a selection of paperbacks at an average price of \$9. Its fixed costs are \$400,000 per year and the unit variable cost of each paperback is \$4 on average.
 - a. Calculate the company's break-even quantity.
 - b. The company's sales target for the year is 100,000 units. What will be its profit?
 - c. At the beginning of the year the unit variable cost rises to \$5. If the company wants to achieve the same profit as obtained in item *b*, how many books will it have to sell?

- d. At the beginning of the year, the company installs new billing equipment. Its fixed costs rise to \$450,000. If the company wants to remain on target to sell 100,000 units and preserve the profit obtained in item *b*, what price will it have to charge per book?
 - e. An early review of competitive prices forces the company to drop its average price to \$8. If it still targets its sales at 100,000 units and will settle for a profit of \$50,000, what is the maximum unit variable cost it can afford? (TFC = \$400,000)
 - f. If the average price during the year is \$8 and the unit variable cost remains at \$4, how many books must the company sell to achieve the profit obtained in item *b*? (TFC = \$400,000)
10. Suppose you own an existing winery (with plenty of excess bottling and storage capacity) and you are considering introducing a new brand of wine in the super-value segment. You want to compete against Two-Buck-Chuck. (If you do not know about Charles Shaw, simply google “Two-Buck-Chuck” and read about the fastest growing wine brand in the history of the wine industry.) Your estimated AVC of production is constant at \$1.50 per bottle. You intend to introduce the wine at \$2.00 per bottle (to the wholesaler), but you realize that in order to sweeten the deal in the first year and to build brand loyalty, you will have to offer a \$0.40 per bottle discount during 4 of the first 12 months. Historically, new wine brands in this category have little consumer loyalty and face substantial competition—both factors lead to a high price elasticity of demand for new wine brands. You estimate the price elasticity of demand for your new wine brands to be $\epsilon_{\text{new brand}} = -5$. (Assume for simplicity that there is no seasonality involved in this product entry: one month of the year is as good as another.) You realize that advertising and promotion costs to launch a new brand should come to \$90,000 in the first year. How many bottles must you sell in order to break even?

Pricing and Output Decisions: Monopolistic Competition and Oligopoly

Learning Objectives

Upon completion of this chapter, readers should be able to:

- Cite the main differences between monopolistic competition and oligopoly.
- Describe the role that mutual interdependence plays in setting prices in oligopolistic markets.
- Illustrate price rigidity in oligopoly markets using the “kinked demand curve.”
- Elaborate on how nonprice factors help firms in monopolistic competition and oligopoly to differentiate their products and services.
- Cite and briefly describe the five forces in Porter’s model of competition.

The Situation

In a meeting with Nicole Goodman, Frank Robinson explained the results of his price analysis. “My concern is that we know our optimal price, but do we really know how our competitors are going to react when we launch our new energy drink, “Go-nuts” at this price point? Furthermore, I’m not sure the major players really consider us as a threat to their business, at least not yet. Therefore, can we assume they will take us seriously?”

Nicole agreed and suggested conducting further research. “One thing that we need is a complete list of prices of the major brands as well as the smaller brands.”

“Good idea,” said Frank. “I’ll get on this right away and have a report for you in a couple of weeks.”

INTRODUCTION

In this chapter, we examine in detail the pricing and output decisions made by managers in monopolistic competition and oligopoly. Economists also label these markets *imperfect competition*. This is in reference to their relative market power, which you will recall is the economist’s main criterion for determining the degree of competition. Perfect competition is, as its name signifies, “perfect” because firms have no market power whatsoever. Monopoly is not competitive because the single firm in the market has absolute market power. Monopolistic competition and oligopoly are considered

“imperfect” competition because firms in these markets have the power to set their prices within the limits of certain constraints. This power and these constraints are principal subjects of this chapter.

Table 9.1 provides an overview of the competitive environment by comparing the four market types on the basis of market power, nonprice competition, ease of market entry and exit, and the degree of interaction among competitors when making decisions (referred to as **mutual interdependence**). The key characteristic that separates perfect competition from the other three markets is market power. Firms in perfect competition are price takers, whereas firms in the other three markets are price makers. The key characteristic that separates perfect competition and monopoly from monopolistic competition and oligopoly is nonprice competition. Because product differentiation exists in the latter two markets, firms have reason to compete on the basis of other factors besides the price. Market entry and exit are easiest in perfect competition and relatively easy in monopolistic competition. It is not possible in monopoly and could be difficult in oligopoly, depending on how dominant the leading firms are in that market.

In our view, the most challenging aspect of the competitive environment is the presence of mutual interdependence. It is one thing to know how to use optimization techniques such as the $MR = MC$ rule to set the best price and output levels for the firm. It is entirely another matter when one has to anticipate and judge the possible reaction by competitors to one’s decision. The presence of mutual interdependence is the one factor that separates oligopoly from the other markets.

We begin this chapter with monopolistic competition and show how the $MR = MC$ rule that was developed in Chapter 8 can be used for firms in this industry. We then proceed to oligopoly. Because of its daunting competitive environment, most of this chapter is devoted to a discussion of the pricing, output, and nonpricing decisions that managers in this type of market must make. The topic of game theory is put off until Chapter 11. Following the discussion of imperfectly competitive markets, we provide a brief introduction to business strategy. We believe that this is a logical extension of the theoretical analysis of the behavior of firms in imperfectly competitive markets. As you will see, the decisions pertaining to a firm’s strategy are particularly important for managers when firms are price makers and are faced with tough price and nonprice competition.

Table 9.1 Overview of Competitive Environment, Four Market Types

	Perfect Competition	Monopoly	Monopolistic Competition	Oligopoly
Market power?	No	Yes, subject to government regulation	Yes	Yes
Mutual interdependence among competing firms?	No	No	No	Yes
Nonprice competition?	No	Optional	Yes	Yes
Easy market entry or exit?	Yes	No	Yes, relatively easy	No, relatively difficult

MONOPOLISTIC COMPETITION

Monopolistic competition is a market in which there are many firms and relatively easy entry. These two characteristics are very similar to those of perfect competition. What enables firms to set their prices (i.e., to be monopolistic) is product differentiation. By somehow convincing their customers that what they are selling is not the same as the offerings of other firms in the market, a monopolistic competitor is able to set its price at a level that is higher than the price established by the forces of supply and demand under conditions of perfect competition.

A good example of how product differentiation can turn a product sold in a perfectly competitive market into one in which a seller is able to exercise some degree of market power is the case of the poultry industry. Those readers who live along the East Coast will recognize the Perdue brand of chicken. For a number of years, Frank Perdue, owner of the company that bears his name, appeared on television telling people why his chickens were better. (His son, who now heads the company, appears in television commercials.) Of course, Perdue chickens were priced higher than the unbranded poultry products sold in the typical supermarket. But in monopolistically competitive markets, entry by newcomers is fairly easy. Soon after Perdue became a well-known brand, along came Bell and Evans as well as Murray chickens. At first, these newcomers could be found primarily in butcher shops and specialty or gourmet food stores. But soon, many supermarkets began to carry these brands.

Another company that was noted for its ability to differentiate a commodity product such as chicken was Tyson. In the mid-1990s, Tyson was applauded for its ability to differentiate its product and maintain stable premium prices by such actions as pricing by the hen rather than by the pound, and by selling breaded chicken parts.¹ But in recent years, Tyson has not been profitable, despite its success in differentiating its products. One reason is that it has not been very successful in managing its costs. To be sure, some of these costs, such as high fuel prices and wholesale poultry prices, are not under Tyson's direct control. In response to this, Tyson has been increasing its efforts to reduce costs as well as to continue to innovate with poultry products.²

In economic analysis, we assume the monopolistically competitive firm follows the $MR = MC$ rule to maximize its profit (or minimize its loss). Therefore, as a matter of convenience, we employ the same graphical illustration for monopolistic competition that we used for the case of a monopoly market, with one exception, demand is now labeled using a lowercase d with subscript i , d_i , to signify that this is individual firm demand as perceived by a representative firm in the market. This is shown in Figure 9.1.

Two short-run possibilities are shown in Figure 9.1*a* and 9.1*b*. In both instances, there is an incentive for change in the market because of nonzero profits. The positive profits in Figure 9.1*a* attract newcomers to the market, and the negative profits in Figure 9.1*b* suggest that some existing firms will exit the market. Entry will cause demand, as perceived by each firm, to contract, and exit will cause demand, as perceived by each remaining firm, to expand. Exactly how it expands or contracts depends on how the entrant enters this differentiated product market. For example, entry of another Chinese restaurant in a community may have a much larger effect on other Chinese restaurants than the entry of the first Mongolian barbeque restaurant. For

¹See Gene Walden and Edmund Lawler, *Marketing Masters: Secrets of America's Best Companies*, New York: HarperBusiness, 1993, pp. 12–17.

²Tyson CFO: Challenging Times Will Ultimately Benefit Company; Tyson Foods Focused on Creating More Demand Through Innovation, found in: [http://www.prnewswire.com/cgi-bin/stories.pl?ACCT=104&STORY=/www/story/09-20-2006/0004437015&EDATE=.](http://www.prnewswire.com/cgi-bin/stories.pl?ACCT=104&STORY=/www/story/09-20-2006/0004437015&EDATE=)

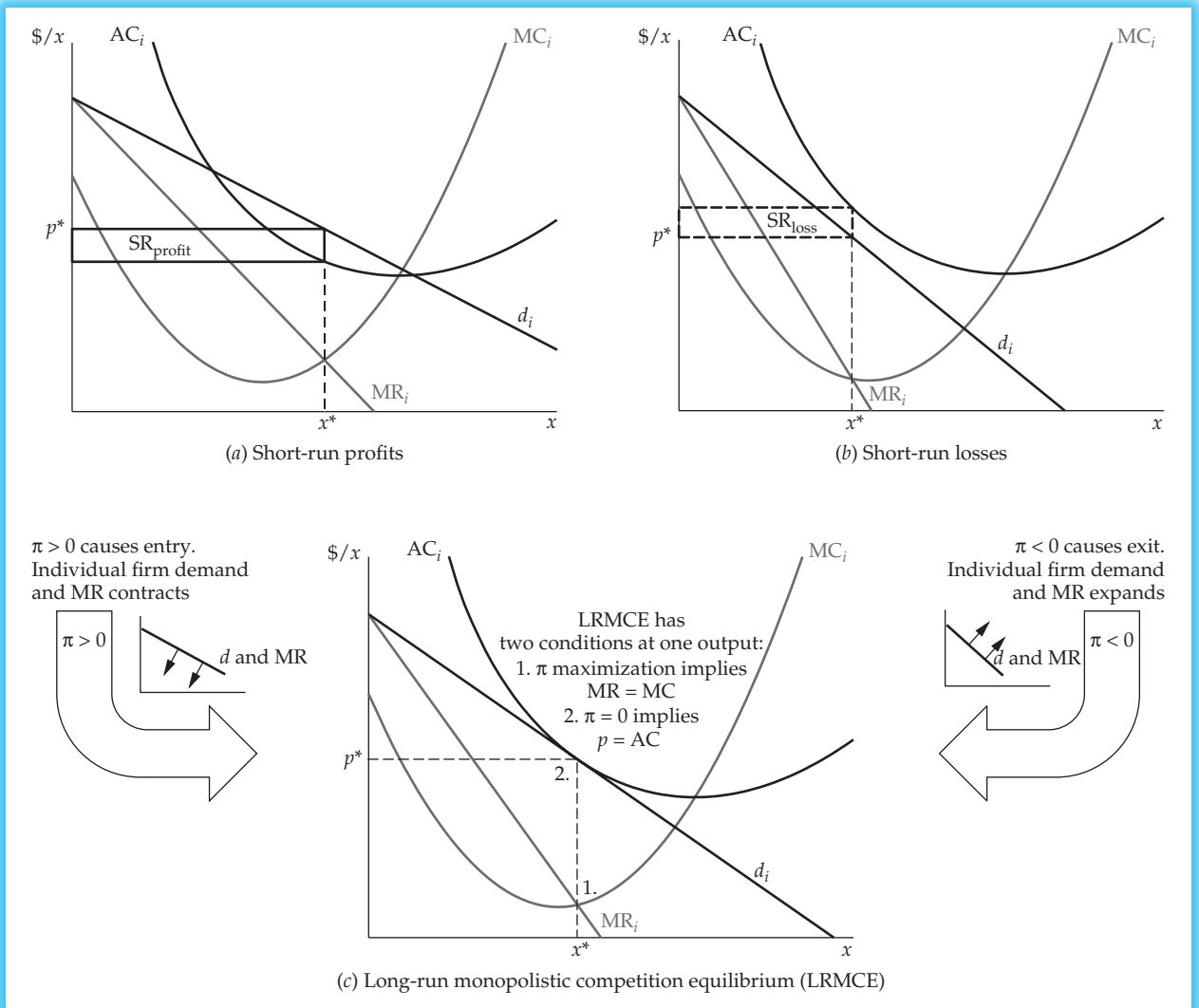


Figure 9.1 Monopolistic Competition. The view of a representative firm in a differentiated product market with free entry and exit

our purposes, we do not need to concern ourselves with exactly how entry affects existing firm demand; the direction of change for the representative firm is sufficient to describe the workings of the monopolistically competitive model. Given easy entry and exit, this process continues until the zero profits result. Entry or exit continues until the representative firm's profits are maximized ($MR = MC$) by having zero profits ($p = AC$). This is the situation shown in Figure 9.1c. Economists call this long-run monopolistic competition equilibrium (LRMCE). Appendix 9A provides an algebraic analysis of LRMCE.

Students are often confused by the lack of a market demand curve in the monopolistically competitive model. Part of the problem is that economists have difficulty in drawing such curves in differentiated product markets. How does one describe the demand for hamburgers when such meals are sold at drive-through fast-food restaurants as well as much more upscale, sit-down establishments? In another example, if an automobile dealer sells two cars, one a Kia and the other a Mercedes, does the dealer treat the two sales equally? If we ignore for a moment these difficulties, we can describe why

we expect demand as perceived by the individual firm to contract with entry and to expand with exit. Start with ten firms in the market and suppose each is the same size: Each has one-tenth of market demand. If a firm enters and becomes equally successful at attracting clients, then each firm will control one-eleventh of market demand. Demand rotates inward. Conversely, if a firm exits and the remaining firms equally split that firm's demand, then each firm will control one-ninth of market demand. Demand rotates outward. Of course, differentiated product entry is different because the successful entrant is likely the one that realizes that new customers can be brought to the market by entering in a niche of the market that is not as well served by existing firms. The new pizza shop opens and offers goat cheese and pine nut pizza in addition to the standard fare. The entry of such a firm may not cause existing firm demand to decline by as much because market demand (number of pizzas sold) may increase: A market niche that was previously unsatisfied is now being satisfied.

Some of the best examples of monopolistic competition can be found in the retail trade or in services. Restaurants, grocery stores, dry cleaners, stationery stores, florists, hardware stores, pharmacies, and video rental stores are all markets in which entry is fairly easy and the number of sellers is relatively large. In these businesses, owner-managers may use location, "service with a smile," or a slightly different mix of product offerings to differentiate their business.

In the United States, as well as throughout the world, Japanese sushi has become very popular.³ In recent years, we have noted a number of Chinese restaurants in New York that have started serving sushi in response to this increasing demand. This increase also points out another important feature of monopolistic competition: relatively easy entry into the market.

In the never-ending effort to keep new and incumbent competitors at bay, a firm in monopolistic competition must constantly strive to maintain product differentiation. An article in the *New York Times* goes well beyond the example of Chinese restaurants selling sushi. In "Borscht, Please, with a Side of Sushi," reporter Sarah Kershaw talks about "Russian sushi, Yemeni fried chicken, Taiwanese fish and chips, Japanese spaghetti and Jamaican jerk chicken with chop suey."⁴ In addition, the author has discovered Indian-Chinese restaurants in Queens. These restaurants are owned by Chinese who used to live in India and who are supposed to serve Chinese food exactly as it is prepared in India. Want more? Then try Pakistani-Italian and Norwegian-Cantonese in Brooklyn and Korean soul food in Harlem. There are even restaurants that serve Irish-Dominican, Greek-Irish, and Dominican-Italian food. We could go on and on, but the point of the article from our perspective is what immigrant restaurateurs told Ms. Kershaw: "In the highly competitive world of New York City ethnic food, it pays to branch out, mix, borrow and blend." Differentiating their products is the key to success for these firms.

OLIGOPOLY

Market Concentration

Oligopoly is a market dominated by a relatively small number of large firms. The products they sell may be either standardized or differentiated. Part of the control that firms in oligopoly markets exercise over price and output stems from their ability to differentiate their products. But market power also comes from their sheer size and

³See Chapter 8, "Global Applications," p. 335.

⁴"Borscht, Please, with a Side of Sushi," *New York Times*, December 21, 2001. One of the authors of this text lives in New York, another lives in Phoenix. We do not want to slight the good people of Phoenix so we should say that we know at least one ethnic-fusion restaurant in Phoenix. It is called "Chino-Bandito," and we allow readers to guess what combination of ethnic foods is offered in this establishment.

Table 9.2 Share of Value of Shipments Accounted for by the Four, Eight, and Twenty Largest Companies in the Three-, Four-, Five-, and Six-Digit Levels of Beverage and Tobacco Product Manufacturing and Sugar and Confectionary Product Manufacturing

NAICS Code	Industry Group and Industry	Companies	Value of Shipments (\$ millions)	Percent of Value of Shipments			HH Index
				4 Largest	8 Largest	20 Largest	
311	Food manufacturing	21,958	421,737	14.3	22.0	34.8	91.0
3113	Sugar and confectionary product mfg.	1,556	24,114	41.9	56.4	75.0	580.3
31131	Sugar mfg.	49	7,399	64.4	84.0	95.9	1,452.2
311311	Sugarcane mills	34	1,457	56.6	71.4	94.3	1,158.7
311312	Cane sugar refining	12	3,209	98.7	99.9	100.0	D
311313	Beet sugar refining	8	2,732	85.0	100.0	N	1,997.6
312	Beverage and tobacco product mfg.	2,237	96,971	45.1	59.1	72.7	777.2
3121	Beverage mfg.	2,169	60,896	40.9	52.1	66.0	531.5
31211	Soft drink and ice mfg.	1,008	32,587	45.5	53.6	68.8	743.3
312111	Soft drink mfg.	388	31,376	47.2	55.6	70.9	800.4
312112	Bottled water mfg.	109	785	51.7	64.4	80.3	986.6
312113	Ice mfg.	514	424	24.4	31.3	43.1	302.3

N: Does not apply

D: Data omitted because of possible disclosure.

Source: U.S. Census of Manufactures, 1997. (Note: The above data from the 2002 Census are not readily available in the format shown in this table. Thus, we continue to use the 1997 data in this edition for purposes of illustration.)

market dominance. There are two good sources of information on the degree of concentration that exists in various markets. The Census Bureau's Survey of Manufactures provides a comprehensive account of market share for the major industrial sectors of the U.S. economy.⁵ However, it does not always provide the most current data (the most recent survey was conducted in 2002). For more current data, as well as market share information on specific brands, we recommend the *Market Share Reporter*.⁶

Table 9.2 presents sample information found in the Census Survey. The Census Bureau classifies industries by level of product specificity. The information that we found in the 1997 report begins at the three-digit level (broadest) and goes to the six-digit level (most specific). In 1997, the Census Bureau introduced a new system for classifying companies called the North American Industry Classification System, or NAICS. Previous to that, the Census Bureau had used the Standard Industrial Classification, or SIC. The categories in the two systems are only about 50 percent compatible with each other, so time comparisons on shifting market shares are difficult to make. Notice that as the product classification becomes more specific, the degree of market concentration among sellers tends to increase. This pattern occurs in many other industries but not in any consistent or precise manner.

⁵For a PDF file of the latest survey, go to www.census.gov.

⁶Robert S. Lazich, ed. *Market Share Reporter*, Detroit: The Gale Press, 2012.

Besides indicating the degree of market concentration by measuring the market share of the top four, eight, twenty, and fifty largest companies, the Census Survey provides a measure of concentration called the Herfindahl-Hirschman (HH) index. The formula for this index is as follows:

$$HH = \sum_{i=1}^n S_i^2$$

where n is the number of companies in the industry, and S is the i th company's market share. The HH index is shown for the fifteen industries in the last column of Table 9.2. The advantages of the HH index relative to concentration ratios are that

1. It uses the market share information about all the firms.
2. The squaring of individual market shares gives more weight to the larger firms.

The maximum HH index is 10,000—when there is just one firm in the industry. Using the prior example, the HH index will differentiate between an industry where four firms equally share the total market ($HH = 2,500$), and an industry where the top firm has 94 percent and the other three 2 percent each ($HH = 8,848$). According to U.S. Department of Justice Merger Guidelines of 1982, “unconcentrated” markets are defined as those with an HH of less than 1,000. Table 9.2 shows selected markets relating to the soft drink industry.

PRICING IN AN OLIGOPOLISTIC MARKET: RIVALRY AND MUTUAL INTERDEPENDENCE

Whether the sellers in an oligopolistic market compete against each other by differentiating their product, dominating market share, or both, the fact that there are relatively few sellers creates a situation where each is carefully watching the other as it sets its price. Economists refer to this pricing behavior as mutual interdependence. This means that each seller is setting its price while explicitly considering the reaction by its competitors to the price that it establishes.

In the 1930s, economist Paul Sweezy provided an early insight into the pricing dynamics of mutual interdependence among oligopoly firms by developing a **kinked demand** curve model.⁷ The basic assumption of the Sweezy model is that a competitor (or competitors) will follow a price decrease but will not make a change in reaction to a price increase. Thus the firm contemplating a price change may refrain from doing so for fear that quantities sold will be affected in such a way as to decrease profits.

If a firm lowers its price, this may have an immediate impact on the competition. This firm takes its action to increase sales by drawing customers away from the higher-priced competitors, but when competitors realize what is happening (i.e., their sales are declining), they will quickly follow the price cut to maintain their market share. If this firm undertakes the opposite action—a price increase—incorrectly assuming competitors will follow suit, its sales will drop markedly if competitors fail to do so.

It is easy to demonstrate the “kink” in such a demand curve with the graph in Figure 9.2. Let us assume the original price and quantity are found at point A . If the firm lowers its price, expecting that quantity demanded will move along the more elastic demand curve D_1 and this result materializes, then it will gain a relatively large quantity of additional sales for a relatively small decrease in price. If it lowers its price

⁷Paul Sweezy, “Demand Under Conditions of Oligopoly,” *Journal of Political Economy*, 47 (1939), pp. 568–73.

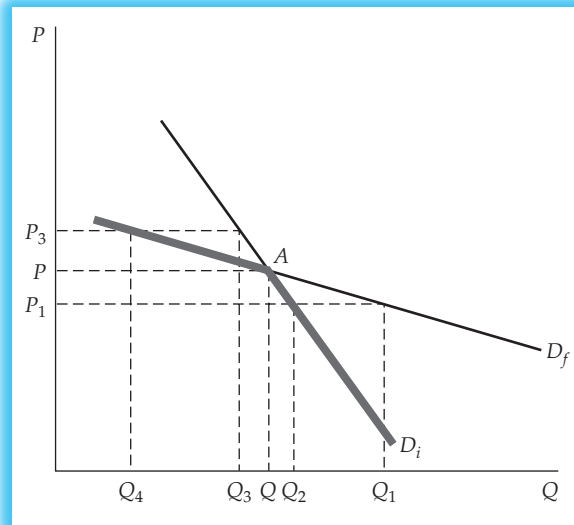


Figure 9.2 Demand Curves for an Oligopoly Considering a Price Increase/Decrease

from P to P_1 , it will expect to increase its sales from Q to Q_1 . This is the relevant demand curve for the firm if other companies do not retaliate. Our firm would thus gain customers at the expense of competition. However, if competitors do react and match the price cut, our company will increase its sales only to Q_2 , along demand curve D_i ; this is the relevant demand curve when all companies in the industry decrease their price equally. There will be a relatively small increase in sales because all prices in the industry are lower, but not nearly as much as the company expected when it reduced its price.

In contrast, suppose our company decides to raise its price, anticipating that competitors will follow the increase. It thus expects to move along D_i to Q_3 when it boosts its price to P_3 . It would thus sustain some loss in sales while benefiting from a significantly higher price. However, suppose its competitors refuse to play along and keep their prices unchanged. The company's situation now becomes more precarious because its quantity sold drops to Q_4 : The demand curve for the firm alone is much more elastic than if all firms raise their prices in unison.

The prospect of being stung by such action will make the company much more loath to change its price from P . From that vantage point, it will appear to the company that the appropriate demand curve is D_i if price is lowered and D_f if the price is increased. The upper portion of D_f and the lower portion of D_i can be seen to form a kinked demand curve around point A ; thus, the name of Sweezy's model. These relevant portions of the two demand curves are boldly outlined in Figure 9.2.

Now that we have developed a demand curve for this oligopolist, we can also derive a marginal revenue curve. This marginal revenue curve will be discontinuous: there will be a gap at the point where the kink occurs. As we know, a company will maximize its profits at the point where marginal cost equals marginal revenue. The two marginal cost curves drawn in Figure 9.3 both imply the same price and quantity at point A . Thus a significant change in costs could occur for our firm, but it will not react by changing its price. Actually, the price may remain unchanged even if the demand curve moves to the right or left, as long as the kink remains at the same price level. Hence, it can be concluded that under the circumstances described, a kinked demand curve will result in price rigidities despite changes in demand and cost.

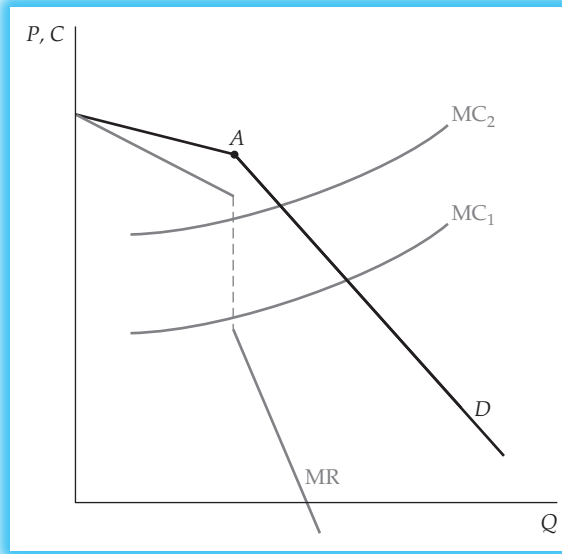


Figure 9.3 The Kinked Demand Curve

Over the years, the kinked demand curve has been challenged by other economists. In particular, Nobel Prize laureate George Stigler investigated several oligopolistic industries and found little empirical support for Sweezy's model. Stigler found that in these industries, price increases were followed as quickly as were price decreases.⁸ Such findings, of course, contradict the existence of the kink. Further, the model does not explain how the price was originally set at the kink. Was it originally set where marginal revenue equaled marginal cost, or was it by some other means, such as tradition?

One commonly held view of how the price at point A is determined involves the concept of an industry **price leader**. This is the firm that dares to break out of the pack without fearing the consequences spelled out in the kinked demand model. If this firm decides to raise its price, it assumes all others will follow. If this firm decides to lower its price, it assumes the others may follow but will not go even lower, thereby triggering a price war that would hurt the entire industry.

In oligopoly markets, the role of the price leader is usually assumed by the company with the largest share of the market. One of the most notable leaders in the past was IBM, when it dominated the mainframe computer market from the 1960s to the mid 1980s. Whatever price IBM set for its computers, the much smaller competitors would usually set their price at about 10 percent lower than IBM's umbrella price. A survey of price leadership can be found in a study by the Strategic Pricing Management Group. Among the price leaders noted in this survey is Procter & Gamble.⁹

COMPETING IN IMPERFECTLY COMPETITIVE MARKETS

Nonprice Competition

The key to the pricing power of firms in monopolistic competition and oligopoly is their ability to differentiate their product so they are not mere price takers who are subject to the tyranny of supply and demand. All efforts to do so are referred

⁸George J. Stigler, "The Kinky Oligopoly Demand Curve and Rigid Prices," *Journal of Political Economy*, October 1947, pp. 432–39. Similarly, see Julian J. Simon, "A Further Test of the Kinky Oligopoly Demand Curve," *American Economic Review*, December 1969, pp. 971–75.

⁹Michael Hurwich, "Price Leader or Price Follower: Where Do You Stand?" <http://www.spmgpricing.com/PDF/Price%20Leadership%20or%20Followship.pdf>

to in economics as *nonprice competition*.¹⁰ We have already provided examples of nonprice competition in our earlier discussion of monopolistic competition. But we now present this concept in a more formal and systematic manner.

A simple, but useful definition of nonprice competition is “any effort made by firms other than a change in the price of the ‘product in question’¹¹ in order to influence the demand for their product.” More specifically, these efforts are intended to influence the nonprice determinants of demand. Here you will see the nonprice determinants of demand (first introduced in Chapter 3) and a suggested list of nonprice variables that firms might choose to use to compete in monopolistic competition or oligopoly. *Nonprice determinants of demand* include any factor that causes the demand curve to shift. They are (1) tastes and preferences, (2) income, (3) prices of substitutes and complements, (4) number of buyers, and (5) future expectations of buyers about product price. (Recall that changes in the price of the product causes movements along the demand curve.) *Nonprice variables* are those factors that managers can control, influence, or explicitly consider in making decisions affecting the demand for their goods and services. These variables are (1) advertising, (2) promotion, (3) location and distribution channels, (4) market segmentation, (5) loyalty programs, (6) product extensions and new product development, (7) special customer services, (8) product “lock-in” or “tie-in,” and (9) preemptive new product announcements.

Readers might recognize the nonprice variables as important tools of marketing. In a typical corporation, the management of these variables (sometimes called the *marketing mix*) is largely the responsibility of the marketing department. A complete discussion of each variable is well beyond the scope of this text and in fact would be found in any marketing text. But we comment briefly on some of these variables in terms of their expected impact on the nonprice determinants of demand:

1. **Tastes and preferences:** *Advertising and promotion* are intended to influence tastes and preferences.¹² Very often, advertising and promotional campaigns are tied into the building or support of a brand. A common means of enhancing these campaigns is the use of celebrities as representatives of a brand. Brand advertising and promotion are often combined with *market segmentation*. In this case, rather than try to influence tastes and preferences, firms try to target different market segments in which they believe their products will have the greatest appeal. In recent years, efforts to retain customers over the long term have become as important to firms as those designed to win new customers. Thus, *loyalty programs* such as points given for using a credit card, for mileage traveled on an airline, or for staying at a hotel have been an important means of influencing tastes and preferences.
2. **Income:** Firms cannot affect income. But *market segmentation* is a way for firms to focus on the income levels that they believe will be most likely to buy their products. If they are selling a “superior” or luxury product, then they would try to target higher income groups. The opposite would hold if they believed their products were “inferior.” The main shopping street in Zurich is Bahnhofstrasse. This street is considered to be the most expensive street for retail shopping in Europe and the third most expensive

¹⁰This somewhat awkward term is an indication of the heavy emphasis in economic analysis of pricing or market power as the basis for competition. In the ordinary language of business, the term to describe the factors discussed in this section would be considered just a part of “competition.” Therefore, outside the confines of economic analysis, a “highly competitive market” is assumed to mean price or nonprice competition or some combination of both.

¹¹The use of the term *product in question* is to remind readers about the difference that we made in Chapter 4 between *own-price* and *cross-price* elasticity. Companies may use the price of complements or substitutes to try to stimulate the demand for the product in question. But this is considered to be a form of nonprice competition.

¹²One must be cautious about the precise nature of advertising’s impact on tastes and preferences. As John Wanamaker, founder of the famous Philadelphia department store, once said, “I believe that half of my advertising is effective, the problem is that I don’t know which half.”

in the world. On a business trip to Zurich in 2012, one of the authors saw that the jewelry stores carrying the most well-known (and most expensive) brands of luxury Swiss watches had discreet, descriptive signs for some of the watches in two languages. Can you guess what these two languages were?¹³

3. **Prices of substitutes and complements:** Efforts to *lock-in* customers will tend to reduce the effect that changes in the prices of substitutes will have on the products that firms sell. This effort may be fairly aggressive, such as wireless phone companies offering discounts or free mobile phones for long-term contracts but levying a penalty if customers want to get out before the expiration date. In contrast, this effort may be rather subtle. In effect, Microsoft locks in the use of its PC applications such as Word or PowerPoint through the “network effect.”
4. **Number of buyers:** This impact of this factor is self-evident. This is precisely why so many companies are working so hard to increase their market presence in the growth markets in Brazil, Russia, India, and China (the BRIC countries). Their large populations (and their rising middle class) offer huge potential demand for goods and services of all kinds. BRIC countries are discussed in Chapter 13.
5. **Future expectations:** Very often, when a new consumer electronics product is launched, sellers price them at a relatively high level, a practice called price skimming (see Chapter 10). Shortly afterward, prices begin to drop as market supply increases. If there are enough consumers who assume this will always happen, the demand for the product at its initial launch stage may actually fall because consumers with future expectations that prices will, in fact, fall hold off on buying the product.
6. **Financing terms and conditions:** In both consumer and industry markets, sellers can use financing terms and conditions as a way of influencing the demand for their product. This initiative does not quite fit into the set of nonprice determinants of demand used in economic theory. But in effect, the use of financing terms and conditions is like changing the price.

Perhaps you have seen advertisements or promotional literature sponsored by electronics stores, furniture stores, or department stores saying “zero interest” or “no monthly payments until next year.” This is a way of reducing the price of their products to those who buy on credit without actually lowering the price of the products themselves. Manufacturers of industrial equipment may use the financing arm of the company for this purpose. During the boom years of the telecommunications industry in the 1990s, there were a number of competitive local exchange carriers (CLECs) that sprouted up to compete against the incumbent Regional Bell Operating Companies (RBOCs). Much of the equipment sold to them by Lucent was done with the help of Lucent-backed financing. When the technology sector crumbled, a number of these CLECs went bankrupt or failed to meet their obligations to Lucent, thereby causing the company to suffer heavy losses. In 2006, Lucent was bought by Alcatel, the French telecommunications equipment manufacturer.

The Reality of Monopolistic Competition and Oligopoly: “Imperfect” Competition

In economic theory, monopolistic competition consists of a large number of relatively small firms that are subject to competitive pressures because of the ease with which newcomers can enter the market and set their own prices. Oligopoly consists of a small number of relatively large firms whose size or power in differentiating their products makes it difficult for newcomers to enter the market. But their size and domination of the market also make it imperative for them to watch each other closely when setting their prices.

¹³Chinese and Russian.

However, the distinction between monopolistic competition and oligopoly can sometimes be blurred in actual markets. For example, mutual interdependence can sometimes be found in monopolistic competition, and there is a possibility for new firms to enter oligopoly markets. Consider the following three examples:

1. **The American automobile industry:** Look at any economics textbook from the 1970s and you will probably find the American automobile market as a prime example of an oligopoly. But today this market has taken on certain key characteristics of monopolistic competition. The increasing dominance of Japanese, German, and, more recently, Korean automakers in the U.S. market make it hard to argue that entry is difficult in this market. In fact, analysts believe that Chinese-made and branded cars will eventually enter the American market. At its peak in the early 1970s, General Motors held more than 50 percent market share. Today it is below 25 percent.
2. **Small retail establishments around the world:** A frequently used example of monopolistic competition is the retail industry, particularly those dominated by small mom and pop establishments. But the competitive landscape of the retail business has changed dramatically in the United States over the past several decades as large big-box chains, led of course by Wal-Mart, have come to dominate the selling of products and services to consumers. These firms (other examples include Home Depot, Sports Authority, and Best Buy) look much more like the giant steel and auto companies originally envisioned by economic theorists as examples of oligopoly. These giant retailers make it increasingly difficult for the traditional mom and pop store to survive. Moreover, the family pharmacist who owns the corner drugstore has also become a relic of the past as chains such as CVS, Rite Aid, and Walgreens increase their presence and domination of this particular retail business.
3. **Global issuers of credit cards: *monopolistic competition or oligopoly?*** On the surface, the credit card industry appears to be an oligopoly dominated by Visa, MasterCard, American Express, and Discover. But a look below the surface reveals that Visa and MasterCard are really associations of financial institutions that issue cards and “acquire” merchants who agree to accept the cards as a means of payment. Visa and MasterCard each have thousands of members, most of whom are banks. This would indicate monopolistic competition. But a still closer look indicates that the card-issuing industry has become increasingly more concentrated. On the basis of card holder receivables (i.e., the amount owed by card holders to issuing financial institutions), the top ten account for about 90 percent market share.

As a way of dealing with the possible blurring of these two types of markets, monopolistic competition and oligopoly are sometimes put together into one category called imperfect competition. We introduced this term at the beginning of this chapter. But let us provide a formal definition of this term just to reinforce our points. Imperfect competition is “a market in which firms are able to exert varying degrees of market power because of their size and/or their ability to differentiate their products from those sold by their competitors.”

STRATEGY: THE FUNDAMENTAL CHALLENGE FOR FIRMS IN IMPERFECT COMPETITION

Since the mid-1990s, textbooks in managerial economics have gradually expanded their coverage of strategy. This topic is a principal component of a business curriculum, and entire courses are devoted to its study. There are some important linkages between managerial economics and strategy. Furthermore, within the context of managerial decision making, we believe that no discussion would be complete without some mention of the strategic challenges facing a firm. As the title of this section implies, the decisions pertaining to a firm’s strategy are especially critical for managers of firms operating in imperfect competition. Strategic decisions are far less important

for perfectly competitive or monopoly firms. The former are price takers, and their strategic decisions are confined primarily to whether they want to remain in a market and how best to survive during the downturns of a market. The latter have no competition, so a strategy is not as critical. Strategy becomes a critical part of the manager's responsibility, primarily when firms are price makers and are faced with tough price and nonprice competition, as well as the threat of new entrants into the market. Such is the case with firms operating in imperfect competition (i.e., monopolistic competition and oligopoly). This is our reason for placing this brief discussion of strategy in this particular chapter of the text.

There are many linkages between managerial economics and strategy because a number of the concepts and tools of analysis used in the making of strategy are rooted in economics.¹⁴ In fact, the two fields of study can be said to be linked almost by definition. In Chapter 1, we define managerial economics as “the use of economic analysis to make business decisions involving the best use of an organization's scarce resources.”¹⁵ Textbooks abound with a number of definitions of **strategy**.¹⁶ Here is a composite definition of strategy that we can provide using selected sources: “Strategy is the means by which an organization uses its scarce resources to relate to the competitive environment in a manner that is expected to achieve superior business performance over the long run.”

The similarities and differences between managerial economics and strategy can perhaps be best seen by comparing the basic questions addressed in the two fields of study. These are the types of questions firms must consider in the study of managerial economics:

1. If we operate in perfect competition, what output level should we set for our product?
2. If we are a monopoly or imperfect competitor, what price (which in turn determines our output level) should we set for our product?
3. If we compete in an oligopoly, what might the reaction of our competitors be to decision 2?
4. Are we prepared to compete as a low-cost producer in the long run as new firms enter the market?
5. If we are not breaking even, should we shut down our operations in the short run? In the long run?
6. How do we ensure our short-run profit is maximized? If we are losing money, how do we ensure our losses are minimized?
7. If we are operating in monopolistic competition or oligopoly, how can we differentiate our products so that we can command a price that is higher than the rest of the competition (unless, like Walmart, our strategy is to compete on the basis of price)?

These are the types of questions that firms must consider in the determination of their strategy:

1. What businesses should we be in? (corporate-level decision)
2. How should we compete in these businesses: product differentiator or cost leader? (group, division, or product-level decision)
3. What are our long-run strategic objectives? How do these relate to our short-run tactics?

¹⁴Two excellent texts on strategy written by economists and intended for strategy courses with an economic slant are Sharon M. Oster, *Modern Competitive Analysis*, 3rd ed., New York: Oxford University Press, 1999; and David Besank, David Dranova, and Mark Shanley, *The Economics of Strategy*, 3rd ed., New York: John Wiley & Sons, 2003. The latter book has been cited in various other sections of this text and is also used as a reference for this section.

¹⁵Refer to Chapter 1.

¹⁶See Besank et al., p. 1, for a list of three definitions of strategy provided by leading academics that are very similar to the one we offer. Also, the introductory section of this book “Primer: Economic Concepts for Strategy,” pp. 9–40, provides an excellent survey of the economic foundations of strategy.

4. What geographic segments of the market should we focus on? (local, regional, national, and international)
5. What demographic segments of the market should we focus on?¹⁷
6. What will be the reaction of our competitors to our decisions? Is there a particular advantage to being a first mover in the market?
7. What are our core competencies and how can we use this to our competitive advantage?
8. In making our strategic decisions, how can we best incorporate the changes that are taking place in the total business environment, which includes suppliers, buyers, potential entrants, substitute products in completely different markets, technology, and macroeconomic factors such as the national business cycle and levels of growth and development around the world?

There are more strategic questions that we could add, but this should be sufficient to show the similarities and differences between the study of managerial economics and the study of strategy.

Our coverage of the linkage between managerial economics and strategy is divided into two sections. The first section deals with industrial organization, a branch of economic study that underlies much of the original work on strategy done by Professor Michael Porter. The second section introduces Porter's basic ideas on strategy.

Industrial Organization

Industrial organization is the branch of economics that studies ways that firms and markets are organized and how this organization affects the economy from the viewpoint of social welfare (i.e., maximizing the well-being of consumers and producers). An important question that affects this discipline directly relates to managerial economics and the study of the behavior of oligopoly firms. *How, if at all, does industry concentration affect the behavior of firms competing in this industry?*¹⁸ Two approaches to the study of industrial organization and to this question in particular have dominated the field. The first, the **structure-conduct-performance (S-C-P) paradigm**, has been predominant since the 1940s until its critics, proponents of the price theory approach, began to challenge it in the 1970s. As so often is the case, the two explanations of industrial behavior were advanced by two of the leading schools of economic thought in the United States, Harvard University and the University of Chicago, respectively. A brief coverage of the two analytical methods follows.¹⁹

The Structure-Conduct-Performance Paradigm

The causality in this theory runs only in one direction. An industry *structure* affects its *conduct*, which in turn affects its *performance*. Hence the label “structure-conduct-performance.” We start with structure that is shaped by the demand and supply conditions prevalent in the industry. For instance, if an industry's product demand tends to be inelastic, then market prices would be higher than if demand were more elastic. Growth patterns and substitutability would also affect structure. On the supply side, technology is an important factor. The existence of economies of scale will determine the number of firms that can operate profitably in the industry.

¹⁷Demographics refers to such well-known characteristics as age, gender, income or education level, and ethnic background or origin. Marketing strategists also use the term *psychographics*, which segments consumers by values and lifestyles.

¹⁸Industry organization specialists who believe that industry concentration does in fact affect the behavior of firms proceed to explore the social welfare implications of this behavior. This part of their work is not within the scope of this text.

¹⁹The following discussion is based on two texts on industrial organization: Dennis W. Carlton and Jeffrey M. Perloff, *Modern Industrial Organization*, 2d ed., New York: HarperCollins, 1994; and William E. Shugart II, *The Organization of Industry*, Homewood, IL: Irwin, 1990.

These basic conditions thus determine industry structure: the number of firms in an industry, conditions of entry, and product differentiation. The structure of the industry then directly influences the way the industry operates—its conduct. Conduct entails primarily pricing strategies and other activities such as advertising, product development, legal tactics, and choice of product, as well as the potential for collusion among companies and mergers, which may further endanger the competitive nature of the industry. “The essence of the structuralist approach is a presumption that industries having fewer (and larger) firms will tend to engage in conduct inconsistent with the norms of perfect competition.”²⁰

The next step in the S-C-P model is to link conduct to performance. As mentioned, the usual normative standard adopted by economists is the maximization of society’s welfare. An industry market with great concentration will fall far short of reaching such a goal. Its performance will be marked by both productive and allocative inefficiencies. Prices will be above marginal costs, the choice and quality of products will not be ideal, technological progress may be slowed down, and ultimately profits will be higher than under competitive circumstances. This high level of profitability arises from the industry’s pricing policy and not because of any cost advantages.

We have now connected the three parts of the S-C-P model. A necessary corollary of the results of this approach is that high industry concentration becomes a reason for government intervention proceeding against possible mergers and even trying to deconcentrate industries.

The “New” Theory of Industrial Organization

A competing view of industrial organization is offered by a theory stating that there is no necessary connection between industry structure and performance that uniquely leads to maximum social welfare. It argues that the study of industry organization should use and apply microeconomic price theory. Thus, for instance, higher profit levels in more concentrated industries may be caused by economic efficiencies rather than pricing strategies. In other words, a small number of efficient firms can survive better than a larger number of firms that cannot take advantage of economies of scale. In fact, certain industries may not be able to support many firms on an acceptable profit level. Another argument involves the potential for collusion in concentrated industries. George Stigler concluded that, although industry concentration makes collusion more likely, collusive agreements (cartels) are inherently unstable because they are expensive to enforce and because participants have incentives to cheat.²¹

A large number of studies examining the links between industry concentration and profits have been produced. The results, particularly in the earlier studies, appear to point to direct correlation between concentration and profit levels. However, the evidence is at best rather weak. Many later studies cast doubt on the earlier results, showing that what appeared to be a link between concentration and profits was actually due to other industry characteristics and often disappeared over the long run. For instance, one researcher found that cost reductions were significantly greater in increasingly concentrated industries, and even though price reductions to consumers were less than cost savings realized, consumers benefited from considerable price advantages.²²

²⁰Shugart, *The Organization of Industry*, p. 9.

²¹G. J. Stigler, “A Theory of Oligopoly,” *Journal of Political Economy*, 72 (February 1964), pp. 44–61. Another argument made by the S-C-P approach is the monopolistic implications caused by tie-in sales. This point is discussed in Chapter 10.

²²S. Peltzman, “The Gains and Losses from Industrial Concentration,” *Journal of Law and Economics*, 20 (October 1977), pp. 229–63.

Another, more recent addition to the previous discussion and the notion that industry profits are not necessarily a function of industry structure is Baumol's theory of "contestable markets."²³ The idea is that performance by firms is ultimately influenced not by the presence of competition, but by the threat of potential competition. Contestability comes in several forms. An extreme case is perfect contestability, which implies that entry by new firms is free and exit is costless. In such cases, existing firms will not be able to sell their product at prices exceeding marginal costs. Of course, in industries where entry and exit are difficult and costly, such competitive threat would not be present. Still, if outsiders have access to the industry's technology and if assets are nonspecialized so their costs can be recouped upon exit, such a competitive threat is powerful.

The debate between the advocates of the S-C-P paradigm and the price theory approach has now continued for many years and neither side has gained clear dominance. Market structure is still very important, particularly from a public policy point of view. But other factors must also be considered in deciding what are the important influences affecting social welfare.

Strategy and the Ideas of Michael Porter

In the late 1970s, Michael Porter, an economics professor from the Harvard Business School, began publishing articles on strategy that would have a profound impact on the study of this subject. As with all innovative ideas, his work has been subject to counterarguments and criticism over the years.²⁴ Nevertheless, many of the terms and concepts that he introduced have become a part of the everyday vocabulary of the business strategist. In our discussion of strategy, we chose to limit coverage to his ideas about strategy because they are most directly related to the study of the economics of the firm and industry economics.

Porter did not dwell on theoretical issues about the links among industry structure, business conduct, and performance and whether a concentrated industry structure leads to a misallocation of a country's resources. Instead, he used the concepts of the S-C-P approach to industry economics as the basis for understanding the strategic challenges facing the managers of firms as they seek to maximize their firm's profit. His **Five Forces model**, shown in Figure 9.4, illustrates the various factors that affect the ability of any firm in the industry to earn a profit.

If buyers and suppliers do not exert much market power and there is little threat from either new competitors or the use of substitute products, firms in the industry are likely to earn relatively high returns on their investment. An extreme example of this is the market for operating systems software for personal computers, as long as you think of Microsoft as *the* company in this industry. It is easy to see the influence of the microeconomic theory of the firm on Porter's ideas when you compare the five forces with the characteristics of the different market types shown previously in Figure 8.1. For example, "Potential Entrants" is related to "free entry and exit." If it is easier to enter, firms in the industry are less able to exercise market power and, hence, are less likely to have a relatively high return on investment.

After establishing a model for analyzing the overall profitability of an industry, Porter proceeded to discuss what kind of strategy would enable a specific firm in the

²³W. J. Baumol, "Contestable Markets: An Uprising in the Theory of Industry Structure," *American Economic Review*, 72 (March 1982), pp. 1–15.

²⁴For a good review of Porter's ideas on strategy together with the developments in the entire field of study during the past 30 years, see Henry Mintzberg, Bruce Ahlstrand, and Joseph Lampel, *Strategy Safari: A Guided Tour Through the Wilds of Strategic Management*, New York: The Free Press, 1998.

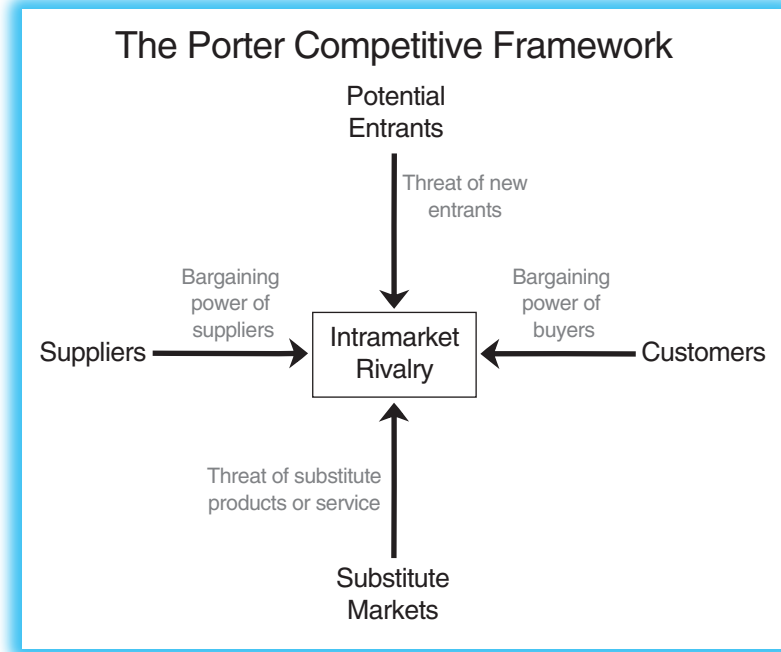


Figure 9.4 The Porter Competitive Framework

industry to earn a return higher than the industry average. This above-average return can be considered akin to “economic” or “above-normal” profit. Porter offers two generic strategies for earning an above-average return on investment. The first is the “differentiation” approach; the second is the “cost leadership” approach. There is actually a third approach, in which a firm could exercise either differentiation or cost leadership for a particular market segment rather than for the entire market of potential buyers.

We can also see the influence of the microeconomic theory of the firm in this aspect of Porter’s ideas on strategy. The cost leadership approach appears to be based on perfect competition, whereas the differentiation approach can be seen as related to the case of imperfectly competitive markets. Recall that in the perfectly competitive market, all firms are price takers. The only way in which a particular firm in this market can earn an economic profit is to keep its cost structure low enough so when $P = MC$, there is still a positive difference between P and AC . In our view, this is where Porter came up with the idea of the cost leadership approach.

In *all* market structures except perfect competition, the demand curve perceived by the firm is downward sloping. This implies that MR is always below demand and that MR and P are related according to the following elasticity rule.²⁵

$$MR = P \cdot \left(1 + \frac{1}{\epsilon} \right)$$

²⁵These assertions are easy to prove with calculus. $TR(Q) = P(Q) \cdot Q$ so $MR(Q) = dTR(Q)/dQ = P(Q) + Q \cdot dP(Q)/dQ$ by the product rule. Because $dP/dQ < 0$, $MR < P$ and MR is always below demand. Since $MR = P + Q \cdot dP/dQ = P + (P/P) \cdot Q \cdot dP/dQ = P \cdot (1 + (Q/P) \cdot dP/dQ) = P \cdot (1 + 1/\epsilon)$.

Because the firm sets $MR = MC$ to maximize profit, we see a ready relation that connects price and marginal cost in imperfectly competitive markets (after regrouping):

$$\frac{p - MC}{p} = \frac{-1}{\epsilon} = \frac{1}{|\epsilon|} \quad \text{Lerner Index of Monopoly Power}$$

The Lerner Index is named after Abba Lerner, who suggested its use as a measure of monopoly power in the 1930s. This is also often called the *inverse elasticity rule*.

Following the $MR = MC$ rule enables a firm to set a price on the demand line that is higher than its AC, assuming, of course, the demand curve itself is high enough to allow this to happen. The extreme case of this was shown in the example of pure monopoly. Not only is the product different, but it is also unique. In the case of monopolistic competition, the above-normal profit enjoyed by the differentiating firm may be reduced as the entry of firms in the long run shift a firm's demand curve to the left. But Porter would say that those firms that have a solid differentiation strategy would not be affected by the new entrants as much as those who allow their products to be "commoditized."

Concluding Remarks on the Linkages Between Managerial Economics and Strategy

The fundamental link between managerial economics and strategy is the decision regarding the allocation of a company's scarce resources. In this chapter, we also show how the tools of strategic analysis such as Porter's "Five Forces" model and his "differentiation versus cost leadership" approach are linked to the economic study of industrial organization and the economic models of a firm's behavior in different market settings. Interested readers can find a thorough discussion of these and many other examples of the linkages between economics and strategy in the textbooks on the subject previously cited.²⁶

GLOBAL APPLICATION: THE WORLD'S MARKET FOR BEER²⁷

The market for beer is a good example of how difficult it is to actually label a particular industry using any of the four basic types of economic markets. At the least, we can say that the market for beer is neither pure monopoly nor perfect competition. But to say it is monopolistic competition or oligopoly depends on whether we are looking at the market from the standpoint of a country, a region, a type (e.g., regular versus light, mass market versus craft-brewed, domestic versus imported), or a brand. To complicate the issue further, as you know, each beer company can produce or own and distribute many different brands. For example, a visit to the website of Anheuser Busch, the leading brewer in the United States, reveals that the company produces or distributes dozens of different brands including its top sellers Budweiser, Bud Light, and Michelob. A leading brewer of craft beer, the Boston Beer Company, produces many other types of beer in addition to its flagship Samuel Adams brand.

²⁶See Oster and Besank et al.

²⁷This section is based on the following sources: Anne Nugent, "The Global Beer Market: A World of Two Halves," *Euroonitor*, February 25, 2005; Julie Bradford, "Just About Everyone Agrees That the Beer Industry Is Having Problems, But the Troubles Are More Frustrating Than They Are Fatal," *Beverage Dynamics*, September 1, 2005; and "The Global Battle of the Bottled Beer," *Grocer*, July 15, 2006.

In the U.S. market, all of the beer brands and types produced by perennial market leader Anheuser Busch hold about 50 percent market share. The second two companies, SABMiller and Molsen Coors, have a combined market share of around 30 percent. The remaining share comprises second-tier brewers such as Stroh's Brewery, craft brewers such as the Boston Beer Company, and imports such as Corona (which in fact is 50 percent owned by Anheuser Busch). Leading imports include Heineken (based in the Netherlands) and Stella Artois (owned by the global giant InBev). So is the U.S. market an oligopoly? Just looking at the market share held by the top three companies, the answer would seem to be yes.

However, based on pricing power, the answer would seem to be more like monopolistic competition. This is because the U.S. market for beer is very much a mature one. The growth rate in recent years was less than 1 percent and demand for mass market regular beer has been falling about 2 percent per year. Only the increase in consumption of craft and imported beers has kept overall market sales from falling. And as a result of the fall in the demand for mass market brands, even the leading manufacturers such as Anheuser Busch have had to drop their prices. The only way that Anheuser Busch can offset this decline in price is to try to sell more of the higher profit margin imports or craft-brewed beer. Another way that this U.S.-based company can offset the stagnating U.S. beer market is to expand globally. But in this respect, it has been outdone in recent years by two leading global companies based in Europe, Heineken and InBev, and a third global player from South Africa, South African Breweries (SAB), now known as SABMiller.

The Heineken brand is probably well known to American consumers because it is one of the leading imported beers. InBev is the result of the merger of Interbrew, a Belgian company, and AmBev, a Brazilian company. Their popular brands include Stella Artois and Brahma. In 2003, SAB made the headlines by buying Miller Brewing from Philip Morris. At the time, Miller was the third largest beer company in the United States. It still is. The fourth largest, Coors, recently merged with the Canadian company Molsen. The new company is now called Molsen Coors Brewing Company. In 2008, InBev bought Anheuser Busch, making it the biggest beer company in the world.

All of the leading global giants are trying to expand by acquisition. As they grow, they hope to use the increasing power of economies of scale and scope to reduce their unit cost of production. Also their expansions are typically in areas of the world such as Brazil, China, Russia, and Central Europe, where beer sales are still expanding at rates 3–4 percentage points above the mature markets in the United States and Western Europe.

In 2007, SABMiller and Molsen Coors Brewing merged their U.S. units and also bought Royal Grolsch, a Dutch company that is Heineken's closest rival in the Netherlands. In January 2008, Heineken and Carlsberg (a Danish company) announced that they had agreed to buy the largest brewery in the United Kingdom, Scottish & Newcastle, for \$15.4 billion. The U.K. brewer is best known for brands such as Foster's and Kronenbourg 1664. Among the key reasons for this acquisition is that Heineken will gain access to Britain's cider market and Carlsberg will gain control of Scottish & Newcastle's holding of the Russian brewer, Baltic Beverages. The cider market in Britain is growing at double-digit rates, and Russia's beer market is expected to grow 5 percent in 2008.²⁸

There are still many hundreds of breweries around the world. So it is safe to say that the global market can still be considered monopolistically competitive. However, at the rate that global acquisitions and mergers are being carried out, perhaps even the global market may become more like an oligopoly.

²⁸Julia Werdigier, "Scottish & Newcastle Agrees to Be Bought and Split," *New York Times*, January 20, 2008.

The Solution

A few weeks later Frank again met with Nicole. He recommended selling “Go-nuts” for \$1.00 instead of the \$1.10 suggested by the $MR = MC$ analysis. His focus group studies showed that people paid more attention to a 25-cent than a 15-cent difference. The average price of the major national brands was \$1.25.

“I believe the 25-cent difference will be really noticeable to the consumer,” Frank explained. “After all, we’re new entrants into the market and we can use this lower price to attract customers away from the competition. At least they’ll be more inclined to give our product a try. But we’re not going to sell this product as a private label or discount product. According to your marketing plan, Nicole, we’re going to be putting a lot of money into advertising to build brand awareness among consumers.”

“That’s just it, Frank, we are in fact going to be positioning “Go-nuts” as a premium product, as good as any other national brand, if not better,” Nicole responded. “If we’re going to position our product in this way, we have got to use our price in support of this. If we set the price at \$1.00 and then decide to raise it to \$1.25 after it gains customer acceptance, they may not go along with this increase. We have to use promotion and advertising more than price to build our brand in the marketplace. Why don’t we just match the competition’s \$1.25?”

“It makes sense to me,” Frank answered. “The one thing we don’t want to do is to trigger a price war. If any of our competitors with established brands decides to match or even beat our lower price, then the whole industry might suffer.”

“Hey, now I’ve really got you thinking like a marketing person, Frank. Remember, demand is based on the customer’s *perception* of value and not the intrinsic value of the product. If we can use advertising to support the price of \$1.25, that’s going to be the key to our success. Besides, you’re right about a price war—I don’t even want to think about starting a price war with Red Bull.”

SUMMARY

This chapter examines the pricing and output decisions faced by firms in monopolistic competition and oligopoly. Firms in oligopolistic markets have a more challenging task because of mutual interdependence. However, if oligopoly firms are of sufficient size or are very effective in differentiating their products, they may not have as much competition from new market entrants as those firms operating in monopolistic competition. In both types of markets, nonprice decisions are an important part of the competitive environment. A critical part of the success of a firm’s operations in imperfectly competitive markets is the development (as well as implementation) of an effective business strategy. Therefore, the final sections of this chapter examined the important elements of business strategy and their linkages to the terms and concepts covered in managerial economics.

IMPORTANT CONCEPTS

Five Forces model. Model developed by Michael Porter that shows the key factors that affect the ability of a firm to earn an economic profit: potential entrants, bargaining power of suppliers, bargaining power of buyers, threat of substitute products or services, and intramarket rivalry. Also referred to as the Porter Competitive Framework. (p. 376)

Kinked demand. A theoretical construction that attempts to explain price rigidities in oligopolistic markets. (p. 367)

Monopolistic competition. A market distinguished from perfect competition in that each seller attempts to differentiate its product from those of its competitors (e.g., in terms of location, efficiency of service, product attributes, advertising, or promotion). Good examples of this type of market can be found in small businesses, particularly those in the retail trades. (p. 363)

Mutual interdependence. A situation in which each firm in the market sets a price based on its costs, price elasticity, and anticipated reaction of its

competitors. This type of pricing situation prevails in oligopolistic markets. (p. 362)

Oligopoly. A market in which there is a small number of relatively large sellers. Pricing in this type of market is characterized by mutual interdependence among the sellers. Products may either be standardized or differentiated. (p. 365)

Price leader. One company in an oligopolistic industry establishes the price, and the other companies follow. Two types of price leadership, barometric and dominant, are discussed in Chapter 10. (p. 369)

Strategy. The means by which an organization uses its scarce resources to relate to the competitive

environment in a manner that is expected to achieve superior business performance over the long run. (p. 373)

Structure-conduct-performance (S-C-P) paradigm. An approach to studying industrial economics that states that an industry's structure determines an industry's conduct, which in turn affects the industry's performance. The key factors that shape industry structure are the number of firms in the industry, the conditions of entry and exit, and product differentiation. (p. 374)

QUESTIONS

1. Explain the key difference between perfect competition and monopolistic competition.
2. Assume firms in the short run are earning above-normal profits. Explain what will happen to these profits in the long run for the following markets:
 - a. Pure monopoly
 - b. Oligopoly
 - c. Monopolistic competition
 - d. Perfect competition
3. In certain industries, firms buy their most important inputs in markets that are close to perfectly competitive and sell their output in imperfectly competitive markets. Cite as many examples as you can of these types of businesses. Explain why the profits of such firms tend to increase when there is an excess supply of the inputs they use in their production process.
4. In the short run, firms that seek to maximize their market share will tend to charge a lower price for their products than firms that seek to maximize their profit. Do you agree with this statement? Explain.
5. Explain why it is sometimes difficult to apply the $MR = MC$ rule in actual business situations.
6. Define *mutual interdependence*.
7. Why do oligopolists often rely on a price leader to raise the market price of a product?
8. How does one determine whether a market is oligopolistic? Is it important for managers to recognize the existence of oligopolistic competitors in the markets in which their companies operate? Explain.
9. In the following list are a number of well-known companies and the products that they sell. Which of the four types of markets (perfect competition, monopoly, monopolistic competition, and oligopoly) best characterizes the markets in which they compete? Explain why.
 - a. McDonald's—hamburgers
 - b. ExxonMobil—gasoline
 - c. Dell—personal computers
 - d. Heinz—ketchup
 - e. Procter & Gamble—disposable diapers
 - f. Starbucks—gourmet coffee
 - g. Domino's—pizza
 - h. Intel—computer chip for the PC
10. Briefly explain the structure-conduct-performance approach to the study of industrial economics.
11. Compare and contrast Porter's Five Forces model with the four basic types of markets first described in Chapter 8 in the section "Market Structure."

12. In 2002, Philip Morris sold its Miller Brewing Division to South African Breweries.
 - a. What impact do you think this transaction had on the market structure of the beer industry in the United States? In world markets? Explain.
 - b. Using the economic concepts presented in Chapters 7, 8, and 9, discuss possible reasons why both parties agreed to this transaction.

PROBLEMS

1. A group of five students has decided to form a company to publish a guide to eating establishments located in the vicinity of all major college and university campuses in the state. In planning for an initial publication of 6,000 copies, they estimated the cost of producing this book to be as follows:

Paper	\$12,000
Research	2,000
Graphics	5,000
Reproduction services	8,000
Miscellaneous	5,000
Personal computer	2,000
Desktop publishing software	500
Overhead	5,500
Binding	3,000
Shipping	2,000

By engaging in this business, the students realized that they would have to give up their summer jobs. Each student made an average of \$4,000 per summer. However, they believed they could keep expenses down by doing much of the research for the book by themselves with no immediate compensation.

They decided to set the retail price of the book at \$12.50 per copy. Allowing for the 20 percent discount that retail stores in their state generally required, the students anticipated a per-unit revenue of about \$10.00. The director of the campus bookstore advised them that their retail price was far too high, and that a price of about \$8.75 would be more reasonable for a publication of this kind.

One of the students, who was a math and statistics major, asked the bookstore manager to provide her with historical data on sales and prices of similar books. From these data, she estimated the demand for books of this kind to be

$$Q = 18,500 - 1,000P$$

where Q = Number of books sold per year

P = Retail price of the books

- a. Construct a numerical table for the retail demand curve, and plot the numbers on a graph. Calculate the elasticity of demand for the interval between \$12.50 and \$8.00.
 - b. Do you think the students should follow the store manager's advice and price their book at \$8.75? Explain. If you do not agree with this price, what would be the optimal price of the book? Explain.
 - c. Assuming the students decide to charge the optimal price, do you think they should proceed with this venture? Explain.
 - d. Assuming the student's demand equation is accurate, offer some possible reasons why the bookstore manager would want to sell the book at the lower price of \$8.75.
2. Use the same data presented in Problem 1 to answer the following questions:
 - a. Explain the impact on the optimal price of designating the "miscellaneous" cost item as fixed versus variable. (*Hint*: Do the pricing analysis assuming miscellaneous is a fixed cost and compare it with an analysis that assumes it is a variable cost.)
 - b. Under what circumstances do you think the average variable cost would *increase* (as is generally expected in the economic analysis of cost)? Do you think the law of diminishing returns would play a role in increasing AVC? Explain.
 - c. Under what circumstances do you think the average variable cost would *decrease*? Explain.

3. A firm in an oligopolistic industry has identified two sets of demand curves. If the firm is the only one that changes prices (i.e., other firms do not follow), its demand curve takes the form $Q = 82 - 8P$. If, however, it is expected that competitors will follow the price actions of the firm, then the demand curve is of the form $Q = 44 - 3P$.
 - a. Develop demand schedules for each alternative, and draw them on a graph.
 - b. Calculate marginal revenue curves for each.
 - c. If the present price and quantity position for the firm is located at the intersection of the two demand curves, and competitors follow any price decrease but do not follow a price increase, show the demand curve relevant to the firm.
 - d. Draw the appropriate marginal revenue curve.
 - e. Show the range over which a marginal cost curve could rise or fall without affecting the price the firm charges.
4. Indicate whether each of the following statements is true or false, and explain why.
 - a. A competitive firm that is incurring a loss should immediately cease operations.
 - b. A pure monopoly does not have to worry about suffering losses because it has the power to set its prices at any level it desires.
 - c. In the long run, firms operating in perfect competition and monopolistic competition will tend to earn normal profits.
 - d. Assuming a linear demand curve, a firm that wants to maximize its revenue will charge a lower price than a firm that wants to maximize its profits.
 - e. In an oligopoly, the firm that has the largest market share will also be the price leader.
 - f. The demand curve facing a firm in a monopolistically competitive market is more elastic than one facing a pure monopoly.
5. A phenomenon in the retail merchandising of food and clothing in the United States and the United Kingdom is the growing popularity of private-label (also called store-brand) products. These products are priced at a lower level than the premium national brands. Use the concepts of price elasticity and relevant cost to explain the profitability of these products from the point of view of
 - a. The retail stores that sell these private-label products.
 - b. The manufacturers of these private-label products.

If you were the manager of a national premium brand, what would you do to fight the growing competition of private labels?
6. Suppose three firms face the same total market demand for their product. This demand is:

P	Q
\$80	20,000
70	25,000
60	30,000
50	35,000

Suppose further that all three firms are selling their product for \$60 and each has about one-third of the total market. One of the firms, in an attempt to gain market share at the expense of the others, drops its price to \$50. The other two quickly follow suit.

- a. What impact would this move have on the profits of all three firms? Explain your reasoning.
 - b. Would these firms have been better off in terms of profit if they all had raised the price to \$70? Explain.
7. A firm has the following short-run inverse demand and cost schedules for a particular product:

$$P = 45 - 0.2Q$$

$$TC = 500 + 5Q$$

- a. At what price should this firm sell its product?
- b. If this is a monopolistically competitive firm, what do you think would happen as the firm moves toward the long run? Explain.

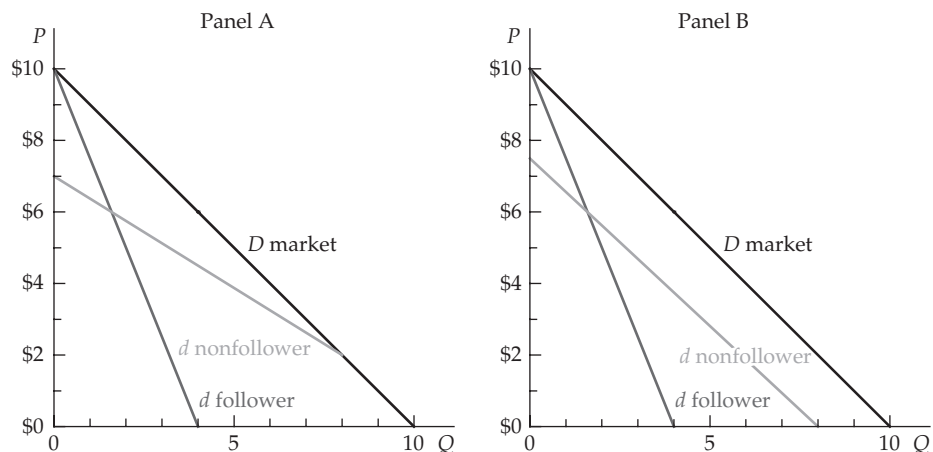
- c. Suppose in the long run, that inverse demand shifts to $P = 25 - 0.2Q$. What should the firm do? Explain. Provide graphs for both the short-run and long-run scenarios. Make sure your graphs include demand, MR, MC, and AC.
 - d. Suppose, in the long run, that inverse demand shifts to $P = 45 - 0.8Q$. What should the firm do? Explain, and provide graphs for both the short-run and long-run scenarios. Make sure your graphs include demand, MR, MC, and AC.
 - e. Does the demand in part c or part d represent a change in market share for the representative firm as a result of the entry or exit suggested in part b?
8. Redo Problem 7a to e with an initial inverse demand given by:

$$P = 45 - Q$$

9. Professor Michael Porter's generic strategy options for competing are the differentiation approach and the cost leadership approach. The first involves competing by having a better product and the second by having a lower cost than one's competitors. Relate this strategy to the monopolistically competitive model presented in this chapter. In particular, use Figure 9.1 to explain the rationale for Porter's generic strategies.
10. Suppose short-run inverse demand in a monopolistically competitive market is represented by: $p(x) = 18 - 0.2x$. Cost is given by: $TC(x) = 320 + 2x + 0.05x^2$.
 - a. Given these demand and cost conditions, what price, output, and profits result in the short run?
 - b. What will happen as the firm moves from the short to the long run?

Suppose entry or exit causes the demand curve to rotate on its p axis. In particular, assume that inverse demand can be described by: $p(x) = 18 - m \cdot x$.

- c. What output and price result in LRMCE?
11. Two natural boundaries to rival reaction are for the rival to not react (or not follow) a firm's actions or to follow a firm's actions. Consider pricing in a differentiated product market. If all firms in the oligopolistic market follow each other's price changes, then the size of the market may change, but the market share controlled by each firm would not appreciably change. On the other hand, if other firms do not follow a price change, then the market share controlled by the firm that has changed its price will change. Suppose we denote market demand by " D market,"; demand given price matching behavior by " d follower"; and demand given price changes are not matched by " d nonfollower," similar to Figure 9.2. Two panels are shown with these three demand curves.



Each panel has the same market and follower demand. Based on these panels, answer the following questions.

- a. What market share does the firm depicted have along the follower demand curve?
 - b. What is this firm's current price?
 - c. Which panel depicts a market in which niche players have a stronger brand identity and are able to maintain at least some of their market despite other firms dropping their price? Answer this problem by making specific reference to what happens in the face of price decreases in each panel.
12. Oligopolistic models are based on behavioral assumptions. One behavioral assumption associated with differentiated product markets is that price increases will not be matched, but price decreases will be matched. This rather pessimistic view of pricing leads to the kinked demand curve discussed in Figure 9.3. To examine why, consider the following simple model: Market inverse demand is given by $P(Q) = 10 - Q$.
- a. Suppose firm A controls 50 percent of the market. What is the demand curve faced by this firm? Write inverse demand in slope-intercept form.

Suppose the current price of the product is \$6.

- b. What is demand faced by firm A given $P = \$6$? Call this quantity Q_b .
- c. Suppose that, if firm A increases price from this point, other firms do not match the price increase. But if A decreases price, other firms decrease price to maintain their market share. The demand in this instance has two segments: the segment above $P = \$6$ and the segment below $P = \$6$. What should happen to market share for prices above \$6? What happens to market share for prices below \$6?

The final question that must be answered is how quickly does market share decline as price increases. Suppose A's demand is linear above $P = \$6$ and A is unable to sell any output above $P = \$8$.

- d. Describe algebraically the inverse demand curve faced by the firm in this instance. Provide a graph that is consistent with your answer. Based on this graph, explain why this is called the kinked demand model. (*Hint:* The equation for the inverse demand curve must be done in two parts.)
13. This problem examines the behavior of marginal revenue in the kinked demand model and requires completion of Problem 12.
- a. What is MR for $Q \leq Q_b$? Provide an equation for MR in this instance and graph this on the graph begun in Problem 12d.
 - b. What is MR for $Q \geq Q_b$? Provide an equation for MR in this instance and graph this on the graph begun in Problem 12d.
 - c. What happens if marginal cost is \$3? How much should firm A produce?
 - d. What if marginal cost increased to \$4 or decreased to \$2? Explain why this happens.
14. (This problem requires completion of Problems 12 and 13.)

The magnitude of the kink in the kinked demand model is often modeled as a function of the market share controlled by the firm. A small-market-share firm in a differentiated product market may well occupy a niche market that will sustain higher prices before all customers are lost to other lower-price firms than a large-market-share firm. Suppose we model this as a linear function of market share. In the previous problems, we posited that a 50 percent market share firm had a \$2 decline in price intercept (from \$10 to \$8). Suppose a 25 percent market share firm has a \$1 decline in price intercept. We can examine what happens for different market share firms given the behavioral assumption underlying the kinked demand model.

Using the market inverse demand of $P = 10 - Q$, re-answer Problems 12 and 13 for a 25 percent market share firm. Because this assumes you have already done those problems, your answers should be brief and point out when substantive differences occur as a result of the change in market share held by the firm.

15. Consider the pizza market in a small college town. Analyze this market based on the following assumptions:
- The market is in long-run equilibrium.
 - Each pizza shop sells 100 pizzas per week.
 - For ease of exposition, suppose that each shop sells *only* pizza and *only* one size.
 - Fixed cost for each shop is \$500 per week.
 - Price and elasticity for Salamandra's (**s**), Genoa's (**g**), Domino's (**d**), and Four Star (**4**) are:
 $P_s = \$11.00, \epsilon_s = -2.2; P_g = \$11.00, \epsilon_g = -2.75; P_d = \$9.00, \epsilon_d = -1.8; P_4 = \$8.00, \epsilon_4 = -2.$
- What market structure best describes the pizza market in this town? Explain.
 - What is average variable cost at this output level for each of the four shops? Explain how you derived this result.
 - Based on your answers to parts a and b and based on assumptions i to v, are any of these four firms earning above normal profit? Explain your answer.
 - What is the marginal cost of production for each shop? Explain how you derived these values.
 - Which firm(s) are producing on the downward sloping part of their AVC curve, and which are producing on the upward sloping parts of their AVC curve? Explain.
 - Which firm(s) are producing on the downward sloping part of their AC curve, and which are producing on the upward sloping part of their AC curve? Explain.
 - Which firm has the lowest percent markup of price over marginal cost and which firm has the highest markup? Explain. Provide a percentage markup for both the highest and lowest firm (to the nearest 0.1 percent).

A Mathematical Restatement of Monopolistic Competition

As the name implies, monopolistic competition has aspects of monopoly as well as competition. The product is differentiated, which leads to each firm having some price-making ability. This is the monopoly aspect of monopolistic competition. Entry and exit is easy, which is the competitive aspect of monopolistic competition. Short-run profits or losses lead to zero profits in the long run. Mathematically, long-run monopolistic competition equilibrium (LRMCE) involves two simultaneous conditions at the same output: LRMCE requires two conditions:

a. $MR = MC$ (for profit maximization) **(9A.1a)**

b. $p = AC$ (for zero profits) **(9A.1b)**

This is shown graphically in Figure 9.1c. These conditions need not always hold, but if they do, then the firm is in LRMCE. This is most readily understood by working through an example.

Suppose demand is represented by $p(x) = 21 - 0.16x$ (the lower case x reminds us that this is demand as perceived by an individual firm). Cost is given by $TC(x) = 400 + x + 0.04x^2$. Given these demand and cost conditions, what price, output, and profits result in the short run?

To find the short-run equilibrium, we must set $MR = MC$ and solve for x . Marginal revenue requires taking the derivative of total revenue:

$$TR(x) = p(x) \cdot x = 21x - 0.16x^2 \quad \mathbf{(9A.2)}$$

$$MR(x) = dTR/dx = 21 - 0.32x \quad \mathbf{(9A.3)}$$

Note that this confirms the MR has the same intercept and twice the slope as demand rule. Marginal cost is the derivative of total cost:

$$MC(x) = dTC/dx = 1 + 0.08x \quad \mathbf{(9A.4)}$$

Profits are maximized where $MR = MC$:

$$21 - 0.32x = 1 + 0.08x$$

Regrouping,

$$20 = 0.4x$$

and solving for x :

$$x = 20/0.4 = 50 \quad \mathbf{(9A.5)}$$

Find price and AC at this output level in order to determine profits:

$$p(50) = 21 - 0.16 \cdot 50 = 21 - 8 = \$13 \quad \mathbf{(9A.6)}$$

$$AC(x) = TC/x = 400/x + 1 + 0.04x$$

so

$$\begin{aligned} AC(50) &= 400/50 + 1 + 0.04 \cdot 50 \\ &= 8 + 1 + 2 = 11 \end{aligned} \quad \mathbf{(9A.7)}$$

Finally, profits are given by $\pi = (p - AC)Q$:

$$\pi = (13 - 11) \cdot 50 = 100 \quad \mathbf{(9A.8)}$$

Short-run profits imply that entry will occur. As entry occurs, individual firm demand will shift inward. How quickly it shifts depends on exactly *how* entry occurs in this differentiated product market. One easy-to-model solution is that entry rotates demand but maintains the p intercept of 21. In this instance demand is given by:

$$p(x) = 21 - mx \quad \mathbf{(9A.9)}$$

Marginal revenue is:

$$MR(x) = 21 - 2mx \quad \mathbf{(9A.10)}$$

Entry will continue until the twin LRMCE conditions in Equation (9A.1) result. These conditions produce two equations in two unknowns, slope, m , and output, x .

$$MR = MC:$$

$$21 - 2mx = 1 + 0.08x \quad \mathbf{(9A.11a)}$$

$$p = AC:$$

$$21 - mx = 400/x + 1 + 0.04x \quad \mathbf{(9A.11b)}$$

Multiplying Equation (9A.11b) by 2 and moving the resulting 42 to the right side of the equation yields

$$\begin{aligned} -2mx &= 800/x + 2 + 0.08x - 42 \\ &= 800/x - 40 + 0.08x \quad (9A.11c) \end{aligned}$$

Substitute Equation (9A.11c) into Equation (9A.11a):

$$21 + 800/x - 40 + 0.08x = 1 + .08x$$

Regrouping,

$$800/x = 20$$

and solving for x ,

$$x = 800/20 = 40 \quad (9A.12)$$

Because $p = AC$, we obtain,

$$p = 400/40 + 1 + 0.04 \cdot 40 = \$12.60 \quad (9A.13)$$

The short-run and long-run solutions are readily graphed using Excel, as shown in the Monopolistic Competition Excel App in worksheet 9A.1. To accomplish this with minimum effort, one final calculation is worthwhile, finding the slope, m , from Equation (9A.9) that occurs in LRMCE. This is most easily obtained using the definition of slope

because we know two points on this line: $(0, \$21)$ and $(40, \$12.60)$.

$$\begin{aligned} m &= \Delta y / \Delta x = (12.6 - 21) / (40 - 0) \\ &= -0.21 \quad (9A.14) \end{aligned}$$

Note that this implies that the x intercept of demand is at $x = 100$ in the LRMCE diagram. The twin conditions required for LRMCE are boxed and highlighted at $x = 40$ in Excel Sheet 9A.1 and are also shown in the lower graph at that same output level.

As a way of checking the solution, readers with a calculus background can verify that, in this instance, demand is tangent to AC at $x = 40$. The slope of AC is

$$dAC/dx = -400/x^2 + 0.04 \quad (9A.15a)$$

Evaluating this at $x = 40$ we obtain

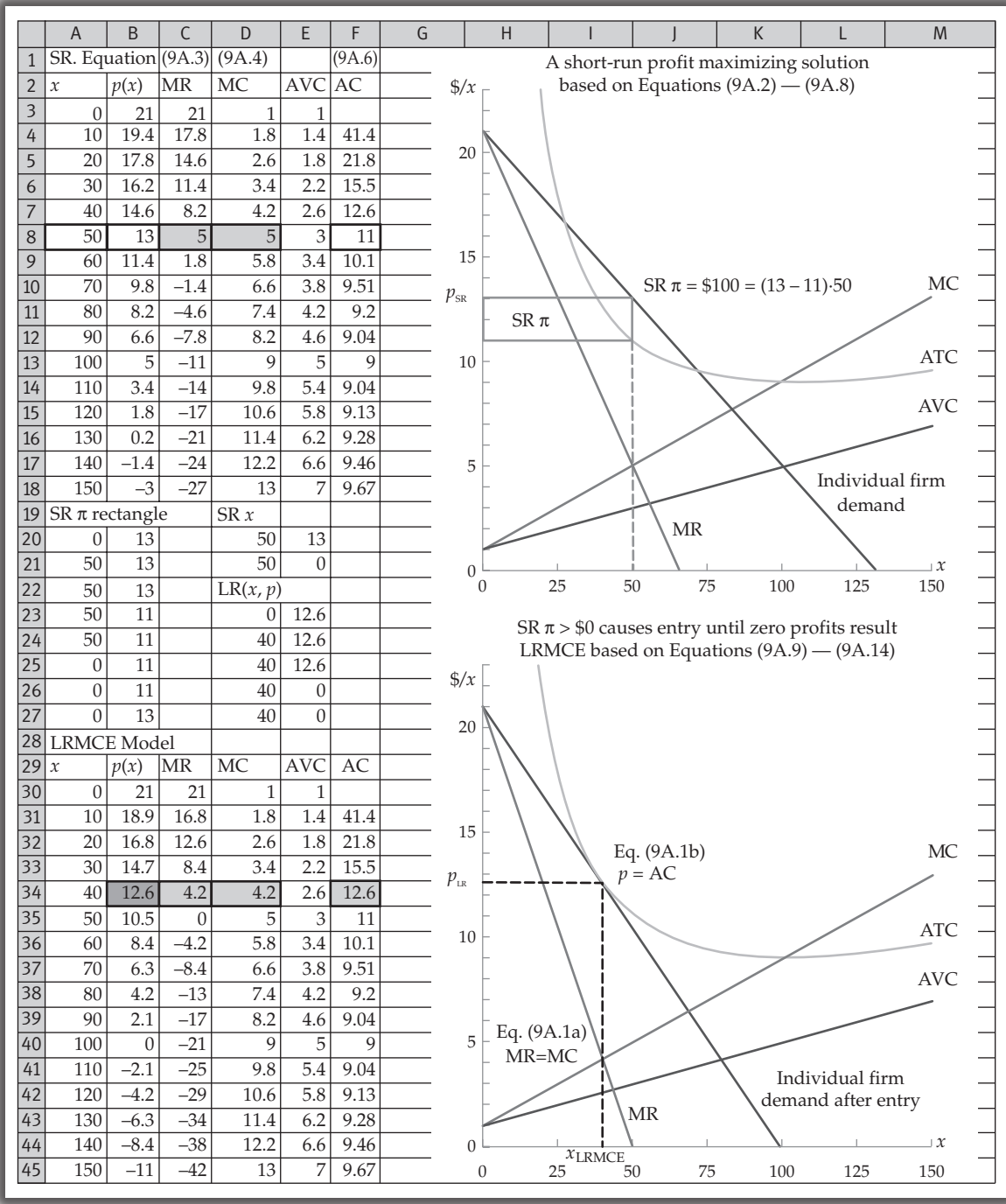
$$\begin{aligned} dAC/dx &= -400/40^2 + 0.04 = -10/40 + 0.04 \\ &= -0.25 + 0.04 = -0.21 \quad (9A.15b) \end{aligned}$$

As expected, the slope of AC in Equation (9A.15b) is the same as the slope of demand derived in Equation (9A.14).



APP

Excel Sheet 9A.1



Source: Monopolistic Competition Excel App, worksheet 9A.1

Special Pricing Practices

Learning Objectives

Upon completion of this chapter, readers should be able to:

- Analyze cartel pricing.
- Illustrate price leadership.
- Understand price discrimination, and explain how it affects production and prices.
- Distinguish between marginal pricing and “cost-plus” pricing.
- Discuss the various types of multiproduct pricing.
- Explain the meaning of “transfer pricing,” and explain how a company should price products that pass from one operating division to another.

The Situation

One of the most difficult challenges in the food and beverage industries is the establishment of effective channels of distribution. Many food-processing and beverage companies rely on food brokers to sell their products to retail outlets such as supermarkets and grocery stores. In the case of energy drinks, the product is shipped from the bottling plants to the individual retail establishments. Obviously, there must first be a willingness on the part of these retail businesses to carry a particular product line.

The task of establishing the relationship between Global Foods, Inc., and the retail stores was given to Rebecca James, assistant vice president of marketing of the beverage division. Because the product, “Go-nuts,” was so new, she found considerable resistance among the major supermarket chains to carrying Global’s line of energy drinks. Thus far, she had been able to sell relatively small volumes to smaller grocery stores, convenience stores, delicatessens, and sandwich shops. Then she learned that a large catering company that provided food service in major airports all across the United States wanted to carry an additional line of energy drinks. This firm had put out a request for bids to all the major beverage companies and also to Global Foods. After all, Global had an established reputation in the food business.

Rebecca was eager to land this major account. But she also realized that her bid would have to be considerably lower than that offered to her present customers. However, this lower price would be more than made up by the potential volume of sales, as well as by the creation of a base from which to further penetrate the market for energy drinks. But she was not quite sure how to decide on the price she should recommend. She decided to consult with Philip Olds, an executive in the company’s foods service division. Philip had considerable experience in preparing bids for large customer contracts.

INTRODUCTION

In Chapters 8 and 9, we discuss output and pricing decisions under different market arrangements. We now continue this discussion and apply our knowledge to pricing decisions made in specific situations. We are also confronted by some complications. Whereas we have previously assumed a firm produces only one product, we now have to allow for the pricing of several products simultaneously. The situations we encounter in this chapter usually occur under imperfectly competitive conditions.

CARTEL ARRANGEMENTS

Competition is a very tough taskmaster. To survive in competition in the long run, a company must operate at its most efficient (minimum) cost point, and it will earn no more than a normal return. Thus there is always an incentive for a company to try to become more powerful than its competitors—in the extreme, to become a monopolist. In an oligopolistic type of industry, where there are several powerful firms, it would probably be impossible for one firm to eliminate all the others. So, to reap the benefits of a monopoly (i.e., higher profits and the general creation of a more certain and less competitive environment), it may be advisable for companies in the industry to act together as if they were a monopoly. In other words, they all agree to cooperate with one another; they form a **cartel**. Cartel arrangements may be tacit, but in most cases some sort of formal agreement is reached. The motives for cartelization have been recognized for many years. Indeed, an early recognition can be found in a passage in Adam Smith's famous book written in 1776: "People of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices."¹

Cartels were made illegal in the United States with the passage of the Sherman Anti-Trust Act of 1890. Thus, most "official" cartels are found in countries other than the United States. Probably the most famous cartel in existence today is the Organization of Petroleum Exporting Countries (OPEC). But there are others, such as the IATA (International Air Transport Association), to which U.S. airlines can belong. Collusive agreements have also existed in the United States. One of the most famous cases of price and market share fixing was in the electrical industry and involved General Electric, Westinghouse, and other large corporations. The case was tried and concluded in 1961 and resulted in prison sentences for several executives and large fines. This case and other more recent ones are described briefly following a discussion of the characteristics and effects of cartels.

Cartels may not flourish in all oligopolistic markets. Following are some of the conditions that influence the formation of cartels.

1. The existence of a small number of large firms facilitates the policing of a collusive agreement.
2. Geographic proximity of the firms is favorable.
3. Homogeneity of the product makes it impossible for cartel participants to cheat on one another by emphasizing product differences.
4. The role of general business conditions presents somewhat contradictory arguments. Cartels are often established during depressed industry conditions, when companies attempt to forestall what they consider to be ruinous price cutting. However, it also appears that cartels disintegrate as demand for the product falls, and each member thinks

¹Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations*, New York: Modern Library, 1937, p. 128.

it can do better outside the cartel. The cartel may then reestablish itself during the recovery period. Thus cartels can form or fall apart during either phase of the business cycle.²

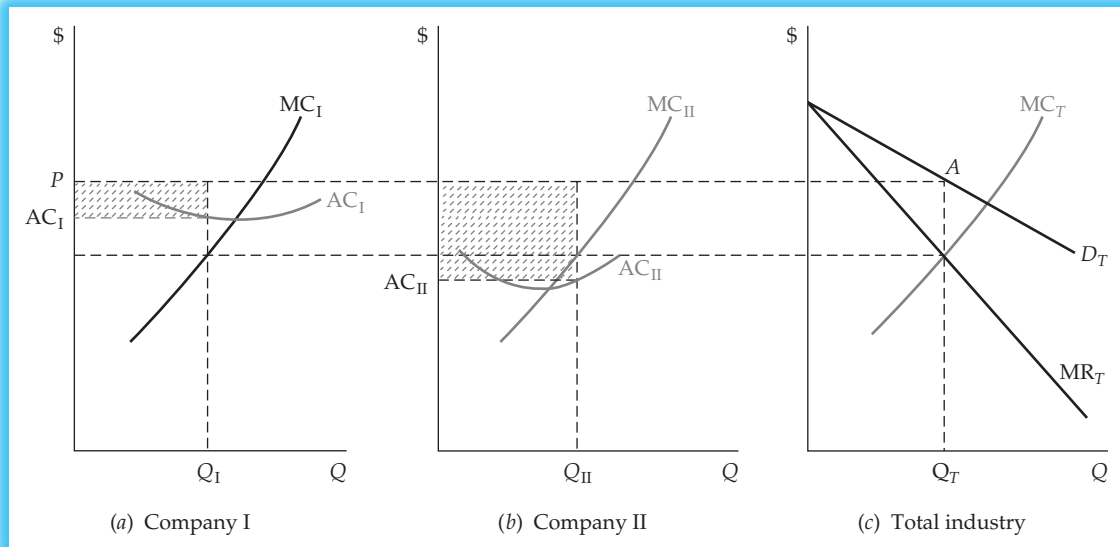
5. Entry into the industry must be difficult. The case of OPEC is a good example. It is impossible for countries that do not possess the basic resource to begin petroleum production and compete for monopoly profits.
6. If cost conditions for the cartel members are similar and profitability thus will not differ greatly among members, cartels will be easier to maintain. Product homogeneity, mentioned earlier, will contribute to cost uniformity.

The ideal cartel will be powerful enough to establish monopoly prices and earn maximum monopoly profits for all the members combined. This situation is illustrated in Figure 10.1. For simplicity, assume there are only two firms in this oligopolistic industry. The total industry demand curve is shown in Figure 10.1c. The marginal revenue curve is constructed for this demand curve in the usual manner. Each of the two competitors (illustrated in Figure 10.1a and b) has its respective average total cost and marginal cost curves, which can differ.

The two individual marginal cost curves are then added horizontally, and the result is plotted on the industry graph (MC_T). Industry output will take place where MC_T equals the industry marginal revenue, and the price charged will be found by drawing a vertical line to the demand curve (point A). This is, of course, the classic monopoly situation, and monopoly profits will be maximized at this point.

The next step is to establish how much each of the two companies will sell at this price. For the entire industry output to be sold, each company will sell that output

Figure 10.1 The Ideal Cartel



²A very cogent description of this phenomenon can be found in George J. Stigler, *The Theory of Price*, New York: Macmillan, 1949, pp. 274–75. Stigler also puts forward the idea that, regardless of the level of business activity, government action in support of such collusion is an important factor in a cartel's success. An important example is the U.S. National Recovery Act (NRA) of the 1930s, later declared unconstitutional by the Supreme Court.

corresponding to the point at which a horizontal line drawn from the $MC_T = MR_T$ intersection on the industry graph crosses the marginal cost curve of each of the two firms. It can be seen that each firm will produce different quantities and achieve different profits, depending on the level of the average total cost curve at the point of production. Generally, the lower average cost company will be the more profitable one (profits for the two companies are shown by the hatched areas in Figure 10.1). This result, although maximizing combined profits, may also be one of the reasons for the subversion of cartels. A very efficient company with low average costs, and most likely with excess capacity under cartel conditions, may find it profitable to cheat by offering its product at a lower price and capturing a larger share of the total business.

Such a cartel may be unstable. Unless strictly enforced, cartels will have a tendency to break down. Secret price cuts may be extremely profitable because (if the product is undifferentiated) the demand curve for an individual firm in a cartel will be quite elastic. Cartel subversion often occurs during slumps in demand because individual members will be looking to increase their share to avoid significant quantity decreases.

It must also be remembered that collusion is costly. First, there is the cost of forming the cartel. Second, there is a cost of monitoring the actions of the cartel members and of enforcing the rules to minimize cheating. There is also the potential cost of punishment by authorities. Thus, in the end, cartelization may not necessarily be profitable. In short, the additional revenues obtained by cartel members due to collusion must exceed the costs just described. We can, therefore, state that, although profit maximization is the incentive that leads to collusion, it may also be the cause of a cartel's breakdown.

Cartels often have agreements specifying the market share of each participant. Such allotments may be based on history, or they can be arranged to give each member a certain geographic area. Collusion can also exist in much more informal ways. Thus, physicians within a geographic area coincidentally charge similar fees for their services. Trade associations are often suspected of collecting and conveying information that will lead to the fixing of prices.

Cases of Price Fixing by Cartels

A classic case of price fixing and market sharing ended in February 1961 in a Philadelphia court, when seven executives of General Electric, Westinghouse, and other companies were sent to prison and fined; twenty-three others were given suspended sentences and fined; and twenty-nine companies were fined a total of approximately \$2 million. Starting shortly after World War II, this conspiracy involved a number of heavy electrical equipment products, such as switching gears, circuit breakers, transformers, and turbine engines. The companies involved pleaded guilty and no contest to the federal indictments.

The story of these collusive practices reads like a mystery story.³ There were meetings in hotel rooms during conventions of the National Electric Manufacturers Association. There were hotel meetings at various locations in which the participants did not register under their company affiliations and recorded trips to other locations in their expense accounts. There were code numbers given to each company. There were telephone calls at the participants' homes and even a conspiratorial round of golf.

³The description of this case has been obtained from two articles in *Fortune* (April 1961, pp. 132–37 ff, and May 1961, pp. 161–64 ff), and from the *Wall Street Journal*, January 10 and 12, 1962.

A more recent case is that of the Archer Daniels Midland Company (ADM), a large agricultural business with annual revenues of approximately \$80 billion in 2011. In October 1996, ADM pled guilty to price fixing of two of its products, lysine, a feed supplement, and citric acid, a food additive. Four Asian firms were implicated in the lysine conspiracy, and four European companies were involved in the citric acid case. The price fixing and division of markets was arranged in a series of secret meetings in many locations, including Mexico City, Paris, Tokyo, and Atlanta. ADM was fined a total of \$100 million, and in addition paid approximately \$90 million to settle customer and stockholder suits.

Two years later, in September 1998, three ADM executives were convicted of participating in the case. One of them was the former vice chairman, Michael Andreas, son of ADM's chief executive officer and his likely successor. He was sentenced to 2 years in prison and assessed a \$350,000 fine.⁴

There is an additional chapter to the saga of Archer Daniels Midland. In 1995, the U.S. Department of Justice began investigating another anticompetitive pricing action involving ADM, Cargill, and two other companies. This investigation concerned the pricing of fructose corn syrup. The plaintiffs comprised some 2,000 companies, including Coca-Cola and PepsiCo. Although the other three companies had settled with the plaintiffs previously (for a total of \$31 million), the case was finally closed in June 2004, when ADM agreed to settle by paying \$400 million.⁵

In 1997, the U.S. Department of Justice began to investigate possible collusion in setting buying and selling commissions by two leading art auction houses. The two companies, Sotheby's and Christie's, together controlled 90 to 95 percent of the art auction market. Although Sotheby's and Christie's appeared to be active competitors in soliciting business, in 1992 they raised their buyers' commissions to identical rates within 6 weeks of one another. Then in 1995 they took similar action on sellers' commissions, again within 6 weeks of one another. In January 2000, Christie's CEO turned over a large number of documents to the Department of Justice and agreed to cooperate with the department; in return, he and his firm were granted immunity from prosecution. In February 2001, a federal judge accepted Sotheby's guilty plea to conspiracy to fix prices and fined the company \$45 million. During the same month, a \$512 million class-action suit settlement was reached between the two auction houses and some 130,000 buyers and sellers. In addition, Sotheby's settled with a group of shareholders who claimed that the auction house had misled them about the size of their earnings; the amount was \$70 million.

In October 2000, the former chief executive officer of Sotheby's, Diana D. Brooks, pled guilty to conspiracy to violate antitrust laws and agreed to cooperate with federal authorities. In May 2001, a federal grand jury indicted both former chairmen, A. Alfred Taubman of Sotheby's and Sir Anthony J. Tennant of Christie's, for collusion. Mr. Taubman was found guilty by a jury in December 2001. He was sentenced to 1 year and 1 day in prison and fined \$7.5 million in April 2002. Mr. Tennant did not stand trial; he did not leave Britain, his home, and the United States cannot extradite

⁴N. Millman, "\$100 Million Fine in ADM Guilty Plea," *Chicago Tribune*, October 16, 1996; G. Burns, "Three ADM Execs Found Guilty," *Chicago Tribune*, September 17, 1998; S. Kilman, "Jury Convicts Ex-Executives in ADM Case," *Wall Street Journal*, September 18, 1998; S. Kilman, "Ex-Officials of ADM Given 2 Years in Jail," *Wall Street Journal*, July 12, 1999; Julie Forster, "A Different Kind of Andreas at ADM," *BusinessWeek*, July 9, 2001, pp. 62–64.

⁵Kurt Eichenwald, "Archer Daniels Settles Price-Fixing Case," *New York Times*, June 19, 2004.

him under British law. Ms. Brooks received no prison term, but was sentenced to home detention, probation, community service, and a fine of \$350,000.⁶

In 1999, two large drug companies, Roche Holding AG and BASF AG, pled guilty in a case brought by the U.S. Justice Department and were fined \$500 million and \$225 million, respectively. The fine incurred by Roche was reported to be the largest ever imposed in an antitrust case. The two companies were accused of fixing the prices of a large array of vitamins. In addition, a Roche executive was fined \$100,000 and sentenced to 4 months in prison. In succeeding years, the companies also paid another \$1 billion in European fines and more than \$1 billion to settle civil lawsuits.⁷

A dynamic random access memory (DRAM) price-fixing conspiracy composed of Korean, Japanese, and German companies was fined more than \$730 million in a U.S. court. The companies raised and stabilized prices of DRAM sold to certain original equipment manufacturers (OEMs) in the United States from 2001 to 2002. The fines were levied from 2004 to 2007. Several executives of Samsung Electronics Co., Ltd. were given prison sentences and fines.⁸

In a very recent case, a San Francisco jury judged AU Optronics Corporation of Taiwan guilty of fixing prices on liquid crystal display (LCD) panels during the period from 2001 to 2006. These panels are used in computer monitors, laptops, televisions, and cell phones. The company's U.S. subsidiary as well as two high-ranking executives were also found guilty. The U.S. Justice Department said that the fine could amount to \$1 billion. At the time of this writing, the case may be appealed. Seven other Asian companies had already pleaded guilty and agreed to pay over \$890 million in fines rather than go to trial. Apparently several large U.S. companies were affected by this cartel action; among them were Apple Inc., Dell Inc., and Hewlett-Packard Co.⁹

In another recent case, in January 2012, two Japanese companies, Yazaki Corporation and Denso Corporation, agreed to plead guilty for conspiring to fix prices of automobile parts, such as wire harnesses used in wiring and circuit board. The fines for the two companies were set at \$470 million and \$78 million, respectively. Four Yazaki executives pled guilty. Earlier, in the same case, Furukawa Electric Company had agreed to pay a \$200 million fine for price fixing, while three of its executives agreed to plead guilty. The price-fixing had gone on for about ten years, from 2000 to 2010.¹⁰

⁶Stories about this case appeared frequently in *Wall Street Journal* during the period 2000 to 2001. A few among these were Laurie P. Cohen and Alexandra Peers, "Christie's Davidge Contacted Sotheby's on Pricing," March 20, 2000; Kathy Kranhold, "Judge Backs Guilty Plea from Sotheby's," February 5, 2001; Kathy Kranhold, "Sotheby's \$70 Million Settlement of Shareholder Suit Is Approved," February 20, 2001; Kathryn Kranhold, "Sotheby's, Christie's Settlement Is Approved by Judge in New York," February 23, 2001; Kathryn Kranhold, "Ex-Chiefs of Sotheby's, Christie's Indicted," May 3, 2001; Kathryn Kranhold, "Sotheby's Chief Is Convicted of Price-Fixing," December 6, 2001; Jerry Markon, "Sotheby's Taubman Is Sentenced to Jail Time, Fined \$7.5 Million," April 23, 2002. Also see Devin Leonard, "First: Feud? What Feud?" *Fortune Magazine*, March 20, 2000, p. 36, and Joshua Chaffin, "Ex-Sotheby's Chief Avoids Prison for Price-fixing Role," *Financial Times*, April 30, 2002. A book published in 2004 tells the story of this case: Christopher Mason, *The Art of the Steal: Inside the Sotheby's-Christie's Auction Scandal*, New York, Putnam, 2004.

⁷"Roche and BASF Fined Heavily for Vitamin Price-Fixing in the USA," *Nutraceuticals International*, June 1, 1999; Edward Iwata, "High Court Rebuffs Foreign Price-Fixing Suits," *USA Today*, June 15, 2004; John R. Wilke, "Price-Fixing Investigations Sweep Chemical Industries," *Wall Street Journal*, June 22, 2004.

⁸Colleen Taylor, "Sixth Samsung Exec Headed for Jail for DRAM Price-Fixing," *Electronic News*, Vol. 53, Issue 18, April 30, 2007, p. 8.

⁹Brent Kendall, "Jury Finds Au Optronics Guilty of Price Fixing," *Wall Street Journal*, March 14, 2012.

¹⁰"Yazaki, Denso to Plead Guilty in U.S. Price-Fixing," Reuters News, January 31, 2012; Matthew Dolan, "Furukawa Electric to Plead Guilty in US to Price Fixing, Pay \$200M," *Dow Jones News Service*, September 29, 2011.

PRICE LEADERSHIP

When collusive arrangements are not easily achieved, another type of pricing practice may occur under oligopolistic market conditions. This is the practice of **price leadership**, in which there is no formal or tacit agreement among the oligopolists to keep prices at the same level or change them by the same amount. However, when a price movement is initiated by one of the firms, others will follow. Examples of such practices abound. You may have observed that at two or more gasoline stations at the same intersection, prices for each grade of gasoline are either identical or almost the same most of the time. Another example is automobile companies, which in recent years have come up with rebate programs. Surely you have seen advertisements offering “\$1,000 cash or 3.9 percent financing.” One company is usually the first to announce such a program; the others follow in short order. Another case is IBM. For many years, in the 1950s and 1960s, IBM was considered to be the price leader in the computer industry. In fact, IBM’s prices were considered to form an “umbrella” for industry pricing. It was said that IBM would establish a price, and because it was the most powerful and preferred manufacturer and thus could command a higher price (an umbrella over the others), its competitors would tend to set their prices at some slightly lower level for similar equipment.

We just described two major variants of the price leadership phenomenon: barometric and dominant price leadership.

Barometric Price Leadership

There may not be a firm that dominates all the others and sets the price each time. One firm in the industry—and it does not always have to be the same one—will initiate a price change in response to economic conditions, and the other firms may or may not follow the leader. If the **barometric price leadership** model has misjudged the economic forces, the other companies may not change their prices or may effect changes of a different, possibly lesser, magnitude. If the firm has correctly gauged the sentiment of the industry, all the firms will settle in comfortably at the new price level. But if this does not happen, the price leader may have to retract the price change, or a series of iterations may be set in motion until a new price level, agreeable to all, is reached. Such a pattern of price changes has been observed in many industries, including automobiles, steel, and paper.

In more recent years, the airline industry has furnished several examples of price leadership that was not followed. An almost bizarre example occurred in August 1998. First, Delta Air Lines and American Airlines raised leisure fares by 4 percent. When Northwest Airlines refused to match the increase, it was rescinded. A few days later, Northwest raised its fares and was matched by the others. Two days later Northwest rescinded the increase, and within a day other airlines followed. Then Northwest raised some of its fares again, only to pull some of them back, and actually decreased leisure fares in some of its markets. Other airlines then realigned their fares with those of Northwest.¹¹

A similar case occurred more recently. Due to the increases in fuel costs in the summer of 2004, American Airlines announced that it was increasing prices on domestic flights by \$5 on one-way trips and \$10 on roundtrips. Some airlines went along with this increase. However, some low-cost airlines, including Southwest and JetBlue, refused to increase their fares. A day later, American and the other airlines retracted the increases.¹²

¹¹This sequence of fare changes was reported in several issues of *Wall Street Journal* between the beginning and end of August 1998.

¹²“American Air Abandons Proposed Fare Increases,” *Wall Street Journal*, August 3, 2004.

For years, Wal-Mart Stores, Inc. has been the price leader by offering customers “everyday low prices.” But since the recent recession, surveys by retail analysts have found that “Wal-Mart’s aura of price leadership has faded. . . .” Apparently consumers who have searched for lower prices have been able to get them at Wal-Mart’s competitors, such as Dollar General Corporation, Aidi Inc., and Amazon. This is despite that various surveys have shown that Wal-Mart’s prices are still lower than those of competitors such as Target. But a recent survey by Morgan Stanley found that 60 percent of Wal-Mart shoppers did not think that Wal-Mart has the lowest prices. “Wal-Mart has responded by returning to its roots. It is cutting the number of ‘rollbacks,’ or planned sales specials, after some customers concluded Wal-Mart was becoming as gimmicky as competitors. Wal-Mart also began restoring thousands of products it had removed from stores as it sought to tidy up cluttered aisles.” Still, retail analysts are questioning whether the company can regain its market share.¹³

Dominant Price Leadership

When an industry contains one company distinguished by its size and economic power relative to other firms, the **dominant price leadership** model emerges. The dominant company may well be the most efficient (i.e., lowest-cost) firm. It could, under certain circumstances, force its smaller competitors out of business by undercutting their prices, or it could buy them out on favorable terms. But such action could lead to an investigation and eventual suit by the U.S. Department of Justice under the Sherman Anti-Trust Act. To avoid such difficulties, the dominant company may actually act as a monopolist, setting its price at the point where it will maximize its profits, and it will permit the smaller companies to continue to exist and sell as much as they want at the price set by the leader. The theoretical explanation of the dominant price leadership model is quite straightforward and is presented in all microeconomics textbooks. We follow its development in Figure 10.2.

The demand curve for the entire industry is D_T . The marginal cost curve of the dominant firm is MC_D , and the sum of all the marginal cost curves of the follower firms

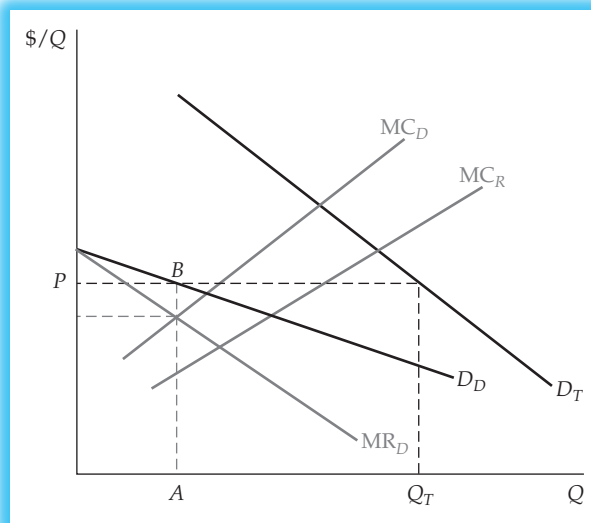


Figure 10.2 Dominant Price Leadership

¹³Miguel Bustillo, “Wal-Mart Loses Edge; Perception That Retailer No Longer Has Best Prices Undercuts Sales Turnaround,” *Wall Street Journal*, August 16, 2011.

is represented by MC_R . The demand curve for the leader, D_D , is derived by subtracting at each point the marginal cost curve of the followers from the total demand curve, D_T . The reason is that if the small firms supply the product along their combined marginal cost curve, MC_B , then the dominant firm will be left with product demand shown along D_D . When the leader's marginal revenue curve, MR_D , is drawn in the usual manner, the leader can establish its profit-maximizing output A and its price at point B . This price is then accepted by the smaller firms in the industry, which will supply the rest of the market at this price. The followers are thus actually faced by a horizontal demand curve at price P .

Such an arrangement is satisfactory to the dominant firm. It maximizes profits and at the same time permits the small firms to exist, thus possibly avoiding legal action. In contrast, the followers will be able to assure themselves of a piece of the market without inviting the possibility of a price war, which they would most likely lose.

As in the case of cartels, dominant price leadership arrangements tend to break down. As markets grow, new firms enter the industry and decrease the interdependence among the firms. Technological changes may bring changes in pricing, and in the long run the leadership of the dominant firm is likely to erode.

REVENUE MAXIMIZATION



Module 10A Another model of oligopolistic behavior was developed some years ago by the American economist William Baumol.¹⁴ Ignoring interdependence, the **Baumol model** suggests that a firm's primary objective, rather than profit maximization, can be the maximization of revenue, subject to satisfying a specific level of profits. He gives several reasons for this objective, among them (1) a firm will be more competitive when it achieves large size (in terms of revenue), and (2) management remuneration may be more closely related to revenue than to profits.

This situation is depicted in Figure 10.3. The figure shows three solid curves. The total revenue curve is the usual one for a firm in imperfect competition, with revenue

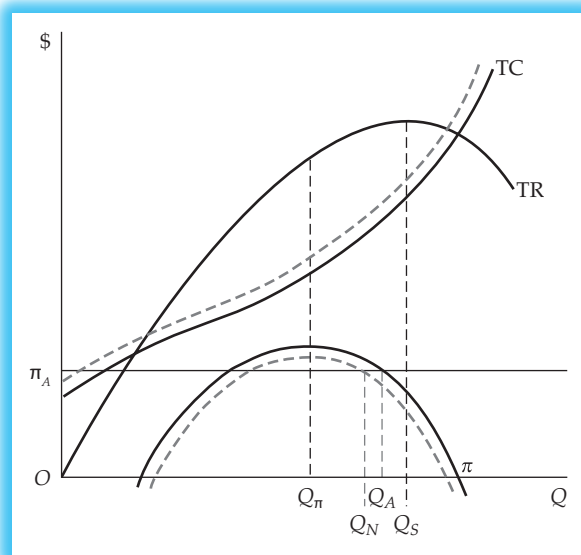


Figure 10.3 Revenue Versus Profit Maximization

¹⁴William J. Baumol, *Economic Theory and Operations Analysis*, 3d ed., Englewood Cliffs, NJ: Prentice Hall International Editions, 1972, chapter 13.

increasing at a decreasing rate because the firm is faced by a downward-sloping demand curve.¹⁵ The total cost curve is also no surprise; it indicates first decreasing and then, at higher production levels, increasing marginal cost. The third line represents profits. It is simply the vertical difference between the revenue and cost lines.

If the firm were a profit maximizer (the traditional economic objective), production would take place at point Q_π , where the profit line hits its peak. In contrast, should the firm be a pure revenue maximizer, equilibrium would occur at output Q_S , where the total revenue (or sales) curve reaches its peak. This point, as we learned earlier, occurs where demand elasticity is unity (i.e., marginal revenue equals zero).

However, revenue maximization may be subject to the constraint that an acceptable profit level exists. This profit will tend to be at a lower level than the maximum achievable. Assuming this acceptable level is π_A output will settle at Q_A . This will achieve the highest possible revenue, while satisfying the profit requirement. Thus, total revenue will be higher than would have been attained under conditions of profit maximization, but lower than if pure revenue maximization (without a minimum profit constraint) had been pursued.

An interesting implication of this model is the effect of a change in fixed costs. Recall that under conditions of profit maximization in the short run, a change in fixed costs will have no effect on price or quantity because neither marginal revenue nor marginal cost is impacted, and thus the maximizing requirement of $MR = MC$ will remain the same. However, in the Baumol model, a rise in fixed costs will raise the cost curve and decrease the profit line. Both new lines will be parallel to the old ones. The two broken lines in Figure 10.3 represent this shift. As can be seen, the existence of the profit constraint will cause output to decrease to Q_N . At this lower output, price will be higher.

Baumol's model is an interesting attempt to present an alternative to the traditional maximization hypothesis. Because his model has not been extensively tested, it is difficult to assess its validity. There have been some empirical studies investigating the relationship between executive pay and revenue (as opposed to profits), but no definitive verdict has been obtained. Some of the studies found a more solid relationship between executive pay and revenue, and others appeared to favor a relationship with profits. Still others arrived at ambiguous answers. One important question remains: Are corporate owners (stockholders) more concerned—and thus determine the market value of a corporation—with revenue or profitability? In the long run, the most likely answer is the latter. Thus it is doubtful that Baumol's model, although possibly applicable to some corporate behavior in the short run, will ever replace the traditional profit-maximizing objective. Another model, growth-maximization, has emerged but it will not be discussed here except to note that the difference between profit-maximization and growth-maximization is the level of retained earnings chosen by the firm.

PRICE DISCRIMINATION

Up to this point, we assume a firm will sell identical products at the same price in all markets. (When the term *identical* is used in this context, it implies that the costs of producing and delivering the product are the same.) But such is not always the case. When a company sells identical products in two or more markets, it may charge different prices in the markets. Such a practice is usually referred to as **price discrimination**. The word *discrimination* here is not used in a normative sense; there is no judgment being made about whether this practice is good or bad. (The term *differential pricing* could be used instead, but the former term has become part of the everyday language of the economist.)

¹⁵Thus each quantity represents a different price.

Price discrimination means one of the following:

1. Products with identical costs are sold in different markets at different prices.
2. The ratio of price to marginal cost differs for similar products.

The practice of price discrimination is not an isolated event. It occurs in many familiar situations. Later in this section, we cite a number of common examples. Here we mention just two to illustrate each general instance just listed. In the first case, price discrimination exists when an adult and a child are charged different prices for tickets (of the same quality and at the same time) at a movie theater. The latter can be illustrated by the selling of cosmetic items, identical except for names on the labels and the quality of packaging, for vastly different prices at department or specialty stores on the one hand and drug and discount stores on the other. The existence of price discrimination is caused by differing demand conditions, not by differences in cost.

Such discriminating price differentials cannot exist under all circumstances. In fact, two conditions are necessary for such a market arrangement:

1. The two or more markets in which the product is sold must be capable of being separated. Specifically, this includes the requirement that there can be no transfer or resale of the product (or service) from one market to the other. That is, there is no leakage among the markets. Only if the markets are sealed off from one another (by natural or contrived means) will the buyers in the various markets be unable to trade the products among these markets. Only in such a case will the seller be able to charge different prices without the price differential being nullified through competition. If the seller incurs costs in creating separate markets, these costs must be less than the additional revenue obtained from discrimination.
2. The demand curves in the segmented markets must have different elasticities at given prices. Without this condition, price discrimination would not be in the firm's best interest.

The reason that companies attempt to engage in price discrimination is that it can enhance profits. From the viewpoint of the consumers of the product, those in the lower-price market may benefit compared with situations where a uniform price is charged. However, consumers in the higher-price market are at a disadvantage.

Economists normally identify three degrees of discrimination. First-degree discrimination is the most profitable for the seller, but it can be achieved only infrequently. Third-degree discrimination, which is not as profitable, is the most commonly observed, and we single it out for brief discussion in the next subsection:

1. First-degree discrimination exists when the seller can identify where each buyer lies on the demand curve and can charge each buyer the price it is willing to pay. Thus, the demand curve actually becomes the marginal revenue curve as faced by the seller. Of course, for the seller—a monopolist¹⁶—to achieve this advantageous position, it must have considerable information on where each buyer can be found on the demand curve, admittedly a herculean amount of market knowledge rarely attained. It is probably almost impossible to find such a pure case in real life, but let us attempt an example. A consumer purchasing a new automobile will generally bargain with the salesperson until they finally agree on a price. If the automobile dealer were clever enough to figure out the highest price that each individual was willing to pay, she could then conclude a deal with each customer at the maximum price (but only if no other dealer offered a lower price). Thus, each price the dealer obtains is on the buyers' demand curve. In reality, automobile dealers (and we must surely be thankful for this) are usually not endowed with such omniscience. We could stretch this example to apply to certain personal services such as medical or legal, where different customers (i.e., patients and clients) could be charged different fees based, for example, on their incomes.

¹⁶The assumption is usually made that a discriminating seller is a monopolist. But price discrimination also commonly exists in markets where competition is not perfect, not just in pure monopoly markets.

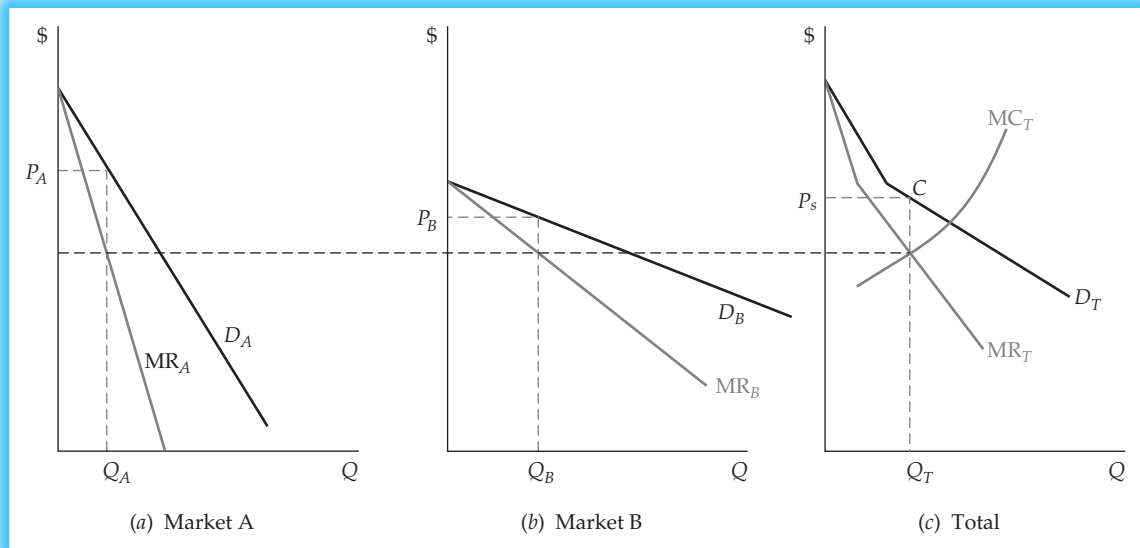
2. Second-degree discrimination, although encountered somewhat more frequently than first degree, also is not commonplace in real life. It involves differential prices charged by blocks of services. An example is the way some public utilities price. They will charge the highest unit price (e.g., per kilowatt of electricity) for small quantities (at the top of the demand curve) and lower prices as the rate of consumption per period increases.¹⁷ Thus, again, only if the monopolist seller has a great deal of information about the demand curve will it be able to roughly “skim” the curve and exact higher revenues from its customer set. To be able to engage in second-degree discrimination, a firm must be able to meter the services consumed by the buyers.¹⁸
3. Third-degree discrimination is by far the most frequently encountered. In this case, the monopolist segregates the customers into different markets and charges different prices in each. Such market segmentation can be based on geography, age, gender, product use, or income, for example.

Third-Degree Discrimination

If the firm can segment markets successfully, it can increase its profits above what they would be if a single price were charged. We show the pricing results with graphs, and subsequently we show a numerical example comparing the profitability of differential pricing versus uniform pricing, as well as a mathematical solution. As shown in Figure 10.4, the company operates in two markets, A and B. In Figure 10.4*a* and *b*, it can be seen that A’s demand curve is less elastic than B’s. Figure 10.4*c* shows the horizontal summation of both demand and marginal revenue curves to represent the company’s total market.

Because we assume the products sold in the two markets are homogeneous, we can draw a marginal cost curve for the firm as a whole, as shown in Figure 10.4*c*. Output will take place at the point where $MR_T = MC_T$. Had a single price been charged,

Figure 10.4 Third-Degree Price Discrimination



¹⁷The price differential here is not related to pricing at peak versus off-peak periods. Such pricing may be related to the cost of producing the service, and therefore, does not represent discrimination.
¹⁸Price discrimination can also be achieved by using a “two-part tariff,” that is, charging a lump-sum fee and then charging for usage. Thus, buyers with greater demand will pay more. Examples of such charges would be a base fee charged by water and telephone companies, and entry to theme parks.

it would have been P_s at point C on the aggregate demand curve. However, the firm can increase its profit by differentiating prices between the two markets. By drawing a horizontal line from the $MR_T = MC_T$ intersection across the graphs for the two separate markets, we can allocate total production to the two markets. For each market, this will be the point where the horizontal line intersects the marginal revenue curve. Marginal revenue is thus the same for both markets. The price charged in each market can be found by drawing a vertical line at the corresponding quantity to the demand curve. Price will be considerably higher in market A, whose demand elasticity is lower.

A numerical example will illustrate third-degree discrimination.¹⁹ Section A of Table 10.1 presents the demand schedule for two markets, as well as the combined schedule for the entire market. Assume fixed costs are \$12,000 per period, and, unlike in Figure 10.4, to simplify the analysis we assume constant per-unit cost of $AVC = MC = \$3$.

If the company were to sell at a uniform price in both markets, it would maximize its profits at a price of \$18. At that point, its profit would be \$10,500. This can be seen in section B of Table 10.1. But if our company can separate the two markets, it can increase its total profit, as shown in sections C and D of the table. If it charges \$24 per unit in market A and \$12 per unit in market B, its profits will be \$12,900 and \$300, respectively. Thus, it will be able to increase its profit by \$2,700.²⁰

Had the company been able to carry out first-degree discrimination and sold to all potential customers at the prices they were willing to pay (with the exception of the last 500 units, which would not have been produced), its profits would have risen to \$25,500, as shown in Table 10.2, assuming that all sales occur at these discrete price points.



A Mathematical Solution for Third-Degree Discrimination

We discuss briefly a simple method to solve for prices and quantities in the presence of third-degree discrimination.

1. Assume there are two markets, A and B, and the demand curves are straight lines, that is,

$$Q_A = a_A - b_A P_A \text{ and } Q_B = a_B - b_B P_B$$

2. Reverse these equations so that P is the dependent variable, thereby obtaining inverse demand.

$$P_A = \frac{a_A}{b_A} - \frac{Q_A}{b_A} \text{ and } P_B = \frac{a_B}{b_B} - \frac{Q_B}{b_B}$$

3. Now calculate total revenue by multiplying by Q :

$$TR_A = \frac{a_A Q_A}{b_A} - \frac{Q_A^2}{b_A} \text{ and } TR_B = \frac{a_B Q_B}{b_B} - \frac{Q_B^2}{b_B}$$

4. Calculate the first derivative of total revenue to obtain marginal revenue:

$$MR_A = \frac{a_A}{b_A} - \frac{2Q_A}{b_A} \text{ and } MR_B = \frac{a_B}{b_B} - \frac{2Q_B}{b_B}$$

¹⁹We have chosen to show a numerical example of price discrimination rather than a mathematical proof. The latter would have been more precise. However, the simpler numerical illustration should be more useful. A brief outline of a mathematical solution is presented in the next subsection.

²⁰We have arbitrarily divided the fixed costs equally between the two markets. But this has no effect on the total profits of the company or on the levels of sales in the two markets that will lead to profit maximization. Indeed, we could have omitted the fixed costs altogether, and there would have been no impact, except that the profits would have been \$12,000 higher in both cases. Also, it should be noted that the quantity sold in this case is somewhat higher than in the one-price case. This is an inaccuracy caused by using discrete numbers in our demand schedule. Had this case been solved mathematically, using calculus, this inaccuracy would not have arisen.

Table 10.1 Numerical Example of Third-Degree Discrimination

<i>A. Demand Schedules</i>			
Price	Market A Quantity	Market B Quantity	Total Quantity
\$36	0	0	0
30	475	25	500
24	900	100	1,000
18	1,100	400	1,500
12	1,300	700	2,000
6	1,450	1,050	2,500
0	1,500	1,500	3,000

<i>B. Total Market</i>							
Price	Quantity	Total Revenue	Marginal Revenue	Fixed Cost	Aver. Var. and Marginal Costs	Total Cost	Profit
\$36	0			\$12,000		\$12,000	\$-12,000
30	500	\$15,000	\$ 30	12,000	\$3	13,500	1,500
24	1,000	24,000	18	12,000	3	15,000	9,000
18	1,500	27,000	6	12,000	3	16,500	10,500
12	2,000	24,000	-6	12,000	3	18,000	6,000
6	2,500	15,000	-18	12,000	3	19,500	-4,500
0	3,000	0	-30	12,000	3	21,000	-21,000

<i>C. Market A</i>							
Price	Quantity	Total Revenue	Marginal Revenue	Fixed Cost	Aver. Var. and Marginal Costs	Total Cost	Profit
\$36	0			\$6,000		\$ 6,000	\$-6,000
30	475	\$14,250	\$ 30	6,000	\$3	7,425	6,825
24	900	21,600	17	6,000	3	8,700	12,900
18	1,100	19,800	-9	6,000	3	9,300	10,500
12	1,300	15,600	-21	6,000	3	9,900	5,700
6	1,450	8,700	-46	6,000	3	10,350	-1,650
0	1,500	0	-174	6,000	3	10,500	-10,500

<i>D. Market B</i>							
Price	Quantity	Total Revenue	Marginal Revenue	Fixed Cost	Aver. Var. and Marginal Costs	Total Cost	Profit
\$36	0			\$6,000		\$ 6,000	\$-6,000
30	25	\$ 750	\$ 30	6,000	\$3	6,075	-5,325
24	100	2,400	22	6,000	3	6,300	-3,900
18	400	7,200	16	6,000	3	7,200	0
12	700	8,400	4	6,000	3	8,100	300
6	1,050	6,300	-6	6,000	3	9,150	-2,850
0	1,500	0	-14	6,000	3	10,500	-10,500

Table 10.2 Profits from First-Degree Discrimination

Price	Quantity	Revenue
\$30	500	\$15,000
24	500	12,000
18	500	9,000
12	500	6,000
6	500	3,000
Total revenue		\$45,000
Fixed cost	\$12,000	
Variable cost (2,500 × \$3)	7,500	
Total cost		19,500
Profit		\$25,500

- Now set the marginal revenue equal to the company’s marginal cost, which we assume to be a constant:

$$MR_A = MC \text{ and } MR_B = MC$$

- Substituting for MR_A and MR_B and solving the two equations gives the quantity sold in each market.
- From here it is easy to find the profit contribution for the two markets and for the combination of the two. Remember that because MC is a constant, average variable cost is constant also, so total variable cost can be calculated simply by multiplying $AVC (= MC)$ by the quantity. Fixed cost is then subtracted.
- If we want to find out what the price and profit would be if a uniform price were charged, we first add the two demand functions found in Step 1. We then reverse the resulting equation in terms of price, and obtain inverse total demand, $P(Q)$. From here, multiply by Q to obtain $TR(Q)$ and take the derivative to obtain marginal revenue. Marginal revenue is then equated to MC . Solving for Q , we obtain total quantity, Q_T . Obtain the uniform price, P_S , using inverse total demand $P_S = P(Q_T)$. Quantity in each market is obtained using the demand equations in Step 1, evaluated at the common price P_S . Output will increase in the more inelastic market (due to the price reduction in that segment) and decrease in the more elastic market (due to the price increase in that segment). Total quantity sold will be the same as if discrimination existed, but the profit will be lower.

Examples of Price Discrimination

Price discrimination is an extremely common practice encountered in all types of situations. A number of rather common examples follow:

- In the past, physicians often set their fees in accordance with patient income. In a way, it could be argued that such a fee arrangement was quite equitable: Those who can afford to pay higher prices will do so. However, as stated before, we are not concerned here with the normative aspects of differential pricing. The result of such a practice will still be an increase in the physician’s income.

Presently, medical price discrimination exists in a somewhat different guise. Physicians frequently charge a patient who has health insurance more for the same services than they charge a patient who does not. The difference cannot be explained by the cost incurred by the physician in filing insurance documents. However, the two conditions necessary for differential pricing exist. The elasticities in the two markets (uninsured and insured) are certainly different, and the markets are sealed from each other (either the patient does or does not have insurance).

- Often, products going into the export market will be priced lower than those sold domestically. A major reason for the differential is that international competition is

stronger than that faced by the firm in its (frequently sheltered) domestic markets. Thus demand curves in international markets are more elastic.

3. Many pubs and bars have “ladies’ hours,” and in the past, major league baseball parks had ladies’ days on Wednesdays. In both cases, the price for women is lower than that charged their male counterparts.
4. Theaters, cinemas, and sports events often charge lower prices for children occupying equal accommodations as adults. The same arrangement is frequently offered to senior citizens.
5. Public transportation systems commonly offer reduced fares to senior citizens.
6. State universities charge higher tuition fees to out-of-state students, although there is no cost differential to the university between in-state and out-of-state students.
7. Public utilities (electric, gas, telephone) customarily charge higher rates to business customers than to residential customers.
8. University bookstores offer 10 to 15 percent discounts to faculty, while charging full prices to students.
9. Individuals can order publications from publishers at lower prices than those charged to libraries and other institutions. Most professional journals are priced in this way.
10. Advancing technology is creating new avenues for practicing price discrimination, as shown in the following two examples:
 - a. Last-minute airline ticket buyers can make special deals through the Web site priceline.com.
 - b. Software permits Web-based merchants to identify individual visitors to their Web sites and study their shopping behavior. Price-sensitive customers may be offered lower prices. This new approach is a variation on catalog retailers’ policies of sending catalogs with different prices to different zip codes.²¹

We leave it to you to think of additional examples.²² In all these instances, different prices are charged at the same time. A senior citizen and a person 40 years of age traveling in the subway together will pay different fares. Differential utility charges to business and residential customers are in effect at the same time during the day.

However, there are also price differences that depend on when products or services are consumed:

1. Theaters charge different ticket prices for matinees and evening performances.
2. Theaters charge higher ticket prices on weekends than on weekdays.
3. Daytime telephone rates are higher than nighttime rates.
4. Hotels catering mostly to business travelers charge lower room rates during weekends.

Are these also examples of price discrimination? Many economics texts subscribe to this notion. (Of course, where a firm’s costs differ at different times, no price discrimination would be claimed.) However, if these are types of discrimination, they do not really belong with the original examples given. After all, movements of a demand curve over time will change the prices of many products without the presence of price discrimination. What appears to be the case here is that weekend theater ticket

²¹S. Woolley, “I Got It Cheaper Than You,” *Forbes*, November 2, 1998, pp. 82–84.

²²Fare differentials in the airline industry are often mentioned as a prime example of price discrimination. A ticket bought well in advance, whose price is not refundable and which includes a weekend, is usually considerably cheaper than one that does not possess these characteristics. Vacation (elastic demand) and business (inelastic demand) travel are usually differentiated in this manner, and thus it may not be unusual for two travelers to sit next to one another on a plane and be paying vastly different fares. However, there are also differences in cost and risk to the airline between these two different fares, so the two tickets are really not the same. If the fare differentials occur due to disparities in cost and risk, then this is not an example of price discrimination.

demand, for example, is considerably higher than demand on a Tuesday or a Wednesday, while the supply curve is essentially vertical. Prices change as demand changes. Whether this qualifies as price discrimination is questionable.

Some Examples of Price Discrimination Practices

It was 1995 when California passed a law banning price discrimination based on gender. The legislation was aimed at retail stores, such as dry cleaners and hair salons, which charged higher prices to women than to men for presumably similar services. However, price differentials still appeared to exist. So in 2001, the California legislature passed a law to further discourage price differentials. In addition to increasing monetary penalties, stores are now required to display prices charged for their fifteen most commonly used services. However, differentials due to differences in cost (due to time required or difficulty of service) are still permitted.²³

Several years later, Lavalife, an online dating service, agreed to pay over \$700,000 to male customers in California because it charged male customers for their services while it allowed female customers to have free access.²⁴ And more recently, the California Supreme Court permitted four men to bring a suit against a Los Angeles nightclub. The four men claimed that the club charged them a \$20 admission charge while women had to pay only \$15.²⁵

In March 2000, McCormick & Co., the leading marketer of spices in the United States, entered into a consent decree with the Federal Trade Commission for violating the Robinson-Patman Act.²⁶ It was alleged that McCormick charged higher prices to some retail stores than others in return for obtaining greater shelf space at the expense of smaller stores. McCormick commonly demanded as much as 90 percent of a store's shelf space. This practice is usually referred to as "slotting fees," and is a standard way of doing business in grocery stores and some other industries. Thus, this decision could have much wider implications for the future. The consent order did not impose any monetary penalties but is effective for 20 years.²⁷

Gasoline wholesalers in several areas employ a method called "zone pricing." For instance, in the Phoenix, Arizona, metropolitan area there are differentials between, for example, prices in the city of Phoenix and Scottsdale, an affluent suburb. Oil companies do not reveal the details of the zone demarcation. However, it is believed that real estate values (i.e., an indication of income levels) are an important factor. Gas consumption and traffic counts are also potential factors. Another area in which zone pricing has been described is in the Long Island, New York, area. Although it may be a matter of supply and demand, according to Ed Silliere, vice president at Energy Merchant Corporation, "[e]ven the affluence of the community could be a factor, not only because higher real costs mean higher rents and taxes, but because wealthier consumers are deemed less likely to go out of their way to shop for price. . . . You might have a marketplace that can afford to pay more, doesn't mind paying more and is unwilling to drive the distance to get a better price." What he is describing, of course, are differences in demand elasticity.²⁸

²³Jim Sanders, "California Lawmakers Pass Bills to Eliminate Retailer Gender Discrimination," *Sacramento Bee*, September 30, 2001.

²⁴Robert Benzie, "Dating Services Hire Top Lobbyist to Fight Gender Pricing Bill," *Toronto Star*, March 13, 2006.

²⁵Court: Nightclub Can Be Sued for Price Discrimination," *Oakland Tribune*, May 31, 2007.

²⁶The Robinson-Patman is discussed briefly in Chapter 14. This law makes it illegal to charge different prices to different customers, unless justified by cost differentials.

²⁷Ira Teinowitz, "FTC, McCormick Reach Accord on Slotting Fees," *Advertising Age*, March 13, 2000, p. 75; Tom Weir, "FTC Raises the Ante with McCormick Order," *Supermarket Business*, April 15, 2000, pp. 45–46.

²⁸Max Jarman, "One Key to Gas Price Puzzle," *The Arizona Republic*, June 16, 2004; Tom Incantalupo, "Many Factors Cause Varying Gasoline Prices in Melville, N.Y., Area," *Newsday*, May 20, 2004.

Toyota Corporation settled several class-action suits alleging price discrimination against Hispanic and African American car buyers. In addition to paying more than \$10 million in attorney fees, Toyota also agreed to grant preapproved credit offers to these car buyers, contribute to nonprofit organizations focusing on educating minority consumers, and pay damages to class members.²⁹ This case illustrates the two necessary conditions for the existence of price discrimination—market segmentation and differences in demand price elasticity.

In 2008, the Nevada Equal Rights Commission ruled that a Las Vegas health club that charged men higher fees than women discriminated against men.³⁰ The European Court of Justice ruled that charging women lower insurance premiums than men “breached EU rules on sex discrimination.” Because of this judgment, insurance companies will not be able to charge different premiums based on gender on products such as car insurance, private healthcare insurance, and pensions.³¹

Pricing in the Hotel Industry: Example of Price Discrimination

Differential room rates have existed in the hotel industry for a long time. However, we do not discuss in this section, for instance, the lower weekend rates that are standard at business hotels or the higher rates charged during the busy season by resort hotels. These intertemporal differences appear to be based on different levels of demand. Here we are concerned with different rates charged to different customers at the same time (e.g., same day). In the following discussion, we must remember an important fact attending the operation of a hotel: a large proportion of the cost of running a hotel is fixed. The variable cost of renting an empty room is relatively small. “To paraphrase one general manager, ‘If I’ve got a warm body with money standing in front of me and cold sheets upstairs, I want to make a deal. As long as the customer is willing to pay more than my variable cost to clean that room, I’m going to make money.’”³² Each additional rented room represents incremental revenue.

This situation leads to different types of price discrimination practices in the hotel industry.

Often a hotel will have several different rates, and the actual rate charged to a particular guest will depend on the bargaining skill of the customer and the knowledge of the innkeeper in estimating the highest price the potential patron is willing to pay. Thus, the revenue of the hotel may be found along the demand curve (if the manager is really proficient at estimating the customer’s willingness to pay), and the result may approach first-degree discrimination. Figure 10.5 illustrates this situation with potential scenarios for three different prices.

However, the more typical way of discriminating is to segment the market. The simplest method is to separate leisure travelers from business travelers. The demand of the former is certainly more price elastic because the room price is an important part of the total vacation expense. At the same time, vacation travelers may be willing to make an advance commitment, may stay for a longer time period, and are more flexible in their arrangements. The business travelers’ demand is less price elastic; they are much more inflexible in making their arrangements (they must be in a certain place at a certain time); their commitments often cannot be made far in

²⁹Jared A. Favole, “Toyota Motor Credit Settles Price-Discrimination Lawsuit,” *Dow Jones Newswires*, February 6, 2007.

³⁰Steve Friess, “Lower Rates for Women Are Ruled Unfair,” *New York Times*, August 13, 2008.

³¹Jane Croft, “Court Outlaws Use of Gender to Assess Risk,” *Financial Times*, March 1, 2011.

³²Richard D. Hanks, Robert G. Cross, and R. Paul Noland, “Discounting in the Hotel Industry: A New Approach,” *The Cornell H.R.A. Quarterly*, February 1992, pp. 15–23. The discussion in this section is based largely on this article. The quote and all that follow also come from this article.

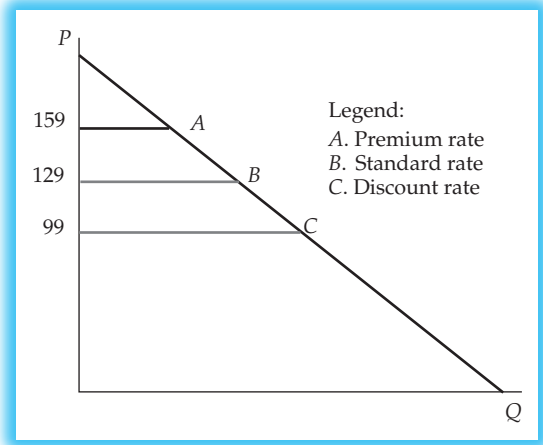


Figure 10.5 First-Degree Discrimination in the Hotel Industry

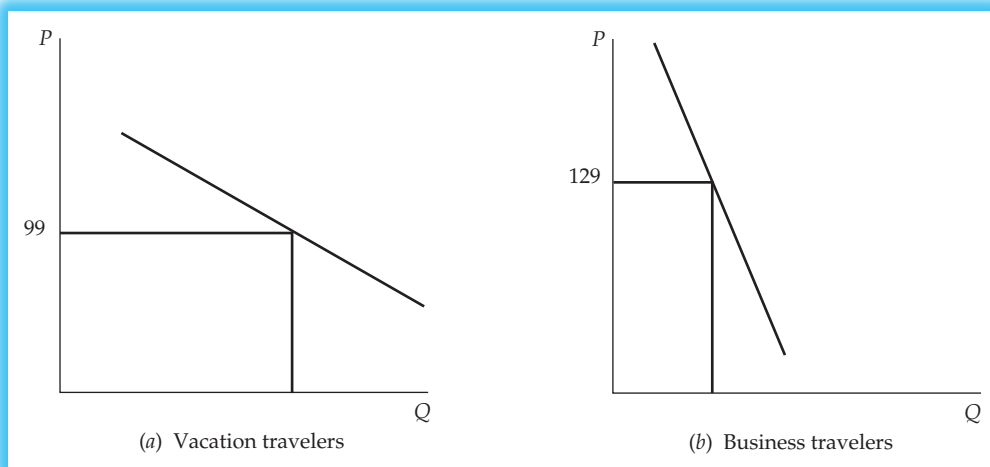
advance; perhaps most important is that they are in most cases not willing to stay over the weekend, when rates are often lowest. Figure 10.6 shows the two demand curves with possible prices.

“As customers have become more sophisticated in manipulating the current pricing system, hotels will eventually be forced to modify their pricing structure.” A new approach being introduced will “fence” customers into different rate categories tailored to specific needs so higher-rate patrons will not be able to “trade down.” Among the various methods to segment the customer sets will be the following:

1. Advance reservations and advance purchase
2. Rates differentiated depending on how many days in advance the reservations are made
3. Refundability
4. Flexibility to make changes in arrangements
5. Required time of stay

Hotels will be able to combine and alter these requirements (“fences”) depending on the level of demand at any specific time. With these procedures, hotels will be able to penetrate their markets more deeply. However, such a multifaceted program

Figure 10.6 Third-Degree Discrimination in the Hotel Industry



also has its problems. First, to carry out this program effectively, hotels must have sophisticated computer-directed reservation systems and must be able to make rapid adjustments to ensure they do not sell too many rooms at unnecessarily discounted rates. The administration of such a system may require additional staffing and additional costs, and management must be sure the additional revenue derived from this system more than compensates for this cost.

Tying Arrangements: A Possible Extension of Price Discrimination

A *tying arrangement* (frequently referred to as a *tie-in sale*) exists when a buyer of one product is obligated to also buy a related (usually complementary) product from the same supplier. U.S. antitrust laws take a very dim view of tying arrangements, and a large number of court cases have declared this practice to be illegal. The reasoning behind this is that a firm will use its existing market power in the first (tying) product to suppress previously existing competition in the second (tied) product. The firm will thus extend its monopoly power.

Typically a tying arrangement will involve a major, often durable product and a low-value item that is usually complementary to the former.³³

The legal argument claiming that tying arrangements cause an extension of monopoly appears flawed. If a company already has market power in the tying product, it can maximize its monopoly profit. To tie another product to it may actually dilute its power, in that customers who buy the two complementary products are mainly interested in the total cost of the two products combined. Thus, if a tie-in sale arrangement causes the price of the tied product to increase, it may have to be made up by a decrease in the price of the tying product.

An alternative economic explanation of a tying arrangement is that it is a type of third-degree discrimination. A company may charge identical prices for the tying good. However, purchasers of the tying product may have different levels of demand for its usage. This quantity of demand can be “metered” by the tied product; purchasers who use the product with greater intensity will use a larger quantity of the tied product. Thus, the seller may charge a relatively reasonable price for the major product and collect monopoly profits on the lower-valued complementary product.

One of the more prominent cases in the history of this litigation was the U.S. government suit against IBM.³⁴ IBM rented its equipment on a monthly basis. All customers paid the same monthly rental whether they used the machine 8 hours a day or 2 hours per week. Those customers who used the equipment with greater intensity had to use a larger number of punch cards. Thus IBM really had a relatively inexpensive way to separate its customers into those who used the rented machines more intensively and those who used them only sporadically. It could practice this discrimination even while charging the same unit price for punch cards to all customers. “It is generally agreed that if the tying and tied goods are complementary in demand, then profit maximization under the price discrimination hypothesis will lead the seller to lower the price of the tying good (below the level that would prevail if the good were sold separately) and sell the tied good at a price above its marginal production cost.”³⁵

³³To illustrate, some of the court cases involving tying arrangements have included riveting machines and rivets, computers and punch cards, camera film and film processing, and condominiums and building management services.

³⁴*International Business Machines Corporation v. United States*, 298 U.S. 392 (1936).

³⁵Meyer L. Burstein, “The Economics of Tie-In Sales,” *Review of Economics and Statistics*, 42 (February 1960), p. 69. Much of the preceding discussion is based on William F. Shughart II, *The Organization of Industry*, Homewood, IL: Irwin, 1990, pp. 307–14.

Although price discrimination is a plausible explanation of tying arrangements, other arguments have been advanced:

1. **Quality control.** Firms have argued that tying arrangements are necessary to ensure the integrity of their product so they do not get blamed if an inferior tied good is used.
2. **Efficiencies in distribution.** A total lower cost can be attained if there are savings to the company in delivering both products.
3. **Evasion of price controls.** If there is a ceiling price on one of the two products, then selling the second product at a higher price will circumvent price control.

Social Welfare Implications of Price Discrimination

As mentioned previously, U.S. antitrust laws look unfavorably at the practice of price discrimination, which is said to lead to a lessening of competition.

Under monopoly conditions, an industry produces a smaller amount of product at a higher price than under competitive conditions. However, as we have seen, under price discrimination of the first and second degree, a company may produce a larger quantity than a single-price monopoly. Given the conditions of first-degree discrimination, the firm will charge prices along the demand curve all the way to the point where demand equals marginal cost. This, of course, results in a larger quantity than produced by a single-price monopoly. A similar situation would hold for second-degree discrimination. Thus, when price discrimination exists, total production may equal that which would exist under competitive circumstances.

In the case of third-degree discrimination, the situation is more ambiguous. As we have seen, as long as the demand curve for the monopolist is a straight downward-sloping line, the quantity produced will be the same under a single-price and a price-discrimination situation. However, under certain circumstances, production under discriminatory pricing may increase. This situation may occur if the demand curve is not a straight line or if a one-price policy would not be profitable enough for a company to produce the product at all.

In the case of third-degree discrimination, customers in the market with lower price elasticity will pay higher prices, while those in an elastic demand market will pay lower prices than under conditions of a one-price monopoly. The implications here are uncertain because it is very difficult to weigh the benefits bestowed on the latter compared with the costs imposed on the former as to their effect on total economic welfare. However, there is no question that the sellers in the discrimination case will benefit from the higher prices charged at least in a part of the market, thereby increasing their profits.

NONMARGINAL PRICING

Throughout this text, we appear to have assumed all businesspeople calculate demand and cost schedules, obtain marginal revenue and marginal cost curves, equate marginal revenue with marginal cost, and thus determine their profit-maximizing selling price and production quantity. But how many business owners or managers actually know how to make these calculations? Even if they have the knowledge, how many have the time and, even more important, sufficient information to make such calculations?

In fact, it is often claimed (as discussed in Chapter 2) that businesses are really not profit maximizers, that they have other objectives. It has been said that

management will seek only satisfactory levels of profit for the owners. The term *satisficing* has been used in this context.³⁶ Other corporate goals may also be important, such as the achievement of a desired market share, a target profit margin (i.e., percent of profits to revenue), or a target rate of return on assets (profit divided by assets) or on equity (profit divided by stockholder equity).

It also appears that one of the most popular pricing methods, believed to be pervasive throughout industry, is the cost-plus or full-cost method, which at first glance seems not to employ the marginal pricing principle. It is this subject that we discuss next.

Cost-Plus Pricing

A researcher questioning a sample of businesspeople on their pricing methods would probably be told by a majority that they simply calculate the variable cost of the product, add to it an allocation for fixed costs, and then add a profit percentage or markup on top of these total costs to arrive at **cost-plus pricing**.³⁷ Thus, for instance, if the direct (variable) cost of a product is \$8, its allocated overhead is \$6, and the desired markup is 25 percent, the price of the product will be \$17.50 ($8 + 6 + 0.25 \times 14$).³⁸

Such a calculation appears to be extremely simple, and the whole method is often described as naive. But this apparent simplicity hides some fairly difficult calculations and assumptions:³⁹

1. How are average variable costs calculated?
2. How are fixed costs allocated? Why are fixed costs included in the price calculation? Economic theory tells us that fixed costs do not affect price.⁴⁰
3. How is the size of the markup determined? Usually, it is said that the markup should guarantee the seller a “fair profit,” or some target profit margin or target rate of return. If this is the case, are demand conditions taken into consideration?

We discuss these problems, and as we go through this analysis, we may find that cost-plus pricing and marginal pricing have a lot in common.

In cost-plus pricing, costs, both variable and fixed overheads, are usually calculated at some standard or normal quantity, as is done by accountants. These are historical costs and do not appear to include an opportunity cost. But economic theory tells us that unit costs tend to vary with quantity, and the expected quantities may not correspond to those that result.⁴¹ Also, as mentioned, fixed costs should not be used in the determination of prices.

However, if we take these criticisms in turn, the shortcomings of cost-plus pricing may not be as serious as they appear. There is no real reason why accounting costs cannot include some measure of opportunity cost. Even if it is not incorporated, a

³⁶See Herbert Simon, “Theories of Decision Making in Economics and Behavioral Science,” *American Economic Review*, 49 (June 1959), pp. 253–83. This condition usually prevails in large corporations, where professional managers may not act in conformance with the wishes of stockholders. In the textbooks on corporate finance, considerable attention is accorded this subject, which has been named “agency theory.” It is based on a famous article by Michael C. Jensen and William H. Meckling, “Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structure,” *Journal of Financial Economics*, October 1976, pp. 305–60.

³⁷One of the original studies was R. L. Hall and C. J. Hitch, “Price Theory and Business Behavior,” *Oxford Economic Papers*, 2 (May 1939), pp. 12–45.

³⁸Markup is ordinarily calculated as a percentage of cost. Profit margin is commonly computed as a percentage of price. Thus, a 25 percent markup is equivalent to a 20 percent profit margin.

³⁹It is said that restaurants usually mark up the food cost four times to arrive at the price of a menu item—a very simple calculation indeed.

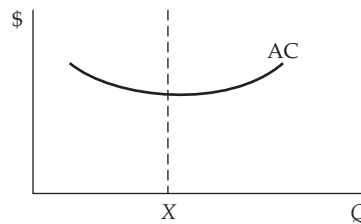
⁴⁰Only in the Baumol revenue maximization, subject to a minimum profit constraint, does fixed cost enter into price determination.

normal profit (another name for opportunity cost) certainly could easily be included in the markup. Now, it is often said that cost-plus pricing is a long-term concept. If that is the case, then, according to economic theory, all costs are variable; a cost allocation is then an estimate of the additional variable costs in the long run. Further, although economists like to draw nice U-shaped cost curves, it is quite possible that in the longer run, the bottom portion of the average cost curve is quite shallow (saucer-shaped) and that over some production range it may appear to be almost horizontal.⁴² In that case, as long as a firm is producing in the range at which standard costs are calculated, the problem of costs varying with quantity is obviated. Also, if the curve is relatively horizontal, marginal cost will be identical or almost identical to the average cost in that interval, and pricing on the basis of average cost will thus be substantially similar to marginal cost pricing. In addition, economic theory tells us that, under perfect competition, in the long run all but normal profits will disappear. The markup, then, must certainly represent normal profit. It is more likely, however, that competition in the real world is not quite perfect, and the firm will therefore be faced with a downward-sloping demand curve.

This brings us to the question of the demand curve. If a markup is applied to obtain a “fair” profit, the implication is that demand conditions are not taken into consideration. But that would indicate almost complete inflexibility regarding the size of the markup. However, it has been observed in innumerable cases that markup percentages differ among different product lines of the same firm. The fact that a company accepts a lower markup for some products than others indicates that demand conditions and the competitive environment are included in the pricing decision making. As is shown, the markup percentage tends to vary inversely with demand elasticity. This makes obvious sense: When a firm faces very strong competition, the demand curve facing it will tend to be nearly or completely horizontal. Under those circumstances, the firm will not be able to afford a very large markup.

One other important point must be made. Not only will different markups be applied for different product lines of a given firm, but the markup will probably change on a given product from time to time. Such changes could be caused by changing demand or changing cost conditions. When this occurs, a firm will adjust its markup and thus its price to meet the new circumstances. Its purpose for such action is most probably to increase or protect its profits. As long as a firm changes its prices to “do better” (i.e., increase its profit or minimize its loss), it is acting as if it has knowledge of its demand and cost curves: that is, it is acting consistently with marginal pricing.

⁴¹The accountant’s cost for a normal quantity can be shown as one point on the economist’s average cost curve:



Point X is the “normal” quantity at which costs would be calculated. But other quantities would be produced at different costs.

⁴²This situation is consistent with constant returns to scale.

It is certainly true that businesspeople do not have sufficient knowledge to estimate marginal revenue and marginal cost curves with any degree of accuracy. Thus, cost-plus pricing can be a substitute for marginal pricing in the absence of sufficient knowledge. But given the proclivity for firms to adjust their markups in response to demand and cost conditions in such a way as to improve profitability, profit maximization and cost-plus pricing can be quite compatible.

Cost-Plus Pricing in Practice

A rather interesting case of cost-plus pricing appeared in a *Wall Street Journal* article. It concerns the pricing practices of the Parker Hannifin Corporation, a manufacturer of industrial parts. “For as long as anyone at the 89-year-old company could recall Parker used the same simple formula to determine prices of its 800,000 parts.” Managers would calculate how much it cost to make and deliver a part, and then add a percentage, much of the time 35 percent.

In 2001 a new chief executive, Donald Washkewicz, was elected. After reviewing the company’s pricing method he decided that prices should not be determined by applying a similar percentage to all products. Rather, prices should be set based on what customers would be willing to pay. (In other words, demand elasticity should be considered.)

With the help of consultants the company classified its products by the amount of competition it faced. It found that a large number of products were priced too low because they were unique or specialized and did not have much competition. The company raised prices on a large number of products an average of 5 percent. There were also some products whose prices were decreased.

Since the introduction of this program, the company has increased its return on invested capital considerably, and the price of its shares has risen faster than the Standard and Poor’s 500 index.⁴³



An Arithmetic Reconciliation of Cost-Plus and Marginal Pricing

It can be shown mathematically that under certain circumstances, cost-plus pricing can be consistent with profit maximization (i.e., $MR = MC$).

Most microeconomics textbooks show the mathematical relationship among price, marginal revenue, and demand elasticity as follows.⁴⁴

$$MR = P \left(1 + \frac{1}{E_p} \right)$$

As profit is maximized when $MR = MC$, we can rewrite the equation as

$$MC = P \left(1 + \frac{1}{E_p} \right)$$

⁴³Timothy Aeppl, “Seeking Perfect Prices, CEO Tears Up the Rules,” *Wall Street Journal*, March 27, 2007.

⁴⁴The equation is derived in the following way: total revenue (TR) equals price times quantity ($P \times Q$). To obtain marginal revenue, total revenue must be differentiated with respect to quantity:

$$MR = \frac{dTR}{dQ} = \frac{d(P \times Q)}{dQ} = P \times \frac{dQ}{dQ} + Q \times \frac{dP}{dQ} = P \times 1 + Q \times \frac{dP}{dQ} = P \left(1 + \frac{Q}{P} \times \frac{dP}{dQ} \right)$$

Note that the product inside the parentheses is the reciprocal of elasticity; hence,

$$MR = P \left(1 + \frac{1}{\epsilon_p} \right)$$

Remember that demand elasticity has a negative sign.

Further, under certain conditions, marginal cost will equal average cost. Thus, our equation becomes

$$AC = P \left(1 + \frac{1}{E_p} \right)$$

and can then be rewritten as

$$AC = P \left(\frac{E_p + 1}{E_p} \right)$$

To show how price is based on average cost, we can rearrange the equation as

$$P = AC \left(\frac{E_p}{E_p + 1} \right)$$

Under conditions of cost-plus pricing,

$$P = AC(1 + M)$$

where M stands for the markup percentage. If the two previous equations are comparable, then

$$(1 + M) = \frac{E_p}{E_p + 1}$$

It can be shown that there is an inverse relationship between markup and demand elasticity. For example, if $E_p = -2$, then $(1 + M) = -2/-1 = 2$ and M is therefore 100 percent. If, however, $E_p = -5$, then $(1 + M) = -5/-4 = 1.25$, and markup is only 25 percent. This result is quite reasonable; it indicates that the less elastic the demand curve, the larger will be the markup.

Thus, under the not infrequent conditions where the average cost curve is constant in the relevant range of production, cost-plus pricing may give results identical to those that would be obtained if managers were pursuing profit maximization.

Incremental Pricing and Costing Analysis

We have just discussed the cost-plus method of pricing; it is considered to be a very popular pricing method. We explained that cost-plus pricing and marginal pricing can generally be reconciled. But there are difficulties with implementing marginal pricing in the real world. To do a good job of marginal pricing, a manager would have to have good estimates of the shape of demand and cost curves. Because it may be quite costly and certainly very difficult to estimate marginal quantities, businesses will often use incremental analysis to achieve the objective of profit maximization.

In a way, marginal and incremental analyses are very similar. Although *marginal* implies that we must estimate the revenue and costs created by one additional unit or the additional revenue obtained from one extra dollar of expenses, incremental analysis deals with changes in total revenue and total costs resulting from a particular decision to change prices, introduce a new product, discontinue an existing product, improve a product, or acquire additional machinery or plant. We discuss the question of incremental costs in Chapter 7.

The important lesson to learn here is that only those revenues and costs that will change due to the decision should be considered. Thus, it is a mistake to include sunk costs in calculating the outcome. Furthermore, if costs such as fixed overhead are already being incurred and will not change (even though accountants may reallocate these costs), they are irrelevant to the decision. In contrast, if a decision results in a

change in revenues or costs of another product (possibly a complementary or substitute good), such an effect must be included in the analysis.

Obviously, this is an important subject to which we are devoting only a small amount of space here. However, incremental analysis is at the heart of the study of long-term investments. This is a subject discussed at greater length in Chapter 12.

MULTIPRODUCT PRICING

In economics, much of the analysis makes use of simplifying assumptions. For example, we know that very few products in our economy are produced under conditions of perfect competition. Nevertheless, a large portion of our text—and all other economics texts—is devoted to its discussion. There are good reasons for this practice. First, perfect competition is the simplest of the economic models and is thus a good starting point for the discussion of more complex systems. Second, many markets, although not perfectly competitive (i.e., firms are faced by downward-sloping demand curves), can be analyzed as such because their behavior resembles perfect competition closely enough. Any predictions based on this analysis will be sufficiently accurate to obviate the need for more complex models.⁴⁵

Another simplification frequently made in economic theory is assuming a firm or a plant produces a single product. Up to this point, we have made the same assumption. Actually, we first assumed single products were sold in single markets, and later extended our analysis to operations in more than one market (price discrimination). Now we provide a brief treatment of **multiproduct pricing**, cases in which a plant or a firm produces two or more products, which are, of course, the norm rather than the exception.

The various products produced by a firm can be independent of one another. This means that neither the demand for nor the cost of one product is affected by the demand for or cost of another product. In such a case, each product will be produced, as usual, at the level where its marginal revenue equals its marginal cost. The analysis can then proceed as if only one good were produced.

In most cases, however, there is some relationship among products produced by one firm. The relationships can exist either on the demand side or the cost side—or both. We can distinguish (at least) four different interrelationships:

1. Products are complements in terms of demand. One company may produce both personal computers and software, or a fast-food restaurant may sell both hamburgers and soft drinks.
2. Products are substitutes in terms of demand: a company may produce different models of a personal computer, or a soft drink company may bottle both cola and lemon-lime soda.
3. Products are joined in production. The extreme case of joint production occurs when two products are produced in fixed (or almost fixed) proportions, such as cattle production, which involves one skin and one carcass per steer.
4. Products compete for resources. If a company making different products that compete for the available resources produces more of one product, it will have to do so at the expense of producing less of other products. The production of different models of the same computer is an example.

Let us now discuss each case.

⁴⁵It is the accuracy of the prediction provided by a model that is important to scientists, not the reasonability or realism of the assumptions. This point has been successfully argued by Milton Friedman in “The Methodology of Positive Economics,” contained in his *Essays in Positive Economics*, Chicago: University of Chicago Press, 1953, pp. 3–43.

Products Complementary in Demand

When two products are complementary, an increase in the quantity sold of one will bring about an increase in the quantity sold of the other. This may be due to an increase in the demand for product A or a price decrease of product A (bringing about an increase in the quantity demanded). Products may be so closely related that they are bought in fixed proportions. An example is kitchen knives, each of which must be made of one wooden handle and one metal blade. Other complementary products are a personal computer and a keyboard, and still another example is an automobile body and a set of four wheels. Somewhat less fixed in proportion but still closely related products are razors and razor blades, tennis rackets and tennis balls, and computers and software. There are also more remotely related products where the demand for one can easily have a beneficial effect on the demand for the other. For instance, a popular textbook in economics published by a particular company may enhance the sales of a finance textbook by the same publisher.

The important point is that the demand for a product is affected not only by its price, by income, and by tastes, for example, but also very strongly by the prices of related commodities. This subject is discussed in Chapter 3, where we define the determinants of the demand curve in general. Here we concentrate on the effects of complementary commodities on the revenues of one firm. Thus, if products A and B are complementary, a change in revenue from A will entail a change in the revenue from B. In both cases, profit maximization will occur at the familiar point where the marginal revenue of each product equals its marginal costs. Because each demand equation will include the prices of both products, the pricing problem will require the solution of simultaneous equations.

If managers had nice, neat demand and cost functions available for each product, they could arrive at the combined profit maximization positions using relatively simple mathematical formulas.⁴⁶ However, because in real life the decision maker would most likely not have sufficient data on hand, the maximization process would proceed along a trial-and-error course, where markups (and thus prices) for the products would be adjusted until the optimal combination is

⁴⁶For the two-product situation, a manager could calculate the marginal revenue for each of the interrelated products x and y given inverse demand curves for those products:

$$P_x = f(x,y) \quad \text{and} \quad P_y = g(x,y)$$

In this instance

$$TR_x(x,y) = xf(x,y) \quad \text{and} \quad TR_y(x,y) = yg(x,y)$$

Total revenue for the firm is

$$TR(x,y) = TR_x(x,y) + TR_y(x,y)$$

Marginal revenue for each product is then given by:

$$MR_x = \frac{\partial TR(x,y)}{\partial x} = \frac{\partial TR_x}{\partial x} + \frac{\partial TR_y}{\partial x} \quad \text{and} \quad MR_y = \frac{\partial TR(x,y)}{\partial y} = \frac{\partial TR_x}{\partial y} + \frac{\partial TR_y}{\partial y}$$

Each marginal revenue is equated to its respective marginal cost:

$$MR_x = MC_x \quad \text{and} \quad MR_y = MC_y$$

These two equalities are solved simultaneously to find the profit-maximizing production bundle of products (x^*, y^*) .

reached. Actually, the process would be even more complex in reality because it is not only the complementary relationship between the firm's two products that has an important influence on the firm's revenue (and profit); competitors' products that are substitutes for our firm's products must also be considered in the process of price setting.

There is another instance in which a company must consider these interrelationships. It is not necessary that a firm produce two related products simultaneously. It may just produce one and be in the process of deciding whether to embark on the production of a complementary product. In calculating the profitability of such expansion, the company must include the increase in sales of and profit earned on the earlier product. If it omitted this positive effect, it would be understating the benefits of the new product. It may decide against the product's introduction when in fact the total profits of the company would increase if the new product were brought to market. As an example, suppose a successful producer of television sets is considering whether to introduce a new line of DVD players. In calculating the potential profitability of producing DVD players, the producer must include the possibility of enhanced sales (and profits) from its television line.

Products Substitutable in Demand

A brief treatment of substitutability and pricing will suffice because this case is extremely similar to that of complements. For substitutes, the effect to be considered is the decrease (increase) in revenue and profits of a second product if quantities bought of the first product rise (fall), either because of changes in demand or changes in price. Examples of such cases abound. Two different sizes of personal computers certainly are substitutes for one another. The different automobile models produced by one manufacturer (sedans versus convertibles, Honda Civics versus Accords, etc.) are relatively close substitutes, so it is necessary to price them jointly. Another example is Global Foods' soft drink division, which produces cola-type and non-cola-type sodas simultaneously.

Just as in the case of complementary products, substitution can occur when a new product is introduced. Thus, a computer manufacturer developing a new generation of computers must consider the impact that the introduction will have on similar but less advanced products now being marketed.

The analysis of these cases is basically the same as for complementary commodities. The marginal revenue of one product will be a function of the quantities sold of both commodities, and the prices of the two will be found by solving simultaneous equations. However, in this case, the sales of one product will have a negative impact on the sales of the other.

Joint Products with Fixed Proportions

Certain products will be produced together from one set of inputs. In some instances, the two products will be produced in fixed proportions to one another. Although precisely fixed proportions may not occur often in the real world, relative fixity is commonly encountered, particularly in the short run. The example given earlier involved the products of a beef carcass and a hide (only one of each can be obtained from one steer). Other examples are soybean meal and soybean oil, and coconut milk and coconut meat. In many cases, there is a principal product and one or more by-products.

Assume products A and B are produced jointly in fixed proportions. Only one cost curve can be constructed in this case. However, the demand curves for the two

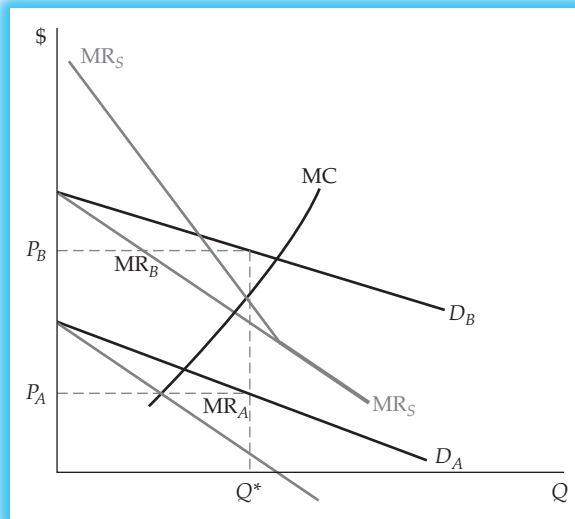


Figure 10.7 Price Determination for Joint Products Made in Fixed Proportions

products are independent (e.g., the demand for coconut meat is not related to the demand for its milk). Thus, the two demand curves and their respective marginal revenue curves can be added vertically to obtain a total demand curve (not shown in Figure 10.7) and a total marginal revenue curve (shown as MR_S). Observe, however, that when one of the separate marginal revenue curves goes negative, it becomes irrelevant to the solution of the problem, because no business would produce at a point where marginal revenue is negative. To the right of this point, the total marginal revenue curve will be coincident with the marginal revenue of the product that is still in the positive range. Production will take place (using our usual maximization rule) where total marginal revenue equals marginal cost. The prices of the two separate products can be found at the quantity indicated on their respective demand curves. Figure 10.7 shows the results. D_A , D_B , MR_A , and MR_B are the demand and marginal revenue curves for the two products, and MR_S represents the vertical summation of the two individual marginal revenue curves. (The summed demand curve is actually irrelevant to the solution of the problem and need not be shown.) As can be seen, MR_S becomes identical with MR_B to the right of the point where MR_A becomes negative.

The curve MC represents the marginal cost of the joint product. Production will take place where marginal cost is equal to MR_S , which is at quantity Q^* on the graph. The prices charged for the two products will be found on their respective demand curves at P_A and P_B . An interesting aspect of this type of construction is that if the optimal production quantity were to the right (i.e., at higher quantities) of the point where one of the marginal revenue curves (in our case, that for product A) becomes negative, it would become profitable for the company to produce this total amount but not to sell quantities of product A beyond the point where its marginal revenue becomes zero. The company should discard the excess product A assuming free disposability.

Another important point is the effect of the change in the demand for one of the two jointly produced products. If the demand for product B rises and the price of B thus increases, there will be a decrease in the price of A (as it will be produced at a lower point on its demand curve).

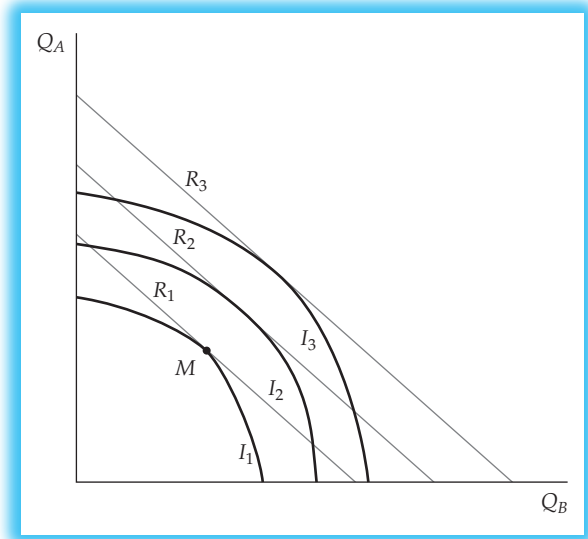


Figure 10.8 Joint Production Using Variable Proportions

Joint Products in Variable Proportions

When we relax the limitation of fixed proportions, we have the usual case of joint production. Indeed, when two products are produced from similar resources in variable proportions (i.e., if we produce more of one product we must produce less of the other) the situation is not dissimilar to the general case of production of different products with limited resources. We are essentially describing the “guns or butter” situation. Under short-run conditions, there is a given amount of resources with which the two products can be produced.

Figure 10.8 illustrates this situation. Curve I_1 is an isocost curve; the total cost of production is the same at each point because all points on this curve are based on the same total inputs in the production process. Different points on the isocost curve relate to how these resources are allocated to product A or B. An essential requisite is that the curve be concave to the origin: As more of one product is produced, progressively larger quantities of the other must be given up. The isocost curve shows the alternative quantities of product A and B that can be produced. If the prices of the two products are constant regardless of quantity (i.e., we are implicitly operating under conditions of perfect competition), then a straight-line isorevenue curve can be drawn. At each point on R_1 in Figure 10.8, identical revenue is obtained. To optimize, the company will produce at the point of tangency between the isorevenue and isocost curves (point M in the figure). This represents the highest revenue that the company can attain for a given total cost. If revenue at this point is greater than cost, economic profit will result.

The company could also move from one isocost curve to another (e.g., from I_1 to I_2 to I_3), and consequently from one isorevenue curve to another, representing the use of additional variable resources. Production would take place at the tangency point that results in the largest difference between total revenue and total cost. This point signifies the maximum economic profit the company can achieve. In the long run, these isocost curves would include changes in all resources, including those (e.g., plant and equipment) that are fixed in the short run. Under conditions of perfect competition, the optimal tangency would occur where total revenue just equals

total cost, with no economic profit. At all other points, economic losses would be sustained.

We have limited our discussion to a simple model, using only two products and assuming perfect competition. Much more complex models could be developed, involving more than two products, noncompetitive conditions, and demand interrelationships, for example. The results would be more difficult to obtain, and relatively complicated mathematical models would have to be introduced, but the principles of economic maximization would still apply. Whenever a company takes action in introducing a new product, producing more of one at the expense of another or eliminating a certain commodity from its product line, and when such action is taken to improve its short-run or long-run profitability, the company is guided by the basic principle of equating marginal cost with marginal revenue.

TRANSFER PRICING

In today's complex industrial world, many companies have subdivided their operations into several groups or divisions. As a product moves from its early stages to the point where it is ready to be sold to consumers, it is passed from one operating division of the company to another. In the automobile industry, for example, various auto parts may be produced in different plants and then assembled into the finished product in yet another plant. Computer and peripheral equipment components may be produced in one plant and assembled into different products at other plants. Then, to sell the products, the marketing arm of the company may have to assemble the various individual machines into complete systems.

To continue this analysis, it is necessary to address the notion of *profit center*, a frequently used term that refers to a situation prevalent in large corporations. The management of each division is charged with a profit objective. Thus, each stage of production must measure its costs and then establish a price at which it will "transfer" its product to the next stage. However, if each intermediate profit center were to set its price to maximize its own profit, the price of the final product may not maximize the profit of the company as a whole, which is the appropriate objective. The price set by the division transferring the intermediate product becomes the cost of the division receiving this product. If that price is set too high, this may start a chain reaction resulting in the final product price being higher than the price which would maximize the company's profit. The **transfer pricing** mechanism must be geared toward maximizing total company profit; therefore, the final pricing policy may be dictated centrally from the top of the corporation.

Such processes can be extremely complicated, particularly if there are more than two steps in the transfer process. Further, the intermediate products may be only for internal usage. In contrast, the producing division may also be selling its product in an external market, and the receiving division may be free to purchase the intermediate product from a competitor, if that would improve the company's profit situation. Let us discuss each case in turn. To simplify matters, we will assume the existence of just two divisions, one that manufactures components (division C), and another that assembles them into the final product and sells it (division A).

No External Markets

If there is no possibility for division A to buy components from a competing firm and no possibility for division C to sell components to other companies, then the two divisions must deal with equal quantities; division C will produce exactly the number of components that will be used by division A for assembly and sales. This situation is illustrated in Figure 10.9. The company will be faced by a demand curve for the

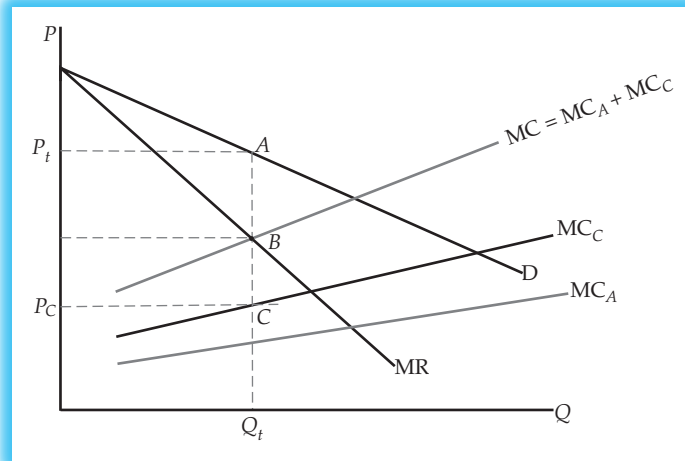


Figure 10.9 Transfer Pricing with No Market for Intermediate Product

final product and two marginal cost curves (D , MC_A , and MC_C). The two marginal cost curves, one for each division, will be summed vertically to obtain total marginal cost (MC), and the company will maximize its total profit by equating the total marginal cost with marginal revenue (MR). Production will take place at that intersection (point B), and the price for the final product will be the corresponding price on the demand curve (point A). The quantity sold is Q_t . The transfer price for the intermediate product is P_C .

External Markets

It may be possible for division C to sell its (intermediate) product in a competitive market and for division A to purchase division C 's product (an identical product) in a competitive market. In that case, the pricing of the product will proceed as follows:

1. Division C will produce at the point where its marginal cost equals the market price. (Because we are assuming the existence of a competitive market, the demand curve is horizontal, resulting in a uniform price regardless of quantity.)
2. The cost of the intermediate product to division A is the market price. This will be added to division A 's marginal cost curve to obtain the total marginal cost for the final product.
3. Production will take place at the quantity where the total marginal cost equals the marginal revenue for the final product.

Figure 10.10 shows a case where there is a competitive market for the intermediate product, and division C produces components in excess of those used by division A . Assume the market price for division C 's product is P_C . The optimal quantity to be produced by C is Q_C , given its marginal cost MC_C . The marginal cost of division C 's product to division A is now P_C and the total marginal cost MC equals MC_A plus P_C . The company will sell its product at optimal price P_t at quantity Q_t . Division C will sell quantity Q_C minus Q_t in the competitive market.

Of course, should division C for some reason attempt to price the intermediate product in excess of the market price, then A would buy all the intermediate product in the external market.

If the optimal output of division C is less than A wants to buy, then A will turn to the external market for the additional units of the intermediate product it needs

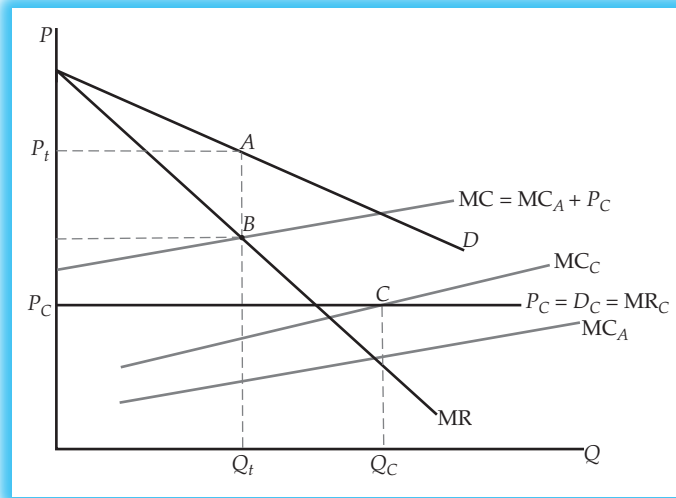


Figure 10.10 Transfer Pricing with Competitive Market for Intermediate Product

to maximize company profits. The graph for this situation would be quite similar to Figure 10.10, and we leave it to the reader to illustrate this particular situation.

In this section, we discuss transfer pricing among different divisions of a company when all the divisions are located in the same country. When different parts of a corporation are located in different countries and products are shipped across borders, the situation becomes even more complex. We discuss this subject in detail in Chapter 13.

OTHER PRICING PRACTICES

Price skimming occurs when a firm is the first to introduce a product. It may have a virtual monopoly, and often will be able to charge high prices and obtain substantial profits before competition enters.

In **penetration pricing** a company sets a relatively low price in order to obtain market share.

Limit pricing exists when a monopolist sets a price below the monopoly price (where $MR = MC$) to discourage potential competitors from entering the market and competing. At the lower price, the monopolist's profit will be below its maximum. Of course, limit pricing may be based on the monopolist's expectation that its cost will decrease because of the existence of a learning curve, so that ultimately the limit price will become the profit maximizing price.

Another special pricing practice is **predatory pricing**, where a company prices below its marginal cost to cause competitors to exit the market. After the competitors have left the market the company will raise its prices. However, this practice is not seen very often in the United States because it is illegal under the Sherman Antitrust Act. Also, it may create considerable losses to the monopolist and may not be worthwhile in the long run. Further, after the monopolist raises prices a new competitive threat may appear. Thus the company may have to lower prices again and incur new losses.

With **prestige pricing**, demand for a product may be higher at a higher price because of the prestige that ownership bestows on the buyer.

Psychological pricing takes advantage of the fact that the demand for a particular product may be quite inelastic over a certain range but will become rather elastic at

one specific higher or lower price. Such a demand curve has the appearance of a step function.

Complete explanations of these pricing practices can be found in many marketing textbooks.

GLOBAL APPLICATION: THE DECLINE OF EUROPEAN CARTELS

European governments have in the past tolerated and supported monopolies, and while pricing cartels have been illegal, they have flourished. This situation has changed, however, and the European Commission has been quite active in pursuing cartels. Several examples of the European Commission's actions over the last decade and more are summarized briefly below.

The European Carton-Board Cartel

A case involving nineteen carton-board producers in ten European countries actually resembles the electrical manufacturers' case discussed previously.

In July 1994, the European Commission imposed record fines of \$159 million on nineteen producers. The cartel involved a market-sharing agreement and orchestrated 6 to 10 percent price increases each 6 months from 1987 to 1990. Monthly "social" meetings were arranged at luxury hotels (most frequently in Zurich). "The cartel's members then compared the state of their order books to judge when best to introduce a price increase. Sometimes the big producers agreed on temporary plant stoppages to keep production under control."⁴⁷ Fake minutes were drawn up to disguise the business that was transacted.

The European Commission acted after complaints by the industry's customers about the continual price increases during a period of sharp economic downturn. In 1991, officials of the commission staged simultaneous raids on the producers and found private notes documenting the dates and amounts of price increases. One of the companies decided to admit to the conspiracy and aided the probe, which resulted in heavy fines.

The European Vitamin Cartel

In November 2001, the European Commission imposed fines totaling more than € 850 million (approximately \$747 million) on thirteen pharmaceutical companies for price fixing of vitamins that had spread over 9 years. Most of the fine was levied on Roche of Switzerland and BASF of Germany. This decision followed 2 years after the United States, who under a similar investigation, had fined the same two companies an amount exceeding \$700 million. It was said that the companies had "virtually operated as a single company." The cartels met each month or quarter, exchanged sales and price data, and reviewed quotas allotted to each company. In some cases, "participants even prepared an annual 'budget' to adjust sales according to each company's quota."⁴⁸

Some Recent Cases of Price Fixing

In January 2007, the European Commission fined ten European and Japanese engineering companies a total of €751 million for price fixing of switchgear used to control

⁴⁷This discussion and quotation are based on Emma Tucker, "Price-Fixing Cartel Given Record Fine by Brussels," *Financial Times*, July 14, 1994, and Emma Tucker, "Rise and Fall of a House of Cards," *Financial Times*, July 15, 1994.

⁴⁸Francesco Guerrera and Birgit Jennen, "European Groups Face Record Fines for Roles in Price-Fixing," *Financial Times*, November 21, 2001; Francesco Guerrera, "Monti Lashes 'Vitamins AG' Cartel," *Financial Times*, November 22, 2001.

electricity to homes, offices, and factories. The commission stated that the companies “rigged bids for procurement contracts, fixed prices, allocated projects to each other, shared markets, and exchanged important and confidential information” between 1988 and 2004.⁴⁹

What was called the “largest ever” fine of €992 million (\$1.3 billion) in the European Union was levied on five European elevator makers “for operating cartels for the installation and maintenance of lifts and escalators in Germany, Belgium, Luxembourg, and the Netherlands.” Of the five companies, the German conglomerate ThyssenKrupp AG received the largest fine, €479 million. The European Commission stated that the cartel rigged bids, fixed prices, and shared markets during the period 1995–2004.⁵⁰

Ten producers of memory chips were fined a total of €331 million for operating as a price-fixing cartel. The cartel appeared to operate from 1998 to 2002. It shared information to set prices for dynamic random access memory chips (DRAMs) in Europe. It was reported that this was the first time companies accused of price fixing decided to agree to settle. Samsung of South Korea received the largest fine of €159 million, while German companies Infineon and Hynix were fined €56 and €51, respectively.⁵¹

Seventeen steel manufacturers were assessed a total of €518 million; Arcelor Mittal incurred the largest fine of €276 million. The price fixing lasted from 1984 to 2002. The scheme “occurred in the market for prestressing steel, which forms the curled steel wires used with concrete and to form the foundations of buildings and bridges.” “The participants fixed quotas and prices and allocated clients and exchanged commercial information.” The cartel affected a majority of European Union member states. The existence of the cartel was revealed by DWK/Saarstahl of Germany; the company was not fined.⁵²

The European Commission fined Procter & Gamble €211.2 million and Unilever €104 million for fixing the prices of laundry detergents during the period 2002–2005. The companies “agreed to make the packages smaller without decreasing prices and, later, to raise prices while keeping the size of their packages the same.” The German company, Henkel, avoided the fine because they informed the European Commission about the cartel.⁵³

Price Discrimination by Airlines

In December 2003, the European Commission asked eighteen European airlines whether they charge different fares on the same flights to residents of different European Union countries. Most of these price differentials occurred in tickets ordered on the Internet and were not connected with different dates of purchase, possibility of changing reservations, or other features that could affect cost. Apparently, this price discrimination existed. After receiving answers from sixteen of the eighteen airlines, the commission ended its action in June 2004 because this pricing practice had been eliminated after the commission’s inquiries.⁵⁴

⁴⁹Russell Hotten, “Engineering Cartel ‘Cheats’ Are Given Record Fines,” *Daily Telegraph*, January 25, 2007.

⁵⁰“EU Slaps U.S. \$1.3 Billion Fine on Four Elevator Makers for Price Fixing,” *Canadian Press*, February 21, 2007.

⁵¹Raff Caser, “EU Slaps \$403 Million Fine on 10 Memory Chip Makers, Including Samsung, in Price Fixing Scheme,” *Associated Press Newsires*, May 19, 2010.

⁵²Nikki Taut, “ArcelorMittal Fined €278 Million for Price-Fixing,” *Financial Times*, July 1, 2010.

⁵³Catherine Boyle, “Washing Powder Cartel Fined for Price-Fixing Eco Products,” *The Times*, April 14, 2011; Graeme Wearden, “Unilever and P&G Fined for Price-Fixing,” *The Guardian*, April 14, 2011.

⁵⁴“Air Transport: Suspicion of Price Discrimination Practices,” *European Report*, December 24, 2003; “Airlines Put End to Price Discrimination on Basis of National Residence,” *Europe Information*, June 9, 2004.

Another case of price discrimination by an airline was reported in 2003. Apparently, foreigners flying domestic routes in Vietnam were charged considerably more (as much as 50 percent) than local fliers. In December 2003, Vietnam's Civil Aviation Administration announced that it would abolish this type of discrimination on January 1, 2004.⁵⁵

AN AIRLINE PRICING EXAMPLE

A very recent article described the pricing created by airlines on international trips. While these pricing methods do not necessarily represent examples of price discrimination (they follow demand patterns), it is interesting to look at how airlines do their pricing and how potential travelers can benefit from lower prices.

"International sales are seasonal, not weekly, and often route-by-route rather than global." Sales for travel in the fall are often announced in late summer. The cheapest fares of the year are in February and appear in late November and early December. Trips on Tuesday, Wednesday, and Thursday are usually cheaper than weekend trips. "For every route, airlines load a dozen or more different fares into reservations systems and then pick which one applies to a specific flight at a specific time." The fares are set based on the demand for that flight. The best time to purchase tickets for peak periods is April.

Fares to London, which was the host for the Summer Olympics in 2012, did not increase dramatically due to the increased demand, according to Travelocity. Airlines had added capacity for flights to London. Travelocity reported that average prices for flights to London increased only 2 percent compared to 2011.⁵⁶

The Solution

Rebecca James went to see Philip Olds of the food division to solicit his advice on the bid she intended to make for the energy drink supply contract with the large airport catering company.

"One thing is certain," he said. "You have to bid at a price considerably lower than the price at which you sell to small retail stores. These stores have a leeway in how much they can charge their customers because they are really selling convenience. A small variance in price is not going to change their sales significantly, so they will not stop buying from us as long as our price is not out of line.

"However, your potential new customer intends to give a contract to only one additional supplier, and 10 cents per case will make a large difference when hundreds of thousands of cases are involved. Thus, you will have to shave your markup as much as possible. It may turn out that you will make precious little profit, if any, on this contract."

Philip was telling Rebecca that the demand price elasticity of the large caterer was quite different from that of the small retailers. As a result, Global Foods could sell to these two markets at different prices.

Rebecca is thus confronted with a case of price discrimination. The demand elasticities in the two markets are probably quite different; thus higher prices can be charged in the market displaying lower elasticity. The separation of the two markets (i.e., no cross-selling) represents the second important condition for the existence of price discrimination. Another important consideration for Rebecca is that the price to be charged must enable the company to gain a foothold in this large market (penetration or entry pricing).

Back in her office, Rebecca begins to work on determining the price. She assumes the average cost per case of "Go-nuts" for the two customers is the same. Although this is certainly not quite accurate (the large shipments the company will make to the new customer if the bid is won will probably create some cost savings), the unit cost differences will probably not be significant, so her analysis will lose little if she assumes equality.

(continued)

⁵⁵"Vietnam to End Airfare Price Discrimination Jan. 1," *Kyodo News*, December 2, 2003.

⁵⁶Scott McCartney, "The Middle Seat: Finding the Best Fares to Europe," *Wall Street Journal*, June 28, 2012.

(continued)

She estimates the company's usual markup to be about 50 percent. After some additional consideration, and a review of some data for the industry that she obtained, she feels that a 20 percent markup would put the company in a good competitive position.

Before making this recommendation to her boss, she will consult with the finance and accounting departments to determine what, if any, profit would be realized on this transaction. She will, of course, include this analysis with her recommendation. It is quite possible that at this low price profit will be extremely marginal, and she will then have to argue that obtaining a foothold in this market will have beneficial long-run consequences.

The decision that she has made has some rather important implications for the estimate of demand elasticities for the two classes of customers. The lower the demand elasticity, the higher will be the markup that a company can obtain. Following the equation developed previously,

$$(1 + M) = \frac{E_p}{E_p + 1}$$

an elasticity of -3 conforms to a markup of 50 percent, whereas an elasticity of -6 corresponds to a 20 percent markup. Of course, these numbers are approximations, but such an estimate could be of great help to Rebecca in her attempt to set the proper price.

SUMMARY

This chapter is built on the foundation laid in Chapters 8 and 9 by applying the principles of pricing and output to specific pricing situations, most under conditions of imperfect competition. Briefly, we learned the following:

1. Cartels are formed to avoid the uncertainties of a possible reaction by one competitor to price and production actions by another. The firms in the industry agree on unified pricing and production actions to maximize profits. However, as history shows, such arrangements are not always stable.
2. Price leadership exists when one company establishes a price and others follow. Two types of price leadership were discussed: barometric and dominant.
3. Baumol's model describes the actions of a company whose objective is to maximize revenue (rather than profits) subject to a minimum profit constraint.
4. Price discrimination (or differential pricing) exists when a product is sold in different markets at different prices. Third-degree price discrimination is the most common. By charging different prices in separate markets that have demand curves with different price elasticities, a firm can increase its profits over what they would be if a uniform price were charged.
5. Cost-plus pricing appears to be a very common method. However, such pricing does not necessarily imply that marginal principles and demand curve effects are not taken into consideration.
6. Multiproduct pricing was examined, because most firms and plants produce more than one product at the same time. Multiple products produced by one firm can be complements or substitutes, both on the demand side and the supply side. Four possible cases were discussed, and it was shown how application of the marginal principle brings about profit maximization.
7. Several other pricing practices were summarized. One was transfer pricing, which is used to determine the price of a product that progresses through several stages of production within a firm.

IMPORTANT CONCEPTS

- Barometric price leadership.** In an oligopolistic industry, a situation in which one firm, perceiving that demand and supply conditions warrant it, announces a price change, expecting that other firms will follow. (p. 396)
- Baumol model.** A model hypothesizing that firms seek to maximize their revenue subject to some minimum profit requirement (i.e., the profit constraint). (p. 398)
- Cartel.** A collusive arrangement in oligopolistic markets. Producers agree on unified pricing and production actions to maximize profits and to eliminate the rigors of competition. (p. 391)
- Cost-plus pricing.** Also called *full-cost pricing*, a practice in which prices are calculated by adding a markup to total cost. (p. 411)
- Dominant price leadership.** In an oligopolistic industry, a firm, usually the largest in the industry, sets a price at which it will maximize its profits, allowing other firms to sell as much as they want at that price. (p. 397)
- Limit pricing.** A monopolist will set price below $MR = MC$ to prevent potential customers from entering the market. (p. 422)
- Multiproduct pricing.** Pricing that reflects the interrelationship among multiple products of a firm that are substitutes or complements. (p. 415)

- Penetration pricing.** Charging a lower price than indicated by economic analysis in order to gain a foothold in the market. (p. 422)
- Predatory pricing.** Setting price below marginal cost to drive competitors out of the market. (p. 422)
- Prestige pricing.** A perception that charging a higher price will increase quantity sold because of the prestige obtained by the buyer. (p. 422)
- Price discrimination.** A situation in which an identical product is sold in different markets at different prices. (p. 399)
- Price leadership.** One company in an oligopolistic industry establishes the price, and the other companies follow. Two types of price leadership are common: barometric and dominant. (p. 396)
- Price skimming.** The practice of charging a higher price than indicated by economic analysis when a company introduces a new product and competition is weak. (p. 422)
- Psychological pricing.** The practice of charging, for example, \$9.95 rather than \$10 for a product, in the belief that such pricing will create the illusion of significantly lower price to the consumer. (p. 422)
- Transfer pricing.** A method to correctly price a product as it is transferred from one stage of production to the next. (p. 420)

QUESTIONS

1. “If a company sets its prices on the basis of a cost-plus calculation, it cannot possibly suffer a loss on its products.” True or false? Comment.
2. Price discrimination is often defended on the basis of fairness. What is meant by this statement? Comment on its validity.
3. Which products in each pair would tend to have higher markups in a supermarket?
 - a. Cigarettes versus tomatoes
 - b. Potatoes versus orange juice
4. Many years ago, a neighborhood lunch counter charged 15 cents for a cup of coffee and 15 cents for a buttered hard roll. One day, a customer ordered the two items and was told that the total price was 35 cents. When the customer asked which of the two items had been raised by 5 cents, the owner’s condescending reply was, “Which do you think?” In your opinion, which of the two items was affected and why?
5. Differentiate *barometric* price leadership and *dominant* price leadership.
6. Is there a similarity between cartel pricing and monopoly pricing?
7. What conditions are favorable to the formation and maintenance of a cartel?
8. Can government be a potent force in the establishment and maintenance of monopolistic conditions? Name and describe such occurrences.
9. Describe the properties of the Baumol revenue maximization model. Do you consider this to be a good alternative to the profit maximization model?
10. Telephone companies charge different rates for calls during the day, in the evening, and at night or weekends. Do you consider this to be price discrimination?

11. Is cost-plus pricing necessarily inconsistent with marginal pricing?
12. Airline ticket prices may differ with respect to when the ticket is bought, how long a passenger remains on the trip (e.g., over a weekend) and other variables. Are these differences a case of price discrimination?
13. Does cost-plus pricing necessarily ignore the demand curve?
14. Define and describe (giving examples):
 - a. Transfer pricing
 - b. Psychological pricing
 - c. Price skimming
 - d. Penetration pricing
15. Under what circumstances would a discriminating monopolist produce a more socially optimal quantity than a nondiscriminating monopolist? Is there any situation under which a discriminating monopolist could produce the quantity that would be produced under competition?

PROBLEMS

1. There are only two firms in the widget industry. The total demand for widgets is $Q = 30 - 2P$. The two firms have identical cost functions, $TC = 3 + 10Q$. The two firms agree to collude and act as though the industry were a monopoly. At what price and quantity will this cartel maximize its profit?
2. An amusement park, whose customer set is made up of two markets, adults and children, has developed demand schedules as follows:

Price (\$)	Quantity	
	Adults	Children
5	15	20
6	14	18
7	13	16
8	12	14
9	11	12
10	10	10
11	9	8
12	8	6
13	7	4
14	6	2

The marginal operating cost of each unit of quantity is \$5. (*Hint:* Because marginal cost is a constant, so is average variable cost. Ignore fixed cost.) The owners of the amusement park want to maximize profits.

- a. Calculate the price, quantity, and profit if
 1. The amusement park charges a different price in each market.
 2. The amusement park charges the same price in the two markets combined.
 3. Explain the difference in the profit realized under the two situations.
- b. (Mathematical solution) The demand schedules presented in Problem 2 can be expressed in equation form as follows (where subscript A refers to the adult market, subscript C to the market for children, and subscript T to the two markets combined):

$$\begin{aligned}
 Q_A &= 20 - 1P_A \\
 Q_C &= 30 - 2P_C \\
 Q_T &= 50 - 3P_T
 \end{aligned}$$

Solve these equations for the maximum profit that the amusement park will attain when it charges different prices in the two markets and when it charges a single price for the combined market.

3. The Bramwell Corporation has estimated its demand function and total cost function to be as follows:

$$Q = 25 - 0.05P$$

$$TC = 700 + 200Q$$

Answer the following questions either by developing demand and cost schedules (*Hint: Use quantities from 1 to 14*) or by solving the equations.

- a. What will be the price and quantity if Bramwell wants to
 1. Maximize profits?
 2. Maximize revenue?
 3. Maximize revenue but require the profit to be a minimum of \$300?
 - b. Now assume the cost function is $TC = 780 + 200Q$, while the demand function remains the same. What will the price and quantity be if Bramwell wants to
 1. Maximize profits?
 2. Maximize revenue?
 3. Maximize revenue but require the profit to be a minimum of \$300?
 - c. Why are the answers the same in $a(1)$ and $b(1)$ but different in $a(3)$ and $b(3)$?
4. The Great Southern Paper Company has the following marginal cost schedule for producing pulp:

Quantity (tons)	Marginal Cost
1	\$18
2	20
3	25
4	33
5	43

Pulp can be bought in the open market for \$25 per ton. The marginal cost of converting pulp into paper is $MC = 5 + 5Q$, and the demand for paper is $P = 135 - 15Q$. Calculate the marginal cost of paper if the company produces its own pulp. What is the profit-maximizing quantity? Should the company purchase pulp from the outside or produce it in-house?

5. Each pair of Fancy Shoes costs Bradbury Footwear Stores \$30. Bradbury's economist has estimated the point elasticity to be -1.8 . What price should Bradbury charge if it wants to maximize profits?
6. An airplane manufacturer has annual fixed costs of \$50 million. Its variable costs are expected to be \$2 million per plane. If the manufacturer wants to earn a 10 percent rate of return on its investment of \$400 million and expects to produce 100 aircraft this year, what will its markup on total cost have to be? If it expects to produce 150 aircraft, what will its markup have to be?
7. Schultz's Orchard grows only two types of annual fruit—melons and strawberries—and over the years it has been able to chart two production levels and resulting total costs. The figures are shown in the following table where quantities are in flats (for strawberries) and cartons (for melons).

Production Level A		Production Level B	
Strawberries	Melons	Strawberries	Melons
2,700	0	4,200	0
2,400	600	3,600	900
1,800	1,200	2,700	1,800
1,200	1,500	2,100	2,100
750	1,650	900	2,550
0	1,800	0	2,700
Total cost of A: \$15,000		Total cost of B: \$25,000	

This year it is expected that the price of strawberries will be \$10/flat and melons will be \$15/carton.

What is the best production mix at each cost level? How much profit should Schultz's Orchard expect at each production level.

8. The Prestige Office Equipment Company produces and sells different types of office furniture. One of the important items it sells is a high-quality desk. During the past year, Prestige sold 5,000 of these at a price of \$500 each. The contribution profit for this line of furniture last year was \$700,000.

A consultant suggests that Prestige decrease the price of each desk by \$30. In his opinion, another 500 desks could then be sold, and the total profit would be maintained. A trade publication that employs an economist has estimated price elasticity of office furniture (including desks) to be about -1.8 .

Assume the variable unit cost per desk in the coming year will remain the same. Evaluate the consultant's proposal. Be sure to include in your answer the price elasticity assumed by the consultant, as well as the published elasticity estimate.

9. The royalties received by an author for writing a college text are frequently set at a rate of about 15 percent of the publisher's book price. This may create a conflict between the goals of a profit-maximizing publisher and those of a royalty-maximizing author. As a student consumer (assuming you have to pay for your textbooks), whose goal would be more beneficial to you? Why? Demonstrate this situation graphically (assume a downward-sloping demand curve).
10. The Prime Company produces two products, X and Y. They are produced jointly so for each X manufactured a unit of Y is also produced. The joint cost function is

$$TC = 50 + 2Q + .5Q^2$$

Q represents the number of joint units produced. The demand equations for the two products are the following:

$$\begin{aligned} Q_x &= 100 - P_x \\ Q_y &= 60 - 2P_y \end{aligned}$$

- How many units should the company produce per period?
- What price should it charge for each of the joint products?
- What will be the company's profit per period?

Assume the company is a profit maximizer.

11. George's Pants Store sells a particular brand of slacks. It recently decreased the price of a popular brand by 5 percent. This decrease was followed by a 20 percent increase in sales. The marginal cost of producing these slacks is \$100.
- What is the point price elasticity of demand for these slacks?
 - What would be the optimal markup and price of these items?
12. A firm makes two products, x and y . Inverse demand for each shows that pricing in one market depends on sales in the other according to the equations:

$$P_x = 1000 - 20x + 3y \quad \text{and} \quad P_y = 500 - 5y + x$$

The firm faces joint fixed cost of \$12,000 and constant marginal cost of production in each product segment, $MC_x = \$200$, and $MC_y = \$100$.

- What bundle of products (x^* , y^*) should the firm produce?
- What prices will the firm be able to charge for each product given production at (x^* , y^*)?
- What profits result in this instance?
- At (x^* , y^*), what are the values of $\frac{\partial TR_y}{\partial x}$ and $\frac{\partial TR_x}{\partial y}$? Provide a short (one- or two-sentence) explanation for each value. That is, explain to your less mathematically sophisticated boss the economic significance of each value.

- e. Check your work by determining the profits that result if x and y are one unit more or less than optimal and fill in the following table.

π given production near optimal (x^*, y^*)				
$y \backslash x$	value of x			
value of y	π at (x, y)	$x^* - 1$	x^*	$x^* + 1$
	$y^* + 1$			
	y^*			
	$y^* - 1$			

Hints: This problem may be solved using the structure described in footnote 46 in this chapter. All answers are whole numbers. Part e is easily accomplished using Excel by setting up a test area where you can try different values of x and y and determine profit.

13. Redo Problem 12 given the following inverse demand functions:

$$P_x = 1000 - 20x - 3y \text{ and } P_y = 500 - 5y - 2x$$

Hint: All answers to this problem are whole numbers, just like the answers to Problem 12.

14. Redo Problem 12 given the following inverse demand functions:

$$P_x = 1000 - 20x \text{ and } P_y = 500 - 5y$$

Hint: All answers to this problem are whole numbers, just like the answers to Problems 12 and 13.

15. This problem requires completion of Problems 12–14. Can you explain what happened to the optimal x and y in the three problems? It may be easiest to explain what happened in moving from the model with negative cross-good coefficients in Problem 12, to zero in Problem 14, and to positive in Problem 13.

Game Theory and Asymmetric Information

Learning Objectives

Upon completion of this chapter, readers should be able to:

- Define game theory, and explain how it helps better understand mutually interdependent management decisions.
- Explain the essential dilemma faced by participants in the prisoners' dilemma game and explain why many oligopolistic situations resemble the prisoners' dilemma.
- Explain the circumstances when the prisoners' dilemma is more of a dilemma for firms.
- Explain the difference between cooperative and noncooperative games, and explain why many of the games played by firms are necessarily non-cooperative in nature.
- Explain the tactics used in bargaining, especially tacit bargaining.
- Explain how focal points and limited war relate to oligopolistic interactions.
- Explain the key problems that arise in a market where buyers and sellers do not have the same information about a product (i.e., when "asymmetric information" exists between buyers and sellers).
- Briefly explain the concepts of "adverse selection" and "moral hazard" and why they exist in markets with asymmetric information.
- Explain how the moral hazard problem relates to the principal-agent problem.
- Explain the essential attribute of a signal as used by Michael Spence and explain how signals can help avoid the lemons problem discussed by George Akerlof.
- Explain how parties that are at an informational disadvantage can sort those on the other side of a transaction into groups based voluntary action.

The Situation

Henry Caulfield did not want to lower the price of his soft drinks even further because he knew it would start a vicious price war between him and the other two store owners in the vicinity.¹ But he believed that this might be the only way he could restore his share of the market for a product that was an important part of his business's profitability.

To his surprise, his daughter Erica wanted to stay at home and work in the family store over the summer to gain some practical business experience. She had just decided to switch her major from art to economics, so she believed that working in her dad's store would be a useful complement to her studies. But when he confided in her about his dilemma regarding the pricing of his soft drinks, he was a bit puzzled by her reaction. "Dad, what you're faced with is a classic game theory problem," she responded.

"I'm not so sure what you mean, Erica. This is not a game we're talking about. This is the real world of business, and the future of my business could well depend on the outcome of this potential price war over soft drinks. If my competitors get the best of me in the pricing of this product, who knows what they'll do in the pricing of coffee, snacks, and gasoline?"

"Don't worry, Dad, I know this is serious. I know that you probably didn't study this topic when you were in college, but when my professor introduced game theory to us in my managerial economics course, I was fascinated. I believe there are some practical applications of the concepts that may be of help to you in this situation."

"Okay, Erica, what's this game theory stuff all about and how can it help me?"

INTRODUCTION

The optimization analysis used throughout this text, such as the use of the $MR = MC$ rule in the setting of price and output, provides the theoretical foundation for managerial decision making. But this type of analysis has two shortcomings when applied to actual business situations. First, it assumes other factors such as the reaction of competitors or the tastes and preferences of consumers remain constant as the firm sets its optimal price and output according to the $MR = MC$ rule.² In reality, managers must often make these decisions in the face of changing competitive reactions and changing consumer tastes and preferences, as Henry Caulfield discovered in "The Solution" vignette in Chapter 4. Second, managers must sometimes make decisions when certain parties have more information about market conditions than others. This chapter presents two theories—game theory and the theory of asymmetric information—that have been incorporated by economists into managerial economics to address the shortcomings of standard optimization analysis. The first part of this chapter talks about **game theory** and shows how this theory can be applied to the setting of prices in oligopolistic markets and can help firms develop their business strategies. The second part presents the basic concepts of the theory of asymmetric information and illustrates when, why, and how managers need to incorporate these particular concepts in making their business decisions.

¹This situation continues from "The Situation" and "The Solution" vignettes presented in Chapter 4.

²The kinked demand model presented in Chapter 9 tries to explain why prices in oligopoly markets tend to be sticky, but it does not explain how and why this "administered price" is initially established. Also, in the early 1900s, the French economist Antoine Augustin Cournot developed a mathematical model to explain how price is determined in a market with only two competitors (i.e., a duopoly). However, his model did not formally describe the adjustment process moving the firms toward Cournot equilibrium.

GAME THEORY: MODELING THE STRATEGY OF CONFLICT³

Economists have extended and refined the analysis of oligopoly markets using game theory to examine strategic interaction. A game is distinguished by the number of players in the game and the number of options or strategies available to each player. A **game** involves players making strategic decisions from an available set of options. **Players** are the decision-making units that play games. Each player chooses among various strategies or options. A **strategy** is an option available to a player.

In the present context, a player is simply a firm, and the options available to that player (firm) are the policy alternatives under the control of the player. For example, there might be two nightclubs in town and each can decide whether to have a deejay or a live band play on a Saturday night. The players are the two nightclubs (or more accurately the managers of the two nightclubs), and the options are whether to have a live band or a deejay. This is a 2×2 game (two players with each player having two options). Each of the four possible alternatives (two players with two strategies each) will have payoffs associated with that alternative for both players. **Payoffs** are the outcomes associated with a set of strategies. Each player will have a payoff for each set of alternatives.

Players need not face the same set of strategies. For example, one of the nightclubs might be a comedy club whose manager faces the choice of having an open-mike night or a professional comedian. The two nightclubs would still, at least theoretically, affect each other via their actions in that case, but the payoffs in this event may well change because a nightclub with comedy does not compete as directly with a nightclub with music as would another nightclub with music.

Although games can have a large number of players who each face a continuum of options, it is useful to restrict our analysis to two players with a finite number of options to describe as simply as possible the outcomes using a payoff matrix. A **payoff matrix** is a table that describes the outcome for each player and for each set of strategic choices. Each cell in a payoff matrix is composed of a payoff to each player based on the set of strategies chosen by all players. A two-player payoff matrix for the nightclub game is shown in Table 11.1. To ease our discussion, we will call the players White and Gray rather than player A and player B. The numbers in the white triangles are payoffs to White and the numbers in the gray triangles are payoffs to Gray. For example, the lower left cell in the payoff matrix signifies that when White uses a live band and Gray uses a deejay, White will make \$800 and Gray will make \$100 in profit for the night.⁴ One other point to note is that the payoff matrix is symmetric in this instance. This need not be the case, but it does simplify our initial discussion.

Each firm faces a choice: Should it spring for a live band or should it go with a deejay? Consider the alternatives from White's perspective. White's payoffs depend on what Gray does. If Gray decides to use a deejay, White is better off with a live band because a profit of \$800 is better than a profit of \$500. If Gray decides to use a live band, then White is better off with a live band because a profit of \$400 is better than a profit of \$100. In this game, regardless of what Gray does, White does better by hiring a live

³Nobel laureate Thomas C. Schelling's classic book, *The Strategy of Conflict*, Harvard University Press, 1960 (reprinted in 1980) provides a wonderfully intuitive introduction to strategic interaction and game theory. Although this was written with the goal of providing insights into nuclear arms bargaining, it has wide application in a variety of fields. Schelling described the breadth of the subject nicely: "In economics the subject covers wage negotiations, tariff negotiations, competition where competitors are few, settlements out of court, and the real estate agent and his customer. Outside economics it ranges from the threat of massive retaliation to taking the right of way from a taxi." p. 21.

⁴Payoff matrices can also be shown as ordered pairs in each cell. You may have seen a game described using the notation (800,100) in the lower left cell instead of the shaded triangles used here.

Table 11.1A Payoff Matrix for the Nightclub Game

Profit Payoffs		Gray's Choices	
		DJ	Live Band
White's Choices	DJ	500 / 500	100 / 800
	Live Band	800 / 100	400 / 400

band. White's dominant strategy is to have a live band. A **dominant strategy (DS)** is a strategy that produces the optimal outcome *regardless* of what the other players do. Because the game is symmetric, Gray's dominant strategy is also to have a live band. Both players have a dominant strategy in the game represented in Table 11.1. When each player has a dominant strategy, the resulting game has a dominant strategy equilibrium. A **dominant strategy equilibrium (DSE)** occurs if each player in a game chooses its dominant strategy.

One way to represent this in a game is to note the location of the dominant strategy (if one exists) for each firm at the top of the column or at the start of the row, as shown in Table 11.1B. We will not typically show this (we wish to have you work through the analysis on your own). This table is provided to show how you might "mark up" a game as you analyze it for yourself.

Dominant strategy equilibria provide a nice starting point for our analysis of game theory, but many games do not have a dominant strategy equilibrium. Consider the game represented in Table 11.2. Rather than talk about strategies such as high or low (price, output, or advertising), simply consider White's strategic choices as *I* and *II* and Gray's strategic choices as *A* and *B*.

Consider the alternatives from White's perspective. White no longer has a dominant strategy; White's optimal strategy depends on what it expects Gray to do. If White expects Gray to follow strategy *A*, then White should follow strategy *II*; however, if White expects Gray to follow strategy *B*, then White should follow strategy *I*. Neither strategy is dominant from White's point of view. This does not mean that the game lacks an equilibrium because White can decide on an optimal strategy if it takes into consideration what Gray is likely to do.

By considering Gray's choices, it becomes clear that Gray is more likely to choose strategy *B* than strategy *A* because *B* is a dominant strategy for Gray. Therefore, the optimal strategy from White's perspective is *I*, a strategy predicated on White's analysis of the options facing Gray. The resulting set of strategies, *IB*, is not a dominant strategy equilibrium, but it is a Nash equilibrium, named after the Princeton

Table 11.1B Marking Up the Nightclub Game

Profit Payoffs		Gray's Choices	
		DJ	Live Band
White's Choices	DJ	500 / 500	100 / 800
	DS Live Band	800 / 100	400 / 400

Table 11.2 A Game with a Nash Equilibrium but Not a Dominant Strategy Equilibrium

Profit Payoffs		Gray's Choices	
		A	B
White's Choices	I	5	4
	II	2	3

mathematician/economist John Nash, who first codified this concept in 1951.⁵ A **Nash equilibrium** occurs if every player's strategy is optimal given its competitors' strategies. A set of strategies is a Nash equilibrium if each player is doing the best it can given the strategies of all its opponents.

In the context of Table 11.2, the pair of strategies *IB* is a Nash equilibrium because strategy *I* is optimal for White if *B* is chosen by Gray, and *B* is optimal for Gray if *I* is chosen by White (in fact as already noted, *B* is optimal for Gray regardless of White's choice). Each player's choice is optimal given the other player's choice.

A game may have a Nash equilibrium even if it does not have a dominant strategy equilibrium. A dominant strategy equilibrium requires that each player be able to choose a strategy that is optimal regardless of rival action, but a Nash equilibrium requires that each player be able to choose an optimal strategy given the optimal strategies of the other players.

Nash equilibrium is a broader concept than dominant strategy equilibrium. Any dominant strategy equilibrium is also a Nash equilibrium, but the reverse is not true. Unfortunately, some games will have no Nash equilibrium and some games will have more than one. Table 11.3 depicts a game with no Nash equilibrium. Suppose this game is played a number of times, and each period the player who did not move last period reacts to the player who did move (like chess or checkers). In this instance, the game has no stable solution. To see why, suppose Gray has the first move and chooses *A*. White responds to *A* by choosing *II*. Next, Gray responds to *II* by choosing *B*. Next, White responds to *B* by choosing *I*. Next, Gray responds to *I* by choosing *A*. Then the cycle begins once again. Each player's best choice is the one they did not choose!

Table 11.3 Chasing Your Tail: A Game with No Nash Equilibrium

Profit Payoffs		Gray's Choices	
		A	B
White's Choices	I	0	1
	II	1	0

⁵Many of you will recognize John Nash as the person who was made famous not by winning the Nobel Prize in Economics in 1994, but by Russell Crowe's portrayal of John Nash in the 2001 movie, *A Beautiful Mind* that received four Oscars including Best Picture and Best Director (Ron Howard).

Table 11.4 A Game with Multiple Nash Equilibria

Profit Payoffs		Gray's Choices	
		A	B
White's Choices	I	-3	5
	II	5	-2

By contrast, the game described in Table 11.4 provides an example of a game with two Nash equilibria. This type of game might be encountered when two firms are deciding how to position a new product in a differentiated product market environment. In this conceptualization, both firms realize that, if they do something similar to what the other firm does, they will saturate the market with two similar varieties and thereby not exploit latent demand as effectively as if they each choose to do something different. The pairs of strategies *IA* and *IIB* represent doing something similar in this context, and the pairs of strategies *IB* and *IIA* represent doing something different. The latter two are both Nash equilibria. The critical factor for each firm is to try to figure out ways to ensure that the Nash equilibrium that is preferable to your firm is chosen.

Consider the situation from White's perspective. If White can signal in some fashion that it plans to choose *I*, then Gray will be forced to choose *B*. The goal for White is to figure out some way to signal a commitment to a strategy. If this can be accomplished, the game turns in favor of White. Of course, Gray has the same goal. Gray must figure out some way to signal that it plans to choose *A*. If this is accomplished, then the game turns in favor of Gray because White must then capitulate and choose *II*. The problem arises if both firms signal their preferred strategy (*I* and *A*) prior to receiving the other firm's signal. If each follows through with its signaled strategy, each firm loses money.

The key to achieving a Nash equilibrium in a game such as the one shown in Table 11.4 is in being able to coordinate actions. Unfortunately, there is a wide array of actions that are illegal to coordinate on a formal basis. Firms cannot, for example, formally agree to hold prices at one level in order to coordinate their pricing activities. This is called price fixing and it is illegal according to Section 1 of the Sherman Act. This does not mean that firms cannot interdependently price; it simply means they cannot enter into formal agreements with regard to pricing.

The Prisoners' Dilemma

Many of the games faced by firms in an oligopoly have a common structure. That structure is called the **prisoners' dilemma**. It is worthwhile to describe the prisoners' dilemma using its classic conceptualization before we show how that structure is identical to a number of oligopolistic interactions.

Suppose two individuals, White and Gray, are suspected of jointly committing a crime, and there is weak evidence to support this suspicion. For example, both have been caught trying to fence (sell) goods that were stolen in an armed robbery. The prosecutors would like to convict the suspects on the more serious charge, but the evidence linking the individuals to that charge is weak. The suspects are picked up and put in separate interrogation rooms, where prosecutors provide each suspect with the following options:

- ▶ If neither suspect confesses, both will be convicted of possession of stolen property and receive a 1-year sentence.

Table 11.5 The Prisoners' Dilemma Game

Jail Time (-Years in Jail)		Gray's Choices	
		Don't Confess	Confess
White's Choices	Don't Confess	-1 / -1	0 / -15
	Confess	0 / -15	-10 / -10

- If only one suspect confesses and is willing to testify at the trial of the other, the one who confesses will get probation on the stolen property charge and receive no jail time. The other will receive a 15-year sentence.
- If both suspects confess, both will be put away for 10 years.

The payoff matrix faced by the suspects is summarized as Table 11.5.

The dilemma is that confession is a dominant strategy for both parties. Confession dominates not confessing for White regardless of what Gray does, and confession dominates not confessing for Gray regardless of what White does. Put in terms of the discussion in this chapter, the set of strategies (confess, confess) represents a dominant strategy equilibrium. Certainly both parties would be better off if they could communicate with one another and agree not to confess. And that is the reason that prosecutors try to “sweat” each suspect separately. But even if they could communicate and reach an agreement, there is an incentive to cheat on the bargain. How can you trust the agreement you have made with your accomplice, especially when you are being told that he or she is being offered the same deal and you know how trustworthy he or she is from past experience? It is not surprising then, that such a strategy leads to confessions with sufficient regularity that it is a standard interrogation technique.

Many oligopolistic interactions have the structure of a prisoners' dilemma. The nightclub game in Table 11.1 is a prisoners' dilemma: Both nightclubs could earn \$100 extra in profit by using a deejay rather than using a live band. The model discussed in the previous chapter where firms face a choice of restricting output and acting like a joint profit-maximizing cartel (JπMax) versus cheating on that bargain by expanding output and capturing higher individual profits (Cheat) is easy to model as a prisoners' dilemma. Consider the numbers in Table 11.6. Although the numbers in Table 11.6 look quite different from Table 11.5, the two tables have the same qualitative structure. Cheating is the dominant strategy for each player, but it is not the best that the firms could do. Both firms would be better served if each could agree to restrict their production in order to achieve joint profit maximum, but each has an

Table 11.6 Analyzing Cartel Behavior as a Game

Profit Based on Chapter 10 Cartel Model		Gray's Choices	
		JπMax: Q = 15	Cheat: Q = 20
White's Choices	JπMax: Q = 15	250 / 250	300 / 175
	Cheat: Q = 20	300 / 175	200 / 200

incentive to cheat on the bargain. And, as we already noted, bargains that attempt to restrain trade are illegal according to U.S. antitrust laws, so tacit coordination techniques must be employed.

Some oligopolistic prisoners' dilemma games are more problematic than others. A key determining factor in being able to avoid the dominant strategy of confession is playing the game more than once. The prisoners' dilemma is a dilemma for the prisoners precisely because they are caught only once, and the prosecutors offer them the deal only once. Similarly, entry into a new market segment can be analyzed as a one-shot game. By contrast, firms that are considering the pricing game or the quantity game play that game repeatedly. When a game is repeated, the players can learn from their mistakes and they can learn from their rivals; they can engage in bargaining with their rivals, which often allows them to avoid the prisoners' dilemma.

The Basics of Bargaining⁶

Bargaining is negotiating over the terms of an agreement. When two individuals bargain over the terms of the sale of a house, they engage in explicit bargaining. Games such as these are called cooperative games. A **cooperative game** is a game in which the players can negotiate explicit binding contracts over various states of nature. The negotiation over the sale of a house involves explicit bargaining over what happens under various states of nature, which makes it an example of a cooperative game. Price is the most obvious negotiation point in the sale of a house, and a higher price for the seller means less consumer surplus for the buyer, so you may be surprised to learn that this is a cooperative game. The fact that \$1 more for one player means \$1 less for the other player signifies that this is zero sum game. A **zero sum game** is a game where the sum of payouts is constant. Whether the constant is zero or another number does not alter how such games are analyzed therefore they are often called *constant sum games*. Most games we examine are, by contrast, variable. A **variable sum game** is a game where the sum of payouts for each set of strategies varies. Such games are also called *non-zero-sum games* or *non-constant-sum games*.

Aspects of the house purchase game involve bargaining beyond price. The sales contract will include an explicitly negotiated agreement detailing what happens in the event that a radon, structural, or termite inspection discovers a problem with the property under negotiation, and it describes what happens if one party or the other backs out prior to completion of the sale. This negotiation is in marked contrast to negotiation between firms in an industry over how they will treat one another. Explicit contracts between firms regarding how they will react to various actions (such as pricing agreements) are *per se* illegal.

Although explicit bargaining is common in economics, much of the bargaining that occurs in oligopolistic markets is, by necessity, not explicit. Therefore, we will focus attention on situations where explicit bargaining is not allowed. If the parties involved in the game cannot negotiate and enter into binding contracts, then they are playing a noncooperative game. A **noncooperative game** is a game in which formal negotiation and entering into a legally binding contract is not possible. If firms are involved in a noncooperative game, then they cannot reach explicit understandings regarding aspects of the game under consideration—tacit bargaining replaces explicit bargaining. **Tacit bargaining** is bargaining that is not openly expressed but is implied by actions. Each party observes its rivals' actions and infers from those observations the intent of the rivals, all the while realizing that its rivals are doing the same and interpreting intent based on the others' actions.

⁶This section is based on Chapters 2 and 3 of Thomas Schelling's *The Strategy of Conflict*.

Noncooperative games involve tacit bargaining because explicit bargaining is not allowed. Certain behavior patterns are deemed illegal according to antitrust laws. Firms cannot enter into binding agreements to fix price, restrict output, or not invade each other's territory. This does not mean that they cannot come to a tacit understanding that price should not be cut, or output should not be increased, or the geographic scope of the market should be expanded. Coming to these tacit understandings is the key to achieving mutually beneficial interdependent outcomes. Economists sometimes describe such tacit coordination as achieving *mutual dependence recognized*. Achieving mutual dependence recognized by coming to such understandings involves a good deal of strategic behavior.

Bargaining involves both conflict and cooperation. Firms would like to cooperate in order to restrict output and raise prices, but there is always the incentive to cheat on that cooperative outcome (by undercutting price or increasing output). Bargaining involves conditioning one's behavior on the behavior of others. Bargaining is based on rationality; however, rationality must be considered in a broad sense that includes the idea that it is sometimes rational to appear irrational. One of the ways that a negotiation can be settled in your favor is if it appears to those with whom you are negotiating that you are truly a stubborn, inflexible, and irrational person who would be willing to do something that goes against your best interest just "because." William Land, the president of Polaroid, defended the instant picture segment of the market with such vehemence that potential entrants were kept out of that segment of the market for many years.

Bargaining power does not necessarily come from having a great deal of negotiating latitude. Indeed, the power to constrain one's adversary may depend on your ability to constrain yourself. When management of a company bargains with a union over a wage contract, the management's bargaining position is strengthened by being able to say "I'd like to offer you that, but the board of directors would not allow me to accept such an offer." *Bargaining may also be strengthened if you are constrained in your ability to communicate. More explicitly, it may be strengthened if you are able to send but not receive messages.*

As we already noted, it is important to distinguish between one-shot and multiple-shot (or continuous) negotiations. With continuous games, you can ask: "What will you think of me if I concede on this point?" The same does not hold true for one-shot games. This distinction allows us to compare negotiations over things like pricing with investment decisions. Pricing decisions are much more continuous in nature, while investment decisions are much more a one-shot negotiation. As a result, we expect the prisoners' dilemma to be a much greater problem for investment decisions than for pricing because pricing offers the ability to learn from past behavior.

Behavior can be altered by the effective use of threats and promises. A **threat** involves undertaking an action that harms yourself if your rival does something you do not want your rival to do. You also harm your rival, but the key attribute of a threat is that you make yourself worse off. A **promise** is a commitment made to a second party in a bargain. The difficulty in invoking a threat is that you need to make the threat *credible*. After all, you are proposing to do something that is against your best interest. In this situation, irrationality or stubbornness helps. Threats can be diffused by destroying communication channels, and they can be made more effective by being decomposed into a series of smaller threats. Promises are required whenever there is an incentive to cheat on the bargain. Table 11.7 describes a situation that requires the use of a threat and a promise to propel the solution from the Nash equilibrium to the joint profit-maximizing solution.

The game in Table 11.7 does not have a dominant strategy equilibrium, but it does have a Nash equilibrium. The Nash equilibrium in *IA* is the result of *A* being the dominant strategy for Gray. This equilibrium is not joint profit-maximizing, however

Table 11.7 The Application of a Threat and a Promise

Profit Payoffs		Gray's Choices	
		A	B
White's Choices	I	5 2	4 4
	II	1 1	5 0

(and neither was the Nash equilibrium in the prisoners' dilemma), and White can do better in this situation if it invokes a promise and a threat. Suppose White says to Gray:

I (White) will choose strategy *II* if you choose A.

I (White) will choose strategy *I* if you choose B.

The first statement is a threat because White would be better off choosing *I* when Gray chooses A. The second statement is a promise because White has an incentive to cheat and choose *II* if Gray chooses B. If White is able to communicate the threat and promise effectively, then Gray will choose B because it would rather face the loss of 1 (4 versus 5) rather than the loss of 4 (1 versus 5) that would occur if Gray sticks with A and White follows through with the threat. Of course, such commitments are not necessarily easy to communicate or enforce, especially when evidence of such coordination provides grounds for lawsuits.

With tacit bargaining, the bargain that is struck is determined by its relative unambiguousness. *The bargain is not necessarily the best; it is just the most obvious.* Tacit bargaining involves the use of focal points. **Focal points** are points chosen due to their prominence of conspicuousness. Focal points are often chosen due to their uniqueness or, even more important, their lack of ambiguity. The important skill for a tacit bargainer is thus in formulating the problem. The goal for a player is to formulate the problem so that the "obvious" choice is the one that is favorable to you. By appropriately setting the agenda, the outcome can be swayed in the direction that is favorable to you.

Focal points can be used to provide recognizable bounds on the bargain. Rivals may look at one another and say: "I must stand firm on a 50-50 split of the market because if I back down and allow my rival to capture more than 50 percent, my rival will take that as a signal that I am willing to roll over and play dead." In negotiations involving numerical magnitudes, the negotiations will often end up at mathematically simplistic solutions; in Schelling's words (p. 67):

. . . [i]n bargains that involve numerical magnitudes, for example, there seems to be a strong magnetism in mathematical simplicity. A trivial illustration is the tendency for the outcomes to be expressed in "round numbers"; the salesman who works out the arithmetic for his "rock-bottom" price on the automobile at \$2,507.63 is fairly pleading to be relieved of the \$.63.

Of course, this example is one of explicit bargaining, but the need for a focal point solution is even greater when tacit methods must be relied upon. As we already noted, the outcome is not necessarily the best, only the most obvious.

A final tactic of bargaining is the use of **limited war**. In the context of armed conflict, limited war means "no nuclear war" or "no chemical or biological warfare." The feature of war that makes it limited is that the bounds are qualitative, not quantitative, in nature. In limited war, the question is, Should we pursue each other on this

dimension? rather than How vigorously should we pursue each other on this dimension? In the market setting, limited war means using focal points to place qualitative bounds on the conflict. Oligopolists ask the question, Should we restrict our competition to nonprice dimensions of rivalry and avoid price competition? Of course, this does not always occur. You often see oligopolists maintain discipline on price but not on other dimensions, such as product variety. The ready-to-eat cereal market is an example of rampant nonprice rivalry with almost no pricing rivalry. The players involved have been able to agree tacitly to limit their war to nonprice dimensions. As a result, we have a cereal aisle in the grocery store that has a maddening array of slightly different choices, all at close to the same price.

A General Framework

The many examples provided so far illustrate two fundamental aspects of game theory: interdependence and uncertainty. First, players are interdependent. There are interactions between your decisions as a manager and the decisions of other people. Decisions that you make as a manager affect other players; decisions that other players make affect you. The second fundamental aspect of the game arises from uncertainty. Actions by other players affect you, but you are not certain what their actions will be, and they may not be certain what your actions will be. This uncertainty arises because other players' actions are not entirely predictable; because information is dispersed, you may not have all the information about the other player's context to help you predict their decision; or the other players may be making decisions without full information. Because of this interdependence and uncertainty, managers must act strategically, that is, make decisions based on what they predict other players are going to do and how those other players are going to react to the first manager's decision.

Studying game theory equips managers with a paradigm for studying a situation, for predicting players' actions, and for making strategic decisions. The PARTS framework developed by Brandenburger and Nalebuff is one such paradigm.⁷ PARTS is an acronym for **Players**, **Added value**, **Rules**, **Tactics**, and **Scope**. By examining the game through this paradigm, the authors believe a manager can design a better strategy for his or her personal success and for the success of the firm. The basic elements of the paradigm are as follows:

Players: *Who are the players and what are their goals?* For a particular firm, it would be customers, suppliers, and those that produce complementary or substitute products. It could also be people who work for the company. *What are the players' goals?* For example, an individual may not have profit as the primary motive. *What alternative opportunities do they have? How important is it to them to move quickly? Can they make irrevocable commitments?*

Added value: *What do the different players contribute to the pie?* For success in the business world, the primary question is often not how to take a greater slice of the pie, at the expense of your competitors, but how to take a bigger slice of a bigger pie. Understanding what each player contributes is key. Ideally, a strategy will work to raise your value added or lower the value added of the other players.

Rules: *What is the form of competition?* It could be price, output, cost, quality, R&D, marketing, or some combination of these. *What is the time structure of the game?* For example, it could be repeated or sequential. *Does one player face higher costs for a lengthy negotiation than another? What is the information structure of the game? How much do the players know about one another's positions?*

⁷See John McMillan, *Games, Strategies and Managers: How Managers Use Game Theory to Make Better Business Decisions*, New York, Oxford University Press, 1992, p. 180.

Tactics: *What options are open to the players? Are there possibilities for credible commitments (i.e., those that remove uncertainty)? Can incentives be offered to change the dynamics of the game?*

Scope: *What are the boundaries of the game? For example, it could be the products, the players, or the region or country. Is it possible these boundaries can be expanded or shrunk?*

The discussion in this section provides an introduction into how game theory can help managers design successful strategies. We encourage you to explore this further in your own reading or in coursework devoted to a more complete discussion of its many aspects.

ASYMMETRIC INFORMATION

Thus far in the text, we have assumed that consumers and producers have all of the relevant information required to make the appropriate economic choices under their consideration. But information itself is a product that has a nonzero cost. Information is costly to gather, process, and produce. If we incorporate imperfect information into our analysis, we find that some of the conclusions of our basic models change. To take the simplest example, one of the hallmarks of the perfectly competitive model is that firms have no price-making ability. A single price clears the market; otherwise, consumers would immediately migrate to lower-cost sellers. But first they have to find the lower-cost sellers. If we assume that consumers know only the prices of products at the stores where they currently shop, then each retailer may be able to charge different prices from other stores without losing all of its customers. Even if consumers find a lower price on a product at store *B*, they may not find it worthwhile to go to the other store if that store is out of the way or if they are already at store *A*. Price dispersion results. This is not to say that stores have complete flexibility to charge whatever they want, but it does mean that firms have some price-making ability.

A different problem arises when buyers and sellers have systematically different information about the product. This is known as **asymmetric information**. A market is subject to an *informational asymmetry* if one party or the other tends to be at a systematic informational advantage relative to the other party in the market interaction. Consider three possible market interactions involving asymmetric information:

- Who knows more about the quality of a used car, the buyer or seller?
- Who knows more about the likely future productivity of a worker, the person who is trying to get hired or the person who is doing the hiring?
- Who knows more about how risky a driver is, the person who is buying auto insurance or the person selling auto insurance?

In the first two markets, the seller knows more than the buyer. In the third, the buyer knows more than the seller. Each of these markets is a classic example that highlights how asymmetric information affects markets for a product.⁸

The Lemons Model

George Akerlof was interested in why used cars sell for so much less than new cars. Even a car that is only a couple of months old and that has been driven for only a few thousand miles will inevitably sell at a substantial discount over what was paid for it.

⁸These examples are based on those used by George Akerlof, Michael Spence, and Joseph Stiglitz, who jointly won the Nobel Prize in Economics in 2001 for their inquiries into markets with asymmetric information. An excellent introduction to their influence on economic analysis is provided at the Nobel website: <http://nobelprize.org/economics/laureates/2001/public.html>.

This is the result of asymmetric information. To see this, consider a simple model in which there are four types of cars: new and used, and good (cherries) and bad (lemons). There is no problem in the new car market because there is no asymmetric information in this market. The new car seller and buyer both know that some new cars are lemons, but neither knows whether a particular car is a lemon. The same cannot be said for the used car market.

The quality of a specific used car is difficult to ascertain prior to purchase (and use). A buyer only gets to know the quality of the car he or she purchases over a period of time. The seller on the other hand, has owned the car and knows whether it is a cherry or a lemon. Because buyers cannot tell a good used car from a bad one, both bad and good used cars will sell for the same price (and, of course, the same is true of new cars because buyers and sellers cannot determine the quality of a specific new car).

The price of a used car will *necessarily* be lower than the price of a new car, even in the absence of wear and tear on the car. This assertion is surprising but simple to justify. If it were not the case, then whenever an individual purchased a car and obtained a lemon, he or she would simply turn around and sell it for no loss. This practice would continue until the individual obtained a good car.

To see the problems that arise due to this informational asymmetry, it is worthwhile first to examine the market in the absence of imperfect information. We start with three very simple assumptions.

- a. With perfect information, equal numbers of lemons and cherries are sold.
- b. Lemons sell for \$4,000 and cherries sell for \$10,000.
- c. Five thousand of each type are sold at these prices (so that there are 10,000 total cars sold given perfect information).

The full information equilibrium is efficient in the sense that supply equals demand in each market segment. Consider each segment individually. Each of the 5,000 buyers of good cars places a value of *at least* \$10,000 on a good car (or else they would not have been willing to buy at \$10,000). Similarly, each of the sellers of good cars values these cars *at most* at \$10,000 (or else they would not have been willing to sell at \$10,000). This just means there are gains from trade in the market as a result. A similar statement can be made for the bad used car segment of the market. Each of the 5,000 buyers values the lemon at least at \$4,000, and each seller values their lemon at most at \$4,000. In both instances, the marginal seller and buyer have a valuation equal to the market-clearing price in that segment.

Suppose we impose the information asymmetry already described. Buyers can no longer tell whether a specific car is a lemon or a cherry; therefore, cherries and lemons must sell at the same price. That price should be a weighted average of the prices that would individually clear the market, with the weights being the frequencies of each alternative, f . Given a 50-50 chance at a lemon or a cherry, a price of \$7,000 should emerge:

$$\$7,000 = f_L \cdot P_L + f_c \cdot P_c = 0.5 \cdot \$4,000 + 0.5 \cdot \$10,000$$

This is a reasonable initial expectation. Unfortunately, that expectation is not likely to be matched by reality for a simple reason: Supply is upward sloping.

Consider first the sellers of lemons. The information asymmetry is great news for them because they can unload the cars at a price that includes a premium based on the possibility that their cars are cherries. We had a full information equilibrium at $(Q, P) = (5,000, \$4,000)$ in the lemons market and $(5,000, \$10,000)$ in the cherries market, but $(10,000, \$7,000)$ is *not* likely to be an equilibrium in the combined market because some existing owners of lemons who valued their lemons at higher than \$4,000 will decide to sell when the price they can receive for their (unknown) lemons

increases from \$4,000 to \$7,000. Suppose there are 1,000 such marginal lemon owners (those whose marginal valuation is between \$4,000 and \$7,000). Contrast this with what happens to good used car sellers: 5,000 were willing to sell at a price of \$10,000. Certainly some of these would not be willing to sell at a lower price. Suppose 2,000 of the good used car owners decide not to sell when the price declines from \$10,000 to \$7,000.

What are the actual proportions of lemons and good used cars at this price? Rather than a 50-50 chance at a lemon or a cherry, you are now twice as likely to have a lemon as a cherry (there are 6,000 lemons and 3,000 cherries sold). But this means that the price calculated above is not valid. Instead, the price should be:

$$\$6,000 = f_L \cdot P_L + f_C \cdot P_C = 2/3 \cdot \$4,000 + 1/3 \cdot \$10,000$$

But, of course, this lower price would cause further flight of the good used car sellers out of the market.

Where the market ultimately ends up coming to rest depends on the supply elasticities in the two segments as well as on how demand responds to a change in price in this instance. *The market equilibrium that results may or may not have good used cars being sold, but it certainly will have a higher proportion of lemons sold than would occur in the absence of the information asymmetry.* Put another way, bad used cars drive good used cars out of the used car market; good cars have a higher value to the owners than the price they could obtain in the used car market. As a result, good used car owners tend to hold onto their good used cars. High-quality sellers cannot be distinguished in the market; therefore, some will tend to withdraw from the market.

Demand need not be downward sloping in this instance. Demand depends not only on price, P , but also on average quality, A , in the used car market. We can write this algebraically as demand = $f(P, A)$. As we already noted, average quality in the market will increase as price increases: we write this as $A = g(P)$ with $\Delta g/\Delta P > 0$. As a result, we see two competing effects on demand as a function of price alone:

$$\text{Demand}(P) = f(P, A) = f(P, g(P))$$

Consider how demand changes for a change in price:⁹

$$\begin{aligned} \Delta \text{Demand}(P)/\Delta P &= \Delta f(P, A)/\Delta P + \{[\Delta f(P, A)/\Delta A] \cdot [\Delta g(P)/\Delta P]\} \\ ? &= - + \{ [+] \cdot [+] \} \\ \text{Total effect (of price change)} &= \text{price effect} + \text{quality effect} \end{aligned}$$

The total effect of the price change (slope of demand) is the sum of two components. The first term on the right-hand side is the price effect (the minus sign signifies the *law of downward-sloping demand*) and the second is the quality effect. The quality effect is the product of two bracketed parts: the effect on demand due to an increase in average quality, $\Delta f(P, A)/\Delta A$, and the change in average quality induced by a price change, $\Delta g(P)/\Delta P$. Both of these parts are positive, but the magnitude of each part is not a constant function of price. For very high prices (close to the new car price), there should be little change in average quality because many good used car owners would be willing to sell their car. In this instance, $\Delta g(P)/\Delta P$ is small, and the quality effect is likely to be smaller than the price effect in magnitude. Demand is downward sloping for high prices. But for low prices, the relative magnitudes may reverse and

⁹Those who have had multivariate calculus will note that this is the total derivative of demand with respect to price:

$$d\text{Demand}/dP = df(P, g(P))/dP = \partial f/\partial P + \partial f/\partial A \cdot dg/dP$$

demand would be upward sloping. If this occurs, demand is backward bending due to asymmetric information.

The lemons model is not unique to the used car market. It applies to any market in which the buyer has more difficulty determining the quality of the product prior to purchase than the seller does. Therefore, it is widely applicable to markets in which it is difficult to determine the quality of the product by inspection prior to purchase. It is worth pointing out that the product does not have to be used in order to have a lemons problem. As long as the seller knows more about the product quality than the buyer, this problem may arise.

Using Signals to Avoid the Lemons Problem

Naturally, if you own a good used car, you would wish to be able to persuade a buyer that yours is not a lemon. Of course, if you owned a lemon you would also wish to be able to persuade a buyer that yours is not a lemon. Both types of owners would like to represent their cars as being cherries—and “cross-my-heart” promises do not offer effective proof of an assertion. Nonetheless, various methods have been developed over the years to reduce the information asymmetry in the market. Individual owners can provide repair and maintenance records to allay fears of prospective buyers. Used car dealerships offer warranties and guarantees as well. And new car dealerships make it a point to tell you that they keep only the best used cars (that they receive as trade-ins) and they sell the rest at auction.

Outside the used car market, manufacturers use their brand name and reputation, in addition to warranties, as implicit signals of product quality. Statements such as “One of the many fine products of General Foods” used in an advertisement is meant to convey that General Foods has enough faith in the quality of this particular product to put the reputation of the company behind it. Even if you do not know the specific product, you know other products produced by the company and you can deduce that this product must be equally good or else it would not be in the company’s product portfolio. Otherwise, the company would quickly lose its reputation as a purveyor of high-quality products, an outcome that the company would go to great lengths to avoid. Firms use brand names to tell consumers that the firm’s reputation is on the line. Of course, such reputational effects cannot be stretched into completely new markets. One should not expect a McDonald’s automobile, for example. But it is not surprising to find Taco Bell has extended its brand image from fast food into the Mexican food aisle of the grocery store. Brand image in one market can help establish an identity in a new market, especially if the connection between markets is apparent.

Job Market Signaling

Michael Spence examined the question of how sellers can send signals to buyers that convey information about their product’s quality. Spence was (initially) interested in how individuals who are involved in the job search process can signal potential employers that they will be highly productive employees. This is once again a situation of asymmetric information—the job seeker has a much better understanding of her or his strengths and weaknesses than does the potential employer.

If productivity could be observed, we would expect that workers would simply be offered their marginal revenue product. If productivity is not directly observable, then all workers will be paid the same wage based on average productivity (rather than marginal). Low-productivity workers are being paid more than they are worth, and high-productivity workers are being paid less than they are worth as a result.

A preferable situation would be *separating* low-productivity workers from high-productivity workers prior to hiring. Those with low productivity could be hired for

more menial tasks at a lower wage, and those with high productivity could be hired for more substantive tasks at a higher wage. The difficulty for the firm is to avoid hiring low-productivity workers for the more substantive tasks (because, of course, all workers would like to receive a higher wage). The firm would like to separate workers even though they cannot directly observe worker productivity prior to hiring. Spence argued that educational achievement could be used as a **signal** to separate workers. In particular, *Spence argued that education could act as a separating device because education is differentially costly to obtain.* Education acts as an observable proxy for the unobservable phenomenon called productivity.

Spence provides a very simple model to make his point. Suppose each year of education costs C_L to a low-productivity worker and C_H to a high-productivity worker. The key to Spence’s analysis is that education is more costly to low-productivity workers than it is to high-productivity workers, $C_L > C_H$. Educational cost includes tuition, books, and fees (which we might initially assume are the same regardless of whether the worker is productive or not). But it also includes the opportunity cost of forgone wages. A year’s worth of education may take longer than a year if a student takes fewer courses per semester or if that student fails courses more frequently. It is reasonable to assume that individuals with low productivity have higher opportunity costs as well as higher tuition costs if they spend more time in pursuit of a degree.

Given a difference in productivity as well as a difference in the cost of education, there will be an education level, e , where it is profitable for the high-productivity worker to obtain this level of education and, at the same time, it is not profitable for the low-productivity worker to obtain this level of education. Suppose the firm offers L in compensation to workers with no education and H in compensation to workers who have an education level of e . You can think of L and H as the (present value of) marginal revenue products of the low- and high-productivity workers, respectively:

$$(H - L)/C_L < e < (H - L)/C_H$$

There must be an educational level e that satisfies this inequality because $C_L > C_H$.

If the firm offers this wage choice, then high-productivity workers will choose to attain the education level e and receive H in compensation. At the same time, low-productivity workers will choose to accept L and avoid e . These assertions are easy to prove by simply considering the net benefits with and without attaining education level e for both types of workers:

	Low-productivity worker	High-productivity worker
No education:	L	L
Education level e :	$H - e \cdot C_L$	$H - e \cdot C_H$

The left-hand inequality can be regrouped to obtain: $L > H - e \cdot C_L$. The low-productivity worker will *voluntarily* choose to accept a wage of L . By contrast, the right-hand inequality can be regrouped to obtain: $L < H - e \cdot C_H$. The high-productivity worker will *voluntarily* choose to obtain an education level of e .

It is worth noting that the firm does not directly observe the worker’s productivity when making the job offer. The firm simply offers a choice of jobs: one at a low wage rate for those who have not obtained a certain education level and another at a higher wage rate for those who have. Workers select themselves, via their educational choices, into the high- and low-productivity groups based on the choices they face and on their own attributes. This is called a separating equilibrium. A **separating equilibrium** is one in which high-productivity and low-productivity segments are separated by workers choosing to use a signal.

Notice that education is *purely* a signal in this model. The result is not based on education increasing worker productivity. This simplification is an overstatement, of course, but it is perfectly consistent with the often-heard admonition that it does not matter what you study in college: You can obtain a wide variety of jobs without having majored in a specific subject in college. This cannot be taken to an extreme: You are unlikely to be able to obtain a job in information technology, biotechnology, or engineering without having a major that is closely related to these fields.

Education is not a single signal. Different colleges provide different signals, as do different degrees from the same college. Graduate school provides another set of signals. And grades as well as the courses taken provide further sets of signals. Certainly there is some aspect of obtaining a degree that involves increasing productivity (after all, you must have learned at least *some* practical skills during college), but even in the absence of such concrete skills, education acts as a worthwhile investment simply for its signaling value.

Just as with Akerlof's analysis in the previous section, Spence's analysis is broadly applicable outside the job market. Spence's seminal point is that sellers can signal quality via actions only if those actions are differentially costly to obtain. *Signals are only effective if high-quality sellers have lower signaling costs than do low-quality sellers.* The warranties discussed in the previous section are a case in point. A warranty is useful as a signaling device because a warranty will be differentially costly to provide. If the product is high quality and unlikely to fail, then the amount of warranty service will tend to be smaller than if the product is low quality and subject to failure. The presence of a warranty (or, given the presence of multiple warranties from competing producers, the coverage provided by the specific warranty) provides a signal to the uninformed buyer about the quality of the product under consideration.

Screening in Insurance Markets

Joseph Stiglitz focused attention on insurance markets. Insurance markets face the reverse asymmetry because buyers of insurance know more about their riskiness than sellers of insurance. Specifically, Stiglitz examined how insurance companies can get insured individuals to reveal information via their insurance choices. The menu of contracts offered by the insurance company can act to screen prospective consumers. By providing individuals with an array of choices, individuals sort themselves into risk categories in much the same way that jobseekers sort themselves into productivity categories in Spence's model.

The price of insurance is based on an insurance company's best guess regarding the riskiness of individuals who will become policyholders. Insurance companies analyze statistics on loss rates for different classes of customers in order to estimate these probabilities. Of course, these estimates, even if accurate for the group as a whole, cannot take into consideration variations in riskiness within a group that has been lumped together. Not all sixty-five-year-old female nonsmokers are equally healthy (and the same would be true if we further narrowed the category). If an insurance company writes life insurance policies for such individuals, they would base their premium on actuarial information for this group of individuals. Unfortunately, this policy will be of greatest interest to the highest risk individuals within the category and will be of least interest to the lowest risk individuals within the category. The distribution of women that choose to purchase the insurance may well become skewed toward the higher-risk members of the group.

Because the insurer cannot distinguish between high-risk and low-risk individuals, it must write a policy to cover the average individual. If this policy costs too much for the low-risk individuals (i.e., they have higher utility by not insuring), then the only individuals to purchase insurance will be high-risk individuals. This is the adverse selection problem. The **adverse selection** problem occurs if the market becomes skewed because of asymmetric information.

Insurance companies set a single price for a given risk category, but those who purchase the policy may well make the policy unprofitable at that price. Therefore, insurance companies are forced to increase their premiums, which in turn will create further adverse selection problems.

The lemons model is also an example of the adverse selection problem. The informational asymmetry causes a single price to prevail in the market. The used car market is therefore skewed toward an average quality level that is lower than would occur in the absence of the information asymmetry. *The situations in the used car market and the insurance market are quite similar despite having the opposite informational asymmetry:*

Insurance: *Low-risk buyers* are less likely to participate in the market because the price they **are charged** cannot accurately reflect their attributes.

Used car: *High-quality sellers* are less likely to participate in the market because the price they *receive* cannot accurately reflect their attributes.

In both instances, the resulting market has a lower quality due to the information imperfection.

Dealing with the Adverse Selection Problem

The adverse selection problem can be reduced in several ways. One method involves the imposition of a government edict requiring all to purchase insurance. All individuals are required to have automobile insurance in many states, for example. Another method involves public provision of universal coverage such as Medicare for the elderly. This insurance avoids the adverse selection problem because everyone is selected (by government edict). A similar argument holds for group health policies. By including all members of a group (all university employees, for example), group members are able to obtain insurance for a lower premium than they would be able to obtain on the open market. The difference between a group policy and an individual policy purchased on the open market is that there is no self-selection in the group policy (all members of the group are covered), but an individual policy is subject to the adverse selection problem.

Screening is a market-oriented method that insurance companies use to deal with the adverse selection problem. Insurance companies can offer a choice of alternative policies at different price points. Some policies will be more expensive but more inclusive, while others will be more restrictive but also less expensive. If the policies are structured appropriately, they create conditions under which the insured parties self-select into groups that avoid the adverse selection problem. This is known as **screening**. A set of insurance policies *screens* individuals if they self-select into different risk classes. Screening occurs without government intervention. The offerings themselves create conditions that induce the insured parties to reveal their risk class via their policy choice. Two common methods of creating such policies are through the use of coinsurance and deductibles. **Coinsurance** is insurance that provides less than full coverage to the insured party. That part not covered by the insurance company is covered by the insured party. Coinsurance typically provides a percentage of coverage; for example, Blue Cross/Blue Shield typically offers 80 percent coverage for health insurance. The policyholder pays the other 20 percent. A **deductible** is a policy that requires the insured party to cover the first X dollars of any loss. Beyond this loss, the insurance company typically pays 100 percent (although co-pay percentages are also possible in this instance).

By providing policy options that include full coverage as well as less-than-full coverage at a lower cost, insurance companies can screen policyholders into groups with similar risk profiles. High-risk individuals expect to have to pay the full deductible. Low-risk individuals expect to be healthy and do not expect to pay the full deductible

on average. Insurance companies can design the choice of premiums charged and deductibles offered so that the expected cost of a high-deductible/low-premium plan is lower for the low-risk individuals, and the expected cost of a low-deductible/high-premium plan is lower for the high-risk individuals. The value of this strategy is that it provides a market solution to the market failure imposed by asymmetric information.

Moral Hazard and Principal-Agent Problems

Insurance markets have another inherent problem. If individuals are insured against losses, they may alter their behavior because they know their losses are covered. This is called the moral hazard problem. The **moral hazard** problem occurs if, as a result of having insurance, the insured party undertakes actions that either increases the magnitude or probability of payment by the insurance company. The moral hazard problem is based on hidden action, while the adverse selection problem is based on hidden information.

The moral hazard problem is quite common and applies to many types of insurance. If individuals are fully insured against the theft of their automobile, they may well be less vigilant about locking their cars. Individuals may well seek a doctor's opinion for an ailment if they have medical insurance, but they may not seek a doctor's opinion for the same ailment if they have to pay for the doctor's visit out of pocket. Similarly, if individuals face a co-payment for a visit to the doctor's office but are fully covered for emergency room visits, then the individuals will tend to use the emergency room more than they would in the absence of this incentive structure.

The moral hazard problem occurs due to asymmetric information. The problem occurs because the actions or behavior of the insured party is hidden. Suppose an insured party's behavior was observable without cost and insurance contracts could be rewritten without cost. In this event, the insurance premium could be adjusted (without cost) to maintain an actuarially fair premium (premium = probability of loss · magnitude of loss). Of course, this is not possible. Policies cannot be rewritten without cost but, more important, insurance is based on the probability of an adverse outcome occurring. If that adverse outcome occurs (if the house burns down), it is not immediately clear to the insurer whether this occurred because the insured party began to be less careful about his or her actions once insured or because of bad luck. If "bad luck" occurs time after time (if an individual wrecks several cars, not just one), then one may become more convinced that this is due to the altered behavior of the insured than due to bad luck.

The moral hazard problem is a specific example of the more general problem called the principal-agent problem. The **principal-agent problem** occurs when the principal cannot monitor the actions of the agent who is charged with making decisions on behalf of the principal. This is also often called the *agency issue*. The classic example of the principal-agent problem involves the conflict that arises between ownership and control of publicly held companies. The principals are the shareholders and the agents are the managers of the firm. If the owner is also the manager, then there is no problem because the goal of one is the goal of the other. If there is a separation of ownership from control, then it is possible that managers can pursue goals that are inconsistent with the owners' goals. The moral hazard problem can be viewed as a principal-agent problem if we consider the insured party as the agent and the insurer as the principal. The agent (insured individual) is able to undertake action that is unobservable by the principal (insurer).

We can assume that the owners (shareholders) wish to have the maximum value for their shares. The problem is that managers may not share the same goal as do owners. Economists often use the term *managerial utility* to describe the sets of factors that are of interest to managers. Elements in the managerial utility function may include objectives such as maximizing growth, maximizing sales, or maximizing compensation.

Further attributes of the managerial utility function may include an interest in perquisites such as lavish offices and a private jet or other fringe benefits. Managers may also wish to avoid effort as well as risk. The goals of managers may therefore not be consistent with those of owners.

The asymmetric information inherent in the principal-agent problem is that the principal cannot monitor the actions of the agent. Owners cannot monitor the actions of the managers. Owners do not know whether the level of performance obtained by the managers represents the best that could have been achieved under the given market conditions. The managers have an insider's view of the situation and can, to a certain extent, paint the picture they wish to paint regarding the firm's performance. (Managers are responsible for creating quarterly and annual reports [which allow the firm to communicate information with shareholders]. Management decides what to highlight in such communications, and the way issues are framed certainly will present management in the best possible light. This happens even in the absence of fraudulent behavior on the part of management in reporting financial data.)

Corporate raiders look for firms that are underperforming because of poor management. The threat of a takeover raid acts as a constraint on managerial utility maximization. If a takeover occurs, then senior management as well as assets owned by the firm are likely to be the targets. If raiders are successful, then they will replace the old managers with new managers who are more willing to pursue shareholders' objectives. Managers have a variety of methods of combating corporate raiders, including the use of poison pills and golden parachutes.¹⁰ A **poison pill** is a device used to reduce the possibility of a hostile takeover by increasing the takeover cost. This typically occurs through the issuance of new preferred shares with severe redemption provisions. A **golden parachute** is a provision in an executive's pay package that has substantial severance compensation in the event of termination. Each of these tactics increases the cost of takeover and thereby reduces the probability of takeover.

The informational asymmetry that exists between managers and owners typically also exists between managers and corporate raiders. Raiders also face the difficult task of buying a large enough block of stock to exert control without substantially running up the stock price.

The informational issues are especially severe if the product produced by the firm is subject to market fluctuations and if information about those fluctuations is not available. It is not clear in this instance if substandard performance is due to shirking on the part of the agent or a downturn in the market for the product being produced.

We see that the market provides bounds on the ability of managers to pursue goals that are at odds with those of the principals. Another way to resolve the principal-agent problem is to create mechanisms that reduce or eliminate the disparity in goals between principal and agent. This can be accomplished by structuring the managerial compensation package so that the manager's goals become those of the owners.

Creating Incentive-Compatible Mechanisms

An important question in the area of management compensation is how to design compensation packages that help align the goals of the agent with those of the principal. If productivity is observable, there is no problem; the problem occurs because effort is not typically observable. In this situation, owners would still like to have managers pursue the owners' objectives, but managers may be able to shirk without being caught. The owners would like to create a structure that induces the manager to pursue the owners' goals even in the absence of being able to measure the manager's

¹⁰Read about these and other mechanisms used by management at <http://www.thecorporatelibrary.com>.

effort. *The problem boils down to creating incentive structures that stimulate effort when effort is not directly observable.* If this can be accomplished, the manager's objectives will become more consonant with the owners'.

Managerial compensation packages have become increasingly sophisticated in being able to align the goals of the principal with those of the agent. Owners wish to have managers maximize profits, so it is natural to tie compensation to profits. The problem with this strategy is that firm profits typically vary due to factors beyond the control of management. In such a situation, the managerial compensation package should have a fixed component as well as a component that is tied to the profitability of the firm.

Profit sharing is one way to create a managerial incentive to maximize profits. If a manager is offered a small base salary together with a fixed percentage of the profits of the company, then the manager's total compensation package is tied to the profit performance of the company. If the profit incentive is large enough (i.e., if the percentage provided to the manager is large enough), then the manager will choose to maximize profits because doing so maximizes managerial utility. In performing these calculations, the manager must balance the disutility of increased effort (i.e., not pursuing the manager's alternative objectives) and the disutility associated with a riskier compensation package with the higher expected income obtained under the profit-sharing plan (relative to a flat salary with low effort).

An alternative to profit sharing is making *stock options* part of the compensation package. The manager is provided with a base salary together with the option to purchase a fixed number of shares of stock at a fixed price at some point in the future. The price at which the stock can be purchased in the future (the exercise price) may be sufficiently high that the likelihood that the option will be exercised is small—and therefore the option may have a small value at the time it is issued. If the stock price rises dramatically and overtakes the exercise price, the option's value is the difference between the market price and the exercise price. Stock options can make managers very wealthy, and they did just that for many managers during the bull market of the late 1990s.

Of course, if the market as a whole is falling, then it is difficult to obtain significant stock price increases even if managers do everything right and act in perfect accord with the goals of the principals. Stock options and profit sharing will not perform as well under these circumstances. In both instances, part of the manager's compensation is due to factors that are beyond the manager's control. This has given rise to the use of *indexed stock options*, which is a stock option that has an exercise price that is not fixed but rather is tied to a broad stock index. If indexed stock options are used, managers are rewarded if their firm does well relative to the market as a whole. This reduces some of the risk associated with stock options and profit sharing because it ties the managers' compensation more closely to factors that are under the managers' control.

Similar strategies occur in writing insurance coverage to reduce the moral hazard problem. Insurance companies may wish to avoid full-coverage insurance because it completely removes the risk to the insured party. If a firm requires a deductible or coinsurance, it forces the insured party to bear at least some of the risk. By facing some risk, the insured party will take "some" level of care. Many prescription drug benefits packages, for example, require differential co-payments for name brand versus generic drugs. If an individual must pay \$50 for the name brand but only \$20 for the generic version of the same drug, then the individual may well opt for the generic version, thereby saving money for both the insured and the insurer.

Another tactic is to provide lower rates to individuals who have devices that automatically reduce the probability or magnitude of loss. Smoke detectors and sprinkler systems reduce the loss due to fire. Such devices do not deal with the underlying moral hazard problem (that the individual may be less concerned about the risk of fire damage if she or he is fully insured), but it does reduce the problem induced by moral hazard.

The Solution

“Let me understand you on this, Erica. You’re telling me that I should construct some sort of payoff matrix to determine the possible consequences to me and my competitors of future price wars?”

“That’s it, Dad. As I was explaining to you with the prisoners’ dilemma example, if we can come up with what it might be worth to keep the prices constant versus lowering these prices, perhaps we can come up with a solution that will be helpful to you and your competitors.”

“I get it. That’s what we noneconomists sometimes call a win-win situation. But I don’t think that it is realistic to create a payoff matrix, as you call it. Whatever I put down, it’ll just be simply made up. Please don’t think I’m criticizing this theory. Just like you, I find this fascinating. In fact, some of the concepts that I learned when you talked about game theory actually have given me an idea. I’ve been thinking for some time now that I want to sell this business and retire. Maybe I should consider auctioning off my business to my two closest competitors. They of all people would know that I’m situated in a great location, which is worth as much as my store itself, maybe more. I could get an assessment of my property and building, and use that as the starting point for the auction. I think the dominant strategy of both my competitors would be to bid up the price closer to their private value, which I know would be much higher than any appraised value that I could get.”

“How do you know this, Dad? You sound pretty sure of yourself,” Erica remarked.

“Well, I have some friends in the Rotary Club, and they’re telling me that they heard from very reliable sources that Wal-Mart wants to open a superstore in this area and that my property and the land right next to mine offer a potential location for them. If these friends have heard this rumor, maybe my competitors have, too. So they’ll be thinking that they could resell my land to Wal-Mart for a good price when the time comes.”

“That’s a great idea, Dad, and by the way, that’s a good use of the game theory concept of dominant strategy. But why don’t you simply wait and sell your land directly to Wal-Mart when the time comes?”

“Well, perhaps you learned in your economics class other concepts besides those used in game theory, such as good old-fashioned *risk*. What if our town, like a few others, decides to try to block Wal-Mart from opening a new superstore and is in fact successful? What if Wal-Mart decides to open in another location and still worse, another location close to ours ... you get the picture?”

“Hey, Dad, I’m pretty impressed with your foresight. But what will you do when you retire?” Erica asked.

“Well, I’ve always wanted to be a teacher. I’m thinking of going back to the university to get an MBA as a sort of credential. Then, I’m sure there are colleges around here that believe their students would benefit from my 35 years of business experience.”

SUMMARY

An inherent uncertainty exists when dealing with oligopolistic markets that does not exist with perfect competition, monopoly, or monopolistic competition. This uncertainty is based on the interdependence of firms within oligopoly markets. Each firm’s actions are predicated on how that firm expects other firms to react. Put another way, each firm’s actions are based on assumptions regarding rival behavior. Common behavioral assumptions include the following: rival firms will match price changes, rival firms will maintain their current prices, rival firms will maintain their current output, rival firms will match price decreases but not price increases, and rival firms will always choose the option that produces the worst outcome for their rivals. Each of these (and many other) behavioral assumptions lead to a different deterministic models of oligopoly interaction. Each implies different market outcomes because each is based on different presumed behavior patterns: Different behavior leads to different outcomes.

Oligopolistic interaction can be modeled using game theory. Games are most easily examined if we restrict our analysis to two players. Each player has a set of possible strategies they can follow. The outcome, or payout, to each player depends not only on which strategy she or he decides to pursue but also which strategy her or his rival decides to pursue. A payoff matrix describes the set of strategies and payouts available in the game.

Sometimes one strategy provides the highest payout regardless of what rivals do. This is called a dominant strategy. If both firms have a dominant strategy, then the result is a dominant strategy equilibrium. Even when there is not a dominant strategy equilibrium, there may be a

“best” outcome for the game. A Nash equilibrium is such a best outcome. A Nash equilibrium occurs if every player’s strategy is optimal given its competitors’ strategies. The problem with Nash as an equilibrium notion is that some games have no Nash equilibrium and some games have more than one.

Many oligopolistic interactions have the structure of a prisoners’ dilemma. The prisoners’ dilemma is a game in which the option of “confessing” dominates “not confessing” for both players, but confessing leads to a worse outcome for both than if both could decide not to confess. In the context of oligopolistic price and quantity setting, for example, the cartel solution has a higher price and lower output than the individual profit-maximizing solution; therefore, there is an incentive for each party to cheat on the cartel solution.

The prisoners’ dilemma is a bigger problem for an oligopolistic rivalry that is “played” once than for those rivalries played over the long term. When oligopolists confront the prisoners’ dilemma on a continual basis, it is possible to learn from their mistakes and to learn from their rivals. They can engage in bargaining with their rivals.

Bargaining is negotiating over the terms of an agreement. Bargaining can be explicit and enforceable, or it can be tacit and formally unenforceable. Many games played by oligopolists are, by necessity, tacit in nature. When such bargaining occurs, the outcome depends critically on the prominence or conspicuousness of that outcome. Focal points are critical to achieving oligopolistic agreement, and the important skill for tacit bargaining is in being able to set the agenda so that the obvious outcome is the one most favorable to your firm.

This chapter has also scratched the surface of an important topic: the economics of information. Information is itself costly to obtain and, as a result, individuals are not perfectly informed despite assumptions to the contrary. When we explicitly consider how markets respond to information imperfections, some of the cornerstone conclusions of the market model are turned upside down. This is especially true when we consider situations when there is an informational asymmetry between buyers and sellers in a market. Three economists jointly shared the Nobel Prize in 2001 for their work in the area of asymmetric information. An informational asymmetry exists when one side knows more than the other in a transaction. One might expect that sellers are more informed than buyers involved in a market transaction, but this is not necessarily the case.

The lemons model developed by George Akerlof involves sellers with more information than buyers. In this setting, an adverse selection problem occurs as the market for used cars becomes skewed toward low-quality cars (lemons) because buyers are not able to distinguish a good used car from a lemon. As a result, used cars sell for a discount purely based on the informational asymmetry that exists in this market.

The lemons problem can lead to markets in which only low-quality cars will be sold, or it may simply lead to a market in which the proportion of low-quality cars is higher than would be sold if buyers and sellers had the same information. A secondary implication of the lemons model is that demand may be backward bending because buyers will recognize that average quality declines as price declines. If used car prices decline, consumers will demand more due to the law of downward sloping demand. But this decline in used car price increases the odds of obtaining a lemon.

Sellers can undertake various actions to persuade buyers that their car is not a lemon. These actions provide signals that the product in question is high rather than low quality. For example, if a used car dealer offers a warranty on one vehicle and sells another “as is,” then the warranty acts as a signal that the seller stands behind the product. Manufacturers offer the same assurance via warranties and by backing up products with their brand name.

Michael Spence examined the issue of signaling more closely. He argued that a signal is effective only if it is differentially costly to obtain. The used car warranty is a good signal because the buyer realizes that the warranty will be costly to the seller to provide if the car for sale is a lemon but not if it is a good used car.

Spence initially examined signaling in the labor market. High-productivity workers are difficult to distinguish from low-productivity workers prior to hiring (and in many instances, for some time after hiring as well). Spence argued that education acts as a separating device because education is differentially costly to obtain. High-quality workers are likely to be able to obtain a certain degree of education at lower cost than low-quality workers. In this event, prospective workers can be offered a choice of jobs: a low-paying job with no education or a higher-paying one with a certain level of education. If these contingent job offers are chosen

appropriately, low-quality workers will self-select into low-paying jobs and high-quality workers will self-select into the high-paying jobs just by obtaining the educational signal.

Joseph Stiglitz examined the reverse asymmetry. In insurance markets, the buyer knows more than the seller because the buyer of insurance has a better idea about his or her behavior than the seller does. Once again, this leads to adverse selection because only those individuals who most need the insurance are likely to purchase the insurance. This will skew the distribution of risk in the market relative to what would occur in the absence of the asymmetric information.

If an insurance company offers multiple options for insurance, then these options can act to screen individuals into risk categories. Low-risk individuals will tend to choose insurance that is less expensive but has a deductible or involves a co-payment. High-risk individuals will tend to choose more expensive and more inclusive coverage.

If individuals alter their behavior as a result of being insured, a moral hazard problem arises. Individuals who are fully covered in the event that their car is stolen may leave the car unlocked when they go into the grocery store, for example. Leaving the car unlocked increases the risk of theft, and the theft may be due to the altered behavior of the insured individual.

The moral hazard problem is a specific example of the principal-agent problem. The principal-agent problem occurs because the principal (the person who has hired an agent to do a task) cannot (fully) monitor the actions of the agent. This allows the agent to pursue objectives that may be at odds with those of the principal. The classic example of the agency issue is the problem faced by shareholders in monitoring the actions of managers of a firm. The separation of ownership (principal) from control (manager) leads to the ability of managers to pursue their own objectives. Managers do not have complete latitude in this regard—if they stray too far from the owner's wishes, they may find themselves without a job either because they are fired or because their firm is acquired by corporate raiders.

The principal-agent problem can be reduced by the appropriate choice of a managerial compensation package. If the goals of the owner are instilled in the manager via the incentive structure built into the compensation package, then the manager's interests will coincide more closely with the owner's. Profit-sharing plans and stock options are two common mechanisms for achieving this goal.

IMPORTANT CONCEPTS

Adverse selection. A type of problem that occurs if the market is skewed because of asymmetric information. (p. 448)

Asymmetric information. A market situation in which one party in a transaction has more information than the other party. (p. 443)

Coinurance. Insurance that provides less than full coverage to the insured party. That part not covered by the insurance company is covered by the insured party. (p. 449)

Cooperative game. A game in which the players can negotiate explicit binding contracts over various states of nature. (p. 439)

Deductible. A policy that requires the insured party to cover the first x dollars of any loss. Beyond this loss, the insurance company typically pays 100 percent (although co-pay percentages are also possible in this instance). (p. 449)

Dominant strategy. A strategy that produces the optimal outcome regardless of what the other players do. (p. 435)

Dominant strategy equilibrium. A type of equilibrium that occurs if all players in a game have a dominant strategy and all choose that strategy. (p. 435)

Focal points. Points chosen due to their prominence, conspicuousness, or lack of ambiguity. (p. 441)

Game. Players making strategic decisions from an available set of options. (p. 433)

Golden parachute. A provision in an executive's pay package that has substantial severance compensation in the event of termination. (p. 451)

Limited war. A war in which the boundaries of the conflict are qualitatively rather than quantitatively delineated. (p. 441)

Moral hazard. A type of problem that occurs when the insured party undertakes actions that he or she would not undertake in the absence of insurance. (p. 450)

Nash equilibrium. A type of equilibrium that occurs if every player's strategy is optimal given its competitors' strategies. (p. 436)

Noncooperative game. A game in which formal negotiation and entering into a legally binding contract is not possible. (p. 439)

Payoff matrix. A table that describes the outcome to each player for each set of strategic choices. (p. 434)

Payoffs. The outcomes associated with a set of strategies. Each player will have a payoff for each set of alternatives. (p. 434)

- Players.** The decision-making units that play games. Each player chooses among various strategies or options. (p. 434)
- Poison pill.** A device used to reduce the possibility of a hostile takeover by increasing the takeover cost. Typically this occurs through the issuance of new preferred shares with severe redemption provisions. (p. 451)
- Principal-agent problem.** A type of problem that occurs when the principal cannot monitor the actions of the agent who is charged with making decisions on behalf of the principal. Also called the *agency issue*. (p. 450)
- Prisoners' dilemma.** A two-person variable sum game in which the dominant strategy equilibrium is not optimal for both players. (p. 437)
- Profit sharing.** A way to create a managerial incentive to maximize profits. If a manager is offered a small base salary together with a fixed percentage of the profits of the company, then the manager's total compensation package is tied to the profit performance of the company. (p. 452)
- Promise.** A commitment made to a second party in a bargain. (p. 440)
- Screening.** A way to allow individuals to self-select into different insurance-risk classes. (p. 449)
- Separating equilibrium.** A situation in which high-productivity and low-productivity segments are separated by workers choosing to use a signal. (p. 447)
- Signal.** An undertaking that is less costly for high-quality agents to invest in than low-quality agents. (p. 447)
- Stock option.** The option to purchase a fixed number of shares of stock at a fixed price at some point in the future. (p. 452)
- Strategy.** An option available to a player. (p. 434)
- Tacit bargaining.** Bargaining that is not openly expressed but is implied by actions. (p. 439)
- Threat.** Undertaking an action that harms yourself if your rival does something you do not want your rival to do. You also harm your rival, but the key attribute of a threat is that you make yourself worse off. (p. 440)
- Variable sum game.** A game where the sum of payouts for each set of strategies varies, also called a *non-zero-sum game* or *non-constant-sum game*. (p. 439)
- Zero sum game.** A game where the sum of payouts is constant, also called a *constant sum game*. (p. 439)

QUESTIONS

1. Is the prisoners' dilemma more of a problem for a one-shot or a repeated game?
2. Is a prisoners' dilemma game a zero sum or variable sum game?
3. Is a prisoners' dilemma game more of a problem for cooperative or noncooperative games?
4. Is it easier for a rational or irrational person to make a credible threat?
5. Why must much of the bargaining in economic situations be done on a tacit basis?
6. Why are focal points important for noncooperative games?
7. Why is it particularly important to be the party who sets the agenda in situations involving tacit negotiation?
8. What attribute about a war makes it a limited war? How does the concept of limited war relate to oligopolistic rivalry?
9. Suppose you are asked to write down the name of a single card in a standard deck of playing cards. You cannot ask your neighbor what he or she is writing down, but you wish to write down the same card. What card would you write down?
10. The responses to Question 9 provide an example of what game theory concept?
11. Why might demand be downward sloping in a market with imperfect information even though the market is otherwise perfectly competitive?
12. Is there an informational asymmetry in the new car market?
13. If there is asymmetric information, do sellers always have better information than buyers do?
14. Which side has the informational advantage that is the key to the *lemons* model?
15. Why might demand be upward sloping in used car markets?
16. Which effect dominates, the price effect or the quality effect of a price change if demand is upward sloping?
17. What does Spence argue is required for a signal to be effective?
18. Who has more information in the job market: the job seeker or the potential employer?
19. How can education act as a separating device?

20. In Spence’s model, does the educated job seeker receive higher wages because education increases the worker’s productivity?
21. How do warranties and guarantees act as signals?
22. Does the buyer or seller of insurance typically have better information?
23. How is the menu of insurance contracts offered by an insurance company like the menu of wage choices available to individuals hired by a company?
24. Is an insurance policy with a high deductible aimed at low-risk or high-risk individuals?
25. Is the *lemons* model an example of the adverse selection problem?
26. Provide two examples of incentive-compatible mechanisms that reduce the principal-agent problem.
27. Does advertising have signaling value?

PROBLEMS

1. In your answer to each part below, include references to specific games discussed in the chapter as appropriate.
 - a. If a game has a dominant strategy equilibrium, does it have a Nash equilibrium?
 - b. If a game has a Nash equilibrium, does it have a dominant strategy equilibrium?
 - c. If one firm has a dominant strategy, can another firm take advantage of that fact in deciding on its optimal strategy?
 - d. Can a game have more than one dominant strategy equilibrium? Can a game have more than one Nash equilibrium?
2. Consider a small town that is served by two grocery stores, White and Gray. Each store must decide whether it will remain open on Sundays or whether it will close on that day. If both stores decide to close, then each has monthly profits of \$21,000. However, if one is open and the other is closed, the open one has profits of \$25,000 and the closed one has profits of \$17,000. If both remain open on Sundays, then each has profits of \$20,000.
 - a. Why might profits be lower if both are open on Sundays than if both are closed on Sundays?
 - b. Place payouts in the appropriate cells of the payoff matrix shown below.
 - c. Does either firm have a dominant strategy? Is there a dominant strategy equilibrium? Is this strategy joint profit maximizing?
 - d. Is the equilibrium discussed in part c likely to be stable over time? In particular, what might firms do to alter this outcome?
 - e. Is this an example of a prisoners’ dilemma?

Monthly Profit Payoffs		Gray’s Choices	
		Open Sundays	Closed Sundays
White’s Choices	Open Sundays	8000	7000
	Closed Sundays	3000	6000

3. Consider another small town that is served by two grocery stores, White and Gray. Each store must decide whether it will remain open on Sunday or whether it will close on that day. Monthly payoffs for each strategy pair are as shown in the table below.

Monthly Profit Payoffs		Gray’s Choices	
		Open Sundays	Closed Sundays
White’s Choices	Open Sundays	4000	7000
	Closed Sundays	10000	12000

- a. Which firm is the most profitable in this market?
 - b. Does the firm you identified in part a have a dominant strategy?
 - c. What should this firm do? Explain your answer.
 - d. Is this an example of a prisoners' dilemma?
4. Suppose you are a purchasing manager for an electronics equipment firm. Your job is to obtain component parts at the best cost from several suppliers and to make sure those parts are received in a timely fashion. If you consider your job as a game, is it a cooperative or noncooperative game?
 5. Why are focal points important for firms involved in tacit bargaining? What is the key skill for a tacit bargainer?
 6. Suppose the prices in each segment under perfect information are as shown in the table below.

Perfect Information Prices	Seller's Reservation Price	Marketing Clearing Price
Lemon: bad	\$3,000	\$4,000
Cherry: good	\$8,000	\$10,000

A reservation price is the lowest price a seller will accept. Also assume that buyers are risk-neutral and therefore are interested only in the average quality of the cars sold in the used car market. The lemons model assumes an equal number of lemons and cherries in the initial analysis in order to simplify calculations. Consider more generally what happens based on different percentages of lemons in the market.

- a. What percentage of lemons and cherries are sold if the price is above the cherry seller's reservation price of \$8,000? What percentage of lemons and cherries are sold if the price is below the cherry seller's reservation price of \$8,000?
- b. Suppose 10 percent of used cars are lemons. What will be the market-clearing price given asymmetric information? What portion of used cars sold will be lemons?
- c. Suppose 25 percent of used cars are lemons. What will be the market-clearing price given asymmetric information? What portion of used cars sold will be lemons?
- d. Suppose 50 percent of used cars are lemons. What will be the market-clearing price given asymmetric information? What portion of used cars sold will be lemons?
- e. Given the reservation and market-clearing prices in the table above, what is the highest proportion of lemons in the used car market, s , that will produce a market equilibrium in which all used cars are sold? What happens if the proportion of lemons in the used car market is higher than s ?
- f. Suppose that, instead of the good used car owners having a reservation price of \$8,000, they had a reservation price of \$7,000 but all other prices are as described in the table in Problem 6. How would your answer to part e change in this instance?

Problems 7 and 8 introduce the more realistic assumption that supply is upward sloping. The important segment in this instance is the high-quality segment of the market because the low-quality segment will certainly be willing to unload their lemons given the blended price that results due to the asymmetric information. As a result, we assume that all lemons will be sold given a market with asymmetric information; the question is, How many of the cherries will be sold under these conditions?

Reader's note: Problem 7 is difficult, even with the hints. You may wish to consider assigning it in two parts: parts a to d on one day, then go over the answers so that student can then solve parts e to g.

7. Assume that one-third of all used cars are lemons. If the supply of good used cars is upward sloping, then it may be the case that the blended price is lower than the reservation price for some good used car owners. Such owners would not sell. Let G represent the fraction of good used car owners who actually sell their good used car, $0 \leq G \leq 1$. Parts a to d of this problem examine the implications of selling only some good used cars on the portion of lemons sold.

- a. If all good cars are sold ($G = 1.0$), what fraction of all used cars are sold? What is the fraction of sold cars that are lemons?
- b. If half the good cars are sold, $G = 0.5$, what fraction of all used cars are sold? What is the fraction of sold cars that are lemons?
- c. If one-quarter of the good cars are sold, $G = 0.25$, what fraction of all used cars are sold? What is the fraction of sold cars that are lemons?
- d. If a general portion G is sold, what fraction of all used cars are sold? What are the fraction of sold cars that are lemons? Make sure that your algebraic answers are consistent with your specific answers in parts a to c.

Suppose the perfect information prices for lemons and cherries are $P_L = \$5,000$ and $P_H = \$10,000$. Consider how good used car owners view their cars. Suppose no owner would be willing to sell at a price of \$6,000, and all good used car owners would be willing to sell at a price of \$9,000. (This contrasts with Problem 6, where we assumed that all would sell at a price of \$8,000.) If we assume that good used car owners are uniformly distributed between these upper and lower bounds, then we can describe the supply curve as being linear between these two endpoints. We can then write the inverse supply of good used cars as a function of G , the fraction of good used cars sold, as: $P_S(G) = 6,000 + 3,000G$. In this context, P_S denotes the price at which good used car *sellers* are willing to sell their car.

- e. Given that one-third of used cars are lemons and the full information prices of $P_L = \$5,000$ and $P_H = \$10,000$, will all used cars be sold in this situation?
 - f. If the fraction G of good used cars is sold, then what price will clear the market? Write this as $P_B(G)$. In this context, P_B denotes the price that *buyers* are willing to pay. *Hint:* It is easy to derive an equation for this price given the answer you derived in part d.
 - g. Use the equation from part f and the equation for inverse supply to determine the market-clearing price in this market. What portion of good used cars is sold? What portion of the cars sold are lemons? *Hint:* You will have to factor a quadratic equation or use the quadratic formula to derive this answer.
8. To answer Problem 8, complete Problem 7 first. Suppose that all of the information from Problem 7 holds except that the portion of lemons in the population is one-half rather than one-third. In this instance, prove that no good used cars will be offered for sale in the market equilibrium. *Hints:* This change affects the equation derived in Problem 7d. Thus, it also changes the weighed average price function in Problem 7f. Once these altered equations are obtained, one need only show that no portion of good used cars sold, G , satisfies the equality required by Problem 7g.
 9. An assertion is made in the lemons discussion in this chapter that demand may well be backward bending in a market in which quality is unobservable prior to purchase. Consider the following model: demand, X , is a function of price, P , and average quality, A . In particular, $X(P, A) = 16000 \cdot P^{-2}A^4$.

- a. What type of demand function is this? What is the price elasticity of demand and the quality elasticity of demand?

Average quality depends on price: A higher price in the market will create a higher average quality level in the market. Let $A(P) = 100 \cdot (P/[P + 10,000])$.

- b. What is average quality at $P = \$0, \$5,000, \$10,000, \$15,000,$ and $\$20,000$?
- c. Based on your answer to part b, what can you say about the shape of the average quality function? In particular, fill in the following sentence.

Average quality _____ at a(n) _____ rate as price increases.

- d. What is demand X at $P = \$5,000, \$10,000, \$15,000,$ and $\$20,000$?
- e. Based on your answer to part d, demand must become backward bending at a price somewhere in what range of prices? In particular, fill in the following sentence.

The price must be more than _____ and less than _____ at the point where demand turns from being upward sloping to downward sloping (as price increases).

- f. This part of Problem 9 requires calculus. Determine the price at which demand turns from being upward sloping to downward sloping (as price increases). *Hint:* Apply the total derivative discussed in footnote 9.
- g. This part of Problem 9 requires Excel, but you need not complete part f. Create two graphs: (1) average quality as a function of price in (P, A) space, and (2) demand as a function of price in (X, P) space. Make sure your average quality and demand functions have the shapes and hit the (P, A) and (X, P) values you derived in parts a to e above.
10. Suppose that a firm finds that its low-productivity workers have a marginal revenue product of \$21,000 per year and that its high-productivity workers have a marginal revenue product of \$35,000 per year. The cost of receiving a year's worth of higher education is \$4,000 for a low-productivity worker and \$3,200 for a high-productivity worker. (This cost can be thought of as repayment on loans that have been incurred in the process of obtaining the education.) The firm plans to offer a wage of \$21,000 to workers without higher education and \$35,000 to workers who attain a "certain" level of higher education.
- If this firm uses this pair of wage offers and wishes to create a separating equilibrium using education as a separating device, how many years of education should be required to receive the higher wage?
 - Why does education work as a signal in this instance?
 - Does education increase the marginal productivity of workers in this instance?
11. Suppose you work for an insurance company. You know that there are equal numbers of individuals who will get in an accident with probability 0.2 and 0.3 and that the loss from getting in an accident is \$4,000. Suppose you wish to screen individuals by offering full-coverage insurance as well as 30 percent coinsurance (i.e., the insurance pays 30 percent of the loss amount, or \$1,200, in the event of an accident). You wish to include a 5 percent premium over the actuarially fair values for these policies to help cover overhead and provide a profit margin for your company.
- What price will the insurance company charge for each type of policy?
 - Will this pair of policies screen individuals into risk classes if we assume that all insured individuals have utility function $U(X) = 2000 \cdot (X/4000)^{0.5}$ where X is income? *Hint:* Calculate utility (or expected utility) for each of the three options (no insurance, full-coverage insurance, and coinsurance) for both types of individuals and compare expected utility levels to see what each will choose.
- Background on Expected Utility:* In this conceptualization of utility, we can compare uncertain outcomes with certain outcomes. For example, if an individual does not have insurance and does not get into an accident, he or she has a utility level of 2,000 (because $X = \$4,000$) or zero if he or she gets into an accident (because then $X = \$0$). The expected utility (EU) is based on the likelihood of each event occurring. A person facing a 10 percent chance of having an accident would have an expected utility of 1,800 using this utility function because $EU = 0.9 \cdot 2,000 + 0.1 \cdot 0$. This person would be indifferent between facing this "lottery" and having \$3,240 with certainty (because $U(3240) = 1,800$). This *certainty equivalent value* was obtained by solving for X in the equation $1800 = 2000 \cdot (X/4000)^{0.5}$. The lottery would be preferred to amounts less than \$3,240 offered with certainty.

Capital Budgeting and Risk

Learning Objectives

Upon completion of this chapter, readers should be able to:

- Identify the types of capital budgeting decisions.
- Show how to calculate the net present value and the internal rate of return, and understand the difference between the two.
- Identify different types of cash flows, and explain how they fit into the capital budgeting calculation.
- Define the cost of capital, and demonstrate how it is calculated.
- Explain the meaning of the capital budgeting model.
- Define capital rationing.
- Define risk and uncertainty.
- Describe and calculate various measures of risk, such as standard deviation and coefficient of variation.
- Explain the meaning of the risk-adjusted discount rate and certainty equivalents.
- Distinguish between sensitivity analysis and scenario analysis.
- Describe how to calculate simulations and decision trees.
- Explain how real options can improve capital budgeting calculations.

The Situation

George Kline, the manager of Global Foods' capital planning department, is responsible for analyzing capital budgeting projects. When the analysis is completed, George and his staff of five make presentations up the company hierarchy—to the treasurer (George's boss) and the vice president of finance. If the proposal is large enough, it may finally have to be approved by the Corporate Management Committee (a group composed of top executives).

George has in front of him a new project proposal that requires extensive analysis by his staff. It is the proposed expansion of company activities into a new geographic region. Global is investigating the possibility of entering a new region where its soft drink and bottled water products have not been marketed previously. Potential annual sales for this area have been estimated at 100,000,000 cases. A 4 percent annual increase in total consumption is forecast. The market research people have estimated that, given an extensive advertising campaign, the first year's market share could reach 1 percent, and it might grow to some 5 percent 4 years later.

For the company to compete in this new area, it must establish a plant. An unused, somewhat obsolete bottling plant is available in the area for \$5 million. The total costs of rebuilding and renovating the plant, and purchasing and

(continued)

(continued)

installing the new equipment, are expected to amount to \$2 million. During the first year, the company will incur expenses of recruiting a new work force and an extensive advertising and promotion campaign. These expenses are estimated at \$750,000.

If Global Foods actually achieves a 1 percent share of the market, it will sell 1 million cases. Each case sells for \$5. Production costs will be \$2.50 per case. General and administrative expenses will be \$650,000 during the first year. Of this amount, \$500,000 is fixed; the remainder is a function of sales revenue. Distribution and selling expenses will be 60 cents per case. Advertising expenses will be 5 percent of sales.

The total cost of the plant, renovation, and new equipment and their depreciation schedule is as follows:

- Land—\$500,000, not depreciable
- Plant—\$3,500,000, depreciated straight-line over 31½ years
- Machinery—\$3,000,000, depreciated by the Modified Accelerated Cost Recovery System (MACRS) over 7 years

Global also expects that it will have to increase the size of its working capital by \$750,000 to cover the additional inventory, accounts receivable, and cash for transactions.

The analysis will span 7 years, including the first year of expenditures and 6 years of operations. This is a relatively conservative assumption—if the company cannot make a viable business of this plant over a 7-year time period, it would consider the operation too risky to undertake. Two other important pieces of information needed to complete the analysis are the following:

The company's marginal income tax rate (federal, state, and local) is 40 percent.

The cost of capital for this project will be 15 percent.

After George completes the evaluation, he will perform a risk analysis to consider potential upside and downside risk. George knows that there are various ways to account for risk. He decides that, although he will make an extensive assessment, he will keep it relatively simple so management will be comfortable in making the final decision.

INTRODUCTION

When we discussed the long-run cost curve, we accepted that each point on the horizontal axis represented a different plant size. We did not ask how the company got from one point to the other. The subject of capital budgeting addresses the question of how a company decides to make investments in additional capacity or in new products, and to replace worn-out fixed assets. We now analyze how a company makes decisions to commit funds for future periods, leading to the maximization of shareholders' wealth.¹

Although capital budgeting is a subject taught at great length in corporate finance courses, we provide a brief treatment in this text. We see that capital budgeting applies the tools of incremental and marginal analysis that we encountered in previous chapters. Capital budgeting is microeconomic analysis extended to multiperiod problems.

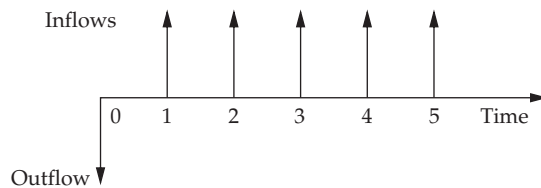
In the latter part of this chapter, we turn to a discussion of risk. Throughout this text, we implicitly treated all quantities as if they were certain. However, risk is always present in any decisions managers make because future events are not known with any degree of certainty. We explore the various methods of incorporating risk into economic analysis.

¹See the discussion on the maximization of shareholder wealth in Chapter 2.

THE CAPITAL BUDGETING DECISION

Capital budgeting describes decisions where expenditures and receipts for a particular undertaking will continue over a period of time. These decisions usually involve outflows of funds (expenditures) in the early periods, and the inflows (revenues) start somewhat later and continue for a significant number of periods.²

The following figure is a simple illustration of the components to be considered in making capital budgeting decisions. The figure shows one outflow at the beginning of the project and five inflows in subsequent periods. This model could represent the purchase of a new machine that will last 5 years and provide new revenues (or savings, if it decreases the cost of production) during that time. We could have shown an example with outflows occurring during, say, the first 3 years. Such flows could represent a company's decision to build a new plant and equip it with new machines, with inflows beginning in year 4.



Types of Capital Budgeting Decisions

Now that we have described the general characteristics of a capital budgeting decision, we can list types of projects that fit this category:

Expansion of facilities. Growing demand for a company's products leads to consideration of a new or additional plant, sales offices, or warehouses.

New or improved products. Additional investment may be necessary to bring a new or changed product to the market.

Replacement. Replacement decisions can be of at least two types: (1) replacement of worn-out plant and equipment or (2) replacement with more efficient machines of equipment that is still operating but is obsolete.

Lease or buy. A company may need to decide whether to make a sizable investment in buying a piece of equipment or to pay rental for a considerable time period.

Make or buy. A company may be faced with deciding whether to make a significant investment to produce components or to forgo such investment and contract for the components with a vendor. Outsourcing, which has become extremely important during the past decade or so, is discussed in Chapter 2 and elsewhere.

Other. A capital budgeting problem exists whenever initial cash outflows and subsequent cash inflows are involved. For instance, an advertising campaign or an employee training program would lend itself to the same method of analysis.

Safety or environmental protection equipment. Such investments may be mandated by law and therefore are not necessarily governed by economic decision making. However, if there are alternate solutions, capital budgeting analysis may be helpful in identifying the most cost-efficient alternative.

²Note that the term *flows*, meaning inflows or outflows, is used in this chapter. In capital budgeting decisions, the quantities considered are the actual expenditures and receipts of cash, not quantities adjusted by accounting conventions, such as accruals. Cash flows are quite objective—either cash goes out or it comes in—and there are no allowances made for when revenues and expenses are recognized in the company's books of account.

TIME VALUE OF MONEY

Because capital budgeting involves cash flows occurring at various times in the future, we must make them equivalent at a particular point in time. This involves the use of the **time value of money**. All this term really means is that a dollar today is worth more than a dollar tomorrow. As long as there is an opportunity to earn a positive return on funds, a dollar today and a dollar a year from now are not equivalent. (In the consumer choice context, there is a time value of money even in the absence of being able to earn a positive return. A dollar today provides extra options to the individual. In particular, if an individual is provided with a dollar today, she or he can either spend it today or save it to spend tomorrow. This option does not exist for the dollar tomorrow.)

Thus to put cash flows originating at different times on an equal basis, we must apply an interest rate to each flow so they are expressed in terms of the same point in time. In capital budgeting calculations, cash flows are usually brought back from various points in the future to the beginning of the project—time zero. It is then said that all cash flows are discounted to the present to obtain a present value. This is a useful convention, although we could discount or compound the flows to any date.



Those of you who are unfamiliar with time value of money calculations or whose memory on this subject is rather hazy will benefit from reading the time value of money material on the Companion Website.

METHODS OF CAPITAL PROJECT EVALUATION

Various methods are used to make capital budgeting decisions, that is, to evaluate the worth of investment projects. Two methods that have been used for many years are mentioned only briefly. They are the payback method and the accounting rate of return method. Although they are still often used in the business world, they have generally been judged to be inadequate.

The **payback** method calculates the time period (years) necessary to recover the original investment. The **accounting rate of return** is the percentage resulting from dividing average annual profits by average investment.³ Among the many drawbacks of these methods is the fact that neither applies the criterion of the time value of money in its computations.⁴ Readers who are not familiar with these two measures and are interested in learning about them will find more lengthy descriptions in any basic corporate finance textbook.

The two major methods that do discount cash flows to a present value are **net present value (NPV)** and **internal rate of return (IRR)**.⁵ Both techniques satisfy the two major criteria required for the correct evaluation of capital projects: use of cash flows and use of the time value of money.

Net Present Value

The net present value of a project is calculated by discounting all flows to the present and subtracting the present value of all outflows from the present value of all inflows. In simple mathematical terms,

$$\text{NPV} = \sum_{t=1}^n \frac{R_t}{(1+k)^t} - \sum_{t=0}^n \frac{O_t}{(1+k)^t}$$

³Note that the accounting rate of return uses profits, rather than cash flows, in its calculation.

⁴Another version of payback, the discounted payback method, uses time value of money calculations. But it does not consider cash flows received after the payback period has been reached.

⁵A third calculation method using discounted cash flows is called the *profitability index* or *index of present value*. This is a derivative of the other two methods and is discussed briefly in this text. For a more in-depth discussion of this method, you should refer to any of the leading textbooks in managerial finance.

where t = Time period (e.g., year)
 n = Last period of project
 R_t = Cash inflow in period t
 O_t = Cash outflow in period t
 k = Discount rate (cost of capital)

Some of these terms must be explained further. Inflows are shown from period 1 to period n ; however, inflows may not occur in all periods. If the project under consideration is the construction of a plant, the time elapsed before the first shipment of product, and thus the first inflow, may not occur until a later period. Remember that in George Kline's expansion project, inflows will not begin until the second period.

Outflows are shown starting in period 0 (i.e., at the very beginning of the project). Indeed, the only outflow may occur in period 0 if the proposal being evaluated is the purchase of a machine that will begin to produce cash inflows upon installation.

The **discount rate**, k , is the interest rate used to evaluate the project. This rate represents the cost of the funds employed (the opportunity cost of capital) and is often called the **cost of capital**. It can also be referred to as the *hurdle rate*, the *cut-off rate*, or the *minimum required rate of return*.⁶

To illustrate the NPV method, we use a simple example. A proposed project requires one initial investment of \$100. The project will last 3 years, and the cash inflows will be \$40, \$60, and \$30 in years 1 to 3, respectively. The cost of capital is 12 percent.

	Inflow X_t	PV factor Table A1.c	PV formula $X_t / (1 + k)^t$	Present Value
Year 1	\$40	× 0.8929	40/1.12	\$ 35.72
2	60	× 0.7972	60/1.12 ²	47.83
3	30	× 0.7118	30/1.12 ³	21.35
				<u>\$104.90</u>
Outflow, $t = 0$				-100.00
Net present value				<u>\$ 4.90</u>

All the estimated cash inflows have been brought back (discounted) to the present at the 12 percent cost of capital and then added. (Two methods are shown in the above table. One uses the factors from Table A1.c in Appendix A at the end of the text. The other uses the present value formula. Both lead to the same outcome, although there could be a small rounding difference.) The total present value of cash inflows is deducted. In this case, there is only one outflow occurring at period 0 (now); thus no discounting is necessary. The net present value equals \$4.90. Should this proposal be accepted?

The answer is yes. The net present value for this program is positive. Stated somewhat differently, if we add all the cash inflows discounted at the cost of capital and deduct the cash outflow, we still have something left over. We expect to earn more than the cost of capital (i.e., the cost of financing this project). The proposal earns what the suppliers of capital require plus an additional amount.

⁶The term *cost of capital* often applies to the overall average cost of funds for a corporation. This cost may differ from the rate used for a specific division of the company or for a particular capital budgeting proposal. One of the major reasons for the difference is risk (both business and financial risk). The corporation's cost of capital represents an average for the whole entity, but specific areas of the business may be more or less risky than the average and thus require higher or lower discount rates. Thus, the discount rate may differ from area to area or project to project within the same company.

If the NPV is positive, the project is financially acceptable. If NPV is negative, rejection is indicated. If NPV is exactly zero, this proposal is just earning the cost of capital; we are on the borderline. However, because the return just equals the required rate of return, the project appears to be acceptable.⁷

Internal Rate of Return

Rather than looking for an absolute amount of present-value dollars, as in the NPV analysis, we solve for the interest rate that equates the present value of inflows and outflows:

$$\sum_{t=1}^n \frac{R_t}{(1+r)^t} = \sum_{t=0}^n \frac{O_t}{(1+r)^t}$$

or

$$\sum_{t=1}^n \frac{R_t}{(1+r)^t} - \sum_{t=0}^n \frac{O_t}{(1+r)^t} = 0$$

The r term in the equations is the internal rate of return—the unknown variable for which we solve. Actually, the IRR solution is only a special case of the NPV technique; the internal rate of return of a project is the discount rate that causes NPV to equal zero (which occurs when the project is just earning its cost of capital).

The calculation of the IRR can be easily accomplished with a handheld business calculator or a computer (using, for instance, an Excel function), and this is what most people do. However, for our purposes, we use the tables in Appendix A at the back of the text. (We could have also shown the formula calculation, but we have omitted it here to keep the table simple.) In this case, the calculation may become rather cumbersome. Unless all cash inflows are uniform and there is only one outflow (in which case we can simply employ the annuity formula), it is necessary to find the answer by trial and error. One must first choose an applicable interest rate—and this may be no more than an educated guess—and then iterate until the correct answer is obtained.

Turn to the example used previously. We found in the NPV analysis that, with a 12 percent cost of capital, net present value is positive (\$4.90), so the internal rate of return must be greater than 12 percent. We first try 14 percent.

The result is still positive, so we try the same calculation with a 15 percent interest rate.

	At 14 percent		At 15 percent	
Inflow, year 1	40×0.8772	\$ 35.09	40×0.8696	\$ 34.78
Inflow, year 2	60×0.7695	46.17	60×0.7561	45.37
Inflow, year 3	30×0.6750	20.25	30×0.6575	19.73
		\$101.51		\$ 99.88
Outflow, year 0		100.00		100.00
Net present value		\$ 1.51		-\$.12



NPV is now negative; 15 percent is too high. Thus the result lies somewhere between 14 percent and 15 percent. It is readily seen that the IRR is much closer to 15 percent than to 14 percent. A more precise answer can be obtained using linear interpolation; this would result in an internal rate of return of about 14.9 percent.

⁷This case is similar to one encountered in Chapter 8. Production takes place at the point where marginal cost equals marginal revenue. If discrete quantities are involved, the cost of the “last” unit produced is the same as the revenue received from it; it earns no economic profit, but only normal profit. In capital budgeting, the situation is really the same: the “last” proposal that would be accepted is the one that just earns the rate required by the suppliers of capital.

The accept/reject criterion for the internal rate of return is based on a comparison of the IRR with the cost of capital of the project. If the internal rate of return is larger than the cost of capital—the required rate of return—of the proposal, it signals acceptance. If $IRR < k$, the proposed project should be rejected. If $IRR = k$, although it could be said that decision makers would be indifferent on whether to undertake the project, an argument can be made that the project is earning its cost of capital and therefore should be accepted at the margin.

The Profitability Index

There is a third method for evaluating projects, which supplements the previous two methods. It is called the **profitability index (PI)**. Its formula is

$$PI = \text{Present value of cash inflows} / \text{Initial investment}$$

The project will be financially acceptable if PI is greater than 1 and not acceptable if PI is less than 1.

NPV Versus IRR



Module 12A

Should one of the preceding methods be preferred to the other? In a large majority of cases, both NPV and IRR will indicate the correct accept/reject decision because when

$$\begin{aligned} NPV > 0, IRR > k, PI > 1 \\ NPV = 0, IRR = k, PI = 1 \\ NPV < 0, IRR < k, PI < 1 \end{aligned}$$

Thus, either of the three measures gives the correct answer. In some cases, however, problems may arise.

When **independent projects** are being analyzed, both IRR and NPV criteria give consistent results. “Independent” implies that if a company is considering several projects at the same time, they can all be implemented simultaneously as long as they pass the NPV or IRR tests and as long as funds are not limited. The adoption of one independent project will have no effect on the cash flows of another.

However, proposals may be **mutually exclusive projects**. This occurs when two solutions for a particular proposal are offered, only one of which can be accepted. Suppose Global decides to acquire a new depalletizer. At this point, sales representatives from two manufacturers of these machines descend on the company, and each offers a new version. But the company needs only one depalletizer. NPV analysis may suggest purchase of machine A, but IRR indicates machine B. Such disparate signals can occur if one or both of the following conditions are present:

1. The initial costs of the two proposals differ.
2. The shapes of the subsequent cash inflow streams differ; for instance, one alternative may have large early inflows with the other exhibiting increasing inflows over time.

The reason for the differences between IRR and NPV results is the implicit reinvestment assumption. In the NPV calculation, as inflows occur, they are automatically assumed to be reinvested at the cost of capital (the project’s k). The IRR solution assumes reinvestment at the internal rate of return (the project’s r).

Conflicting accept/reject signals may not occur frequently in capital project analysis, but they do occur, and they may cause the analyst some anxious moments.

An Example

As mentioned previously, conflicting results can be caused by a difference in project size. Table 12.1 shows such a case. Project A involves an original outlay of \$1,500; its

Table 12.1 Two Mutually Exclusive Projects That Differ in Size

Project	t = 0	t = 1	t = 2	t = 3	t = 4	t = 5
A	\$(1,500)	\$580	\$580	\$580	\$580	\$0
B	\$(1,000)	\$400	\$400	\$400	\$400	\$0
Cost of capital			15.0%			
	<Internal Rate of Return>			<Net Present Value>		
Project A		20.1 %			\$156	
Project B		21.9 %			\$142	
Delta Project						
(A - B)	(500)	180	180	180	180	0
Internal rate or return	16.4%		Net Present Value			14
<i>Project Evaluation at Several Discount Rates</i>						
	0.00%	5.00%	10.00%	15.00%	20.00%	25.00%
Project A	\$820	\$557	\$339	\$156	\$1	\$(130)
Project B	600	418	268	142	35	(55)

cash inflows are \$580 per year for 4 years. Project B (which is a substitute for A) is somewhat less expensive—only \$1,000—but the four cash inflows are also smaller, at \$400 each. Each project has a 4-year life and no salvage value. The cost of capital is 15 percent. The IRR of project B exceeds the IRR of project A. However, the NPV of project A is greater than that of project B.

To resolve this dilemma, we calculate the NPV and IRR for an “incremental” (or *delta*) project. That is, we take the differences between the two project cash flows and create a delta project. If we undertake A rather than B, we must incur an additional original cash outflow (at $t = 0$) of \$500 and receive in turn additional cash inflows of \$180 for each of the next 4 years. Such a project, evaluated separately, would have an NPV of \$14 and an IRR of 16.4 percent. This means that the additional outlay of \$500 provides an incremental positive NPV (\$14) and an internal rate of return that exceeds the 15 percent cost of capital. Thus both criteria indicate that the additional investment of \$500 is worthwhile. It follows that the NPV rule, which suggested project A, was the correct indicator, and that project A should be chosen over project B.

Figure 12.1 illustrates the relationship between the two projects. At the bottom of Table 12.1, the two projects have been evaluated at six different discount rates, and their respective NPVs have been plotted on the graph. When the discount rate is 0, the NPV is merely the sum of all the cash flows. This is shown on the vertical axis. On the horizontal axis, the projects’ NPVs become 0 at their respective IRRs. The two projects’ NPVs cross at \$111, at a discount rate of 16.4 percent, which is the IRR of the incremental project. To the left of this intersection, project A’s NPV exceeds that of project B’s, which creates a conflict because project B’s IRR (21.9 percent) is always higher than project A’s (20.1 percent). Had the cost of capital been higher than 16.4 percent, project B would have been preferred under both methods because it also would now have the higher NPV.

The second instance when the two measures may give conflicting rankings occurs when the shapes of the cash inflow streams differ significantly. Although the initial outlays for both mutually exclusive projects may be the same, one of the projects has a cash flow pattern that starts slowly and builds up, whereas the other has cash inflows that are initially large but decline over time. Readers can verify that this is so by entering data into Module 12A and calculating results.

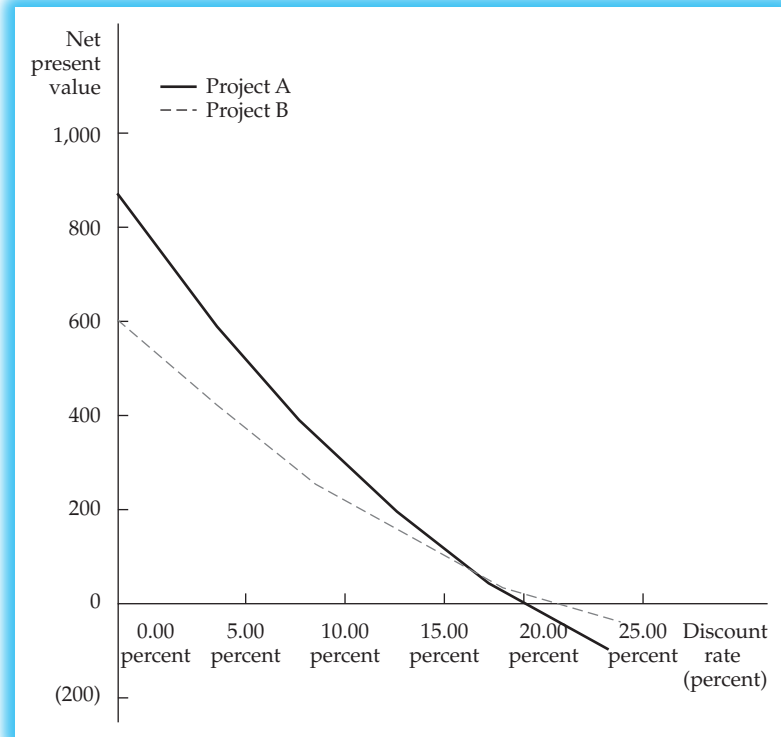


Figure 12.1 Net Present Value Profiles for Projects of Different Size

Another problem that may occur concerns the case of nonconventional cash flows. Conventional cash flows occur when cash outflows are followed by a series of cash inflows for the remainder of the project's life. In other words, over time, there is only one change from negative flows (outflows) to positive (inflows). But suppose there are two or more changes. If a project starts with a cash outflow followed by a series of cash inflows and then ends up with a cash outflow (i.e., two changes in sign), two different rates of return will result. Such an answer is obviously not satisfactory. If NPV analysis is used, a single answer will be obtained.

Writers of financial and economic literature almost unanimously recommend NPV as the theoretically more correct measure. Arguments to support this choice are

1. The financial objective of the firm is the maximization of stockholder wealth. Projects with the largest NPVs add up to the highest present value for the business.
2. The NPV reinvestment assumption, at k , appears to be more realistic in most cases than reinvestment at r of a particular project.⁸

⁸There is a method that has gained favor and that corrects some of the problems encountered with the internal rate of return. It is called the *modified internal rate of return* (MIRR) and is calculated by discounting, at the cost of capital, all cash outflows to year zero and compounding all inflows to the end of the project. The discount rate that equates the sum of the ending values to the sum of the beginning values is the MIRR. Using this method reinvests the cash flows at the cost of capital. Also, the possibility of obtaining more than one solution (nonconventional cash flows) is eliminated. However, accept/reject signal conflicts with NPV can still occur. For a longer explanation, see Stephen A. Ross, Randolph W. Westerfield, and Jeffrey Jaffe, *Corporate Finance*, 9th ed., New York: McGraw-Hill Irwin, 2010, pp. 147–149.

Capital Budgeting in Practice

According to numerous surveys, payback and accounting rate of return have been the popular methods in the past. Over the past 40 years numerous surveys of capital budgeting practices have been conducted. During this period there has been a significant increase in use of the more sophisticated methods—NPV and IRR. This is particularly true among large companies. IRR has been the favored method, despite what financial experts have recommended. Why? Businesspeople are probably more comfortable with a percentage figure than with a dollar figure that is not immediately recognizable.

A 2001 study of 392 companies found that about 75 percent used the IRR and NPV methods “always or almost always.” Large firms (with sales of more than \$1 billion) emphasized discounted cash flow analysis even more. Payback was more prominent among small firms.⁹

Another study covered 232 small manufacturing companies (defined as those with sales of less than \$5 million per year and with fewer than 1,000 employees). Here the payback and accounting rate of return methods were most frequently used, while IRR and NPV lagged far behind.¹⁰

A more recent study investigated capital budgeting practices of very small firms, those with 249 employees or fewer. There were 792 companies in the sample. Here, as should be expected, the use of the older methods was predominant. Actually, the most frequent answer—of 26 percent of the firms in the sample—was “gut feel.” This was followed by the payback period with 19 percent and the accounting rate of return with 14 percent. Discounted cash flow analysis was used by only 12 percent of the companies surveyed.¹¹

In summary, as one looks at the numerous surveys that have been conducted over the last 30 or so years, there is an unmistakable significant upward trend in the use of the more sophisticated capital budgeting analysis methods.

CASH FLOWS



Module 12B Module 12C

Up to this point, the discussion has concentrated on the methods and mechanics of the capital budgeting process. Cash flows were assumed and put into the appropriate formulas for processing. The term *cash flow* has been used in abundance but has not been explained thoroughly.

When confronted with a capital budgeting proposal, the analyst’s most difficult task is to enter the best estimates of cash flows into the analysis. Because all the inflows and outflows are in the future, their amounts and timing are uncertain. Some of them, such as the cost of the new machine, can be assessed with relative certainty. But as the analyst tries to assess future annual benefits and costs, the amount of uncertainty increases.

In most cases, capital budgeting analysts do not generate the inputs for the model. They obtain the estimates from other parts of the organization, such as market research, marketing, manufacturing, engineering, or service. These data have to be examined for potential bias. Market forecasts may be too high because the people who prepare them are interested parties. Costs often are underestimated. In general,

⁹John R. Graham and Campbell R. Harvey, “The Theory and Practice of Corporate Finance: Evidence from the Field,” *Journal of Financial Economics*, 61 (2001), pp. 1–53.

¹⁰Stanley Block, “Capital Budgeting Techniques Used by Small Business Firms in the 1990s,” *The Engineering Economist*, Summer 1997, pp. 289–302.

¹¹Morris G. Danielson and Jonathan A. Scott, “The Capital Budgeting Decision of Small Businesses,” *Journal of Applied Finance*, Fall/Winter 2006, Vol. 16, Issue 2, pp. 45–56.

estimated cash flows tend to be optimistic and must be adjusted to make them more realistic.¹²

The person organizing the data must understand the following points:

1. All revenues and costs must be stated in terms of cash flows.
2. All cash flows should be incremental. Only those flows that will change if the proposal is accepted should be recorded.
3. Sunk costs do not count. Costs incurred prior to the time the project is being evaluated and that cannot be recouped should be omitted from consideration.
4. Any effect on other parts of the operation must be taken into account. If the introduction of a new diet soda will have an adverse impact on sales of current soft drinks, this amount must be subtracted from the cash flows planned for the new product. In contrast, impact may be positive. For example, if the company also sells alcoholic beverages, the introduction of a “light” tonic could enhance sales of the company’s gin. These are, of course, examples of the familiar cases of substitutes and complements.
5. Generally, in capital budgeting analysis, interest paid on debt is not considered. Because interest is included in the discount rate, showing it as a cash outflow would amount to double counting.

Types of Cash Flows

Cash flows come in many varieties. Some of the most common and important types are discussed next.

Initial Cash Outflows

Cash flows occur at the inception of the project. If a new machine is installed, this represents a one-time outflow. But initial outflows can also be spread over a period of time, as mentioned previously.

Operating Cash Flows

When a new project goes online, it begins to generate cash inflows (revenues). Of course, it also generates cash outflows (costs and expenses), which must be subtracted from the inflows. In the early years of operations, outflows can exceed inflows; thus the annual net outflow can continue even after the initial investment stage.

One expense that is recorded by accountants on income statements is depreciation. However, depreciation does not represent a cash flow. The cash flow occurred when the fixed asset was acquired, and depreciation is an accounting entry that shows the decrease in its value. Although no cash flow occurs when depreciation is recorded, it decreases the profits reported to tax authorities and thus qualifies for a tax deduction. Therefore, in capital budgeting analysis we must add depreciation back to the aftertax profit to arrive at cash flow. Here is a simple example:

Sales	\$100
Costs and expenses	–50
Depreciation	–20
Total costs and expenses	–\$ 70
Net profit before tax	30
Income tax	–12
Net profit after tax	\$ 18
Depreciation	20
Net cash flow	\$ 38

¹²Stephen W. Pruitt and Lawrence J. Gitman, “Capital Budgeting Forecast Biases: Evidence from the *Fortune* 500,” *Financial Management*, Spring 1987, pp. 46–51.

For each period (years), operating cash flows have to be calculated and the results discounted to the present.

Additional Working Capital

In the case of an expansion proposal, in addition to new plant and equipment, increased working capital may be required. Inventories may be larger, accounts receivable may grow, and more cash may be needed to finance transactions. Investing in working capital is a cash outflow that is similar in nature to an investment in brick and iron. We must account for it. When the proposed project comes to the end of its life, inventories will be used up, accounts receivable will be collected, and the additional cash will no longer be needed. So the same amount that was expended at the beginning of the program may be returned at the end—with one big difference: The cash outflow occurs at or near the beginning of the operations, whereas the inflow of the same amount occurs at the end and must be discounted to the present.

Salvage or Resale Values

At the end of the project's life, a machine that has been completely depreciated (i.e., has an accounting book value of zero) may turn out to have a residual resale value or some scrap value. If cash can be obtained for it, a cash inflow will result.

But care must be exercised in including this cash flow. If the market value is greater than the book value, a profit will result, with inevitable tax consequences. The formula for computing the cash flow in such a case is

$$\text{Cash flow} = SV - (SV - BV) T$$

where SV = Salvage or resale value

BV = Book value

T = Tax rate

Noncash Investment

Sometimes a new project involves an investment that does not require a cash flow. For instance, suppose an old, fully depreciated machine is standing on the factory floor. This machine is not needed for present production requirements. But then a new expansion proposal is accepted, allowing this old machine to be used. Does it represent a cash outflow? Yes, if the machine has a salvage value; no, if it has no market value. Thus, as in all cases of capital budgeting decisions, the alternatives have to be considered.

COST OF CAPITAL

In each capital budgeting decision described, a certain cost of capital was assumed. Its derivation has not been explained. Much space is usually devoted to this subject in finance textbooks—an entire chapter or more. Such an exposition is beyond the scope of this text. However, a brief explanation of this important concept is essential.

To invest in capital projects, a company must obtain financing. Financing, of course, comes from different sources. There is debt, either short term or long term. Then there is equity. A company may retain earnings, which then become part of its equity, or it can issue new shares. Each type of financing must be paid for; each has its cost. It is these costs that establish a company's cost of capital. When all the costs have been identified, they are combined to arrive at an average cost of capital for a given debt/equity mix.¹³

¹³Admittedly, there are other financial instruments with which a company obtains funds, such as preferred stocks or convertible bonds. However, this short description limits itself to debt and common equity.

Debt

The cost of debt is easier to explain. It is simply the interest rate that must be paid on the debt. But because interest expense is tax deductible, the actual cost of the debt to the company is the aftertax cost. The expression for the cost of debt is

$$\text{Interest rate} \times (1 - \text{Tax rate})$$

Which interest rate should be used? If a company already has debt outstanding, it pays a certain rate. But the rate being paid on past debt is not relevant. What is important to the company in measuring its cost of capital is the interest it would have to pay if it were to borrow today. Thus, the present rate being charged in the market for the kind of debt the company would issue (e.g., life to maturity, risk category) determines the company's cost of debt.

Equity



Module 12D

The cost of equity is more difficult to obtain. A large body of literature exists on this subject, and there are different methods to arrive at this cost. Two of these are described here.

The Dividend Growth Model

Because we discussed the derivation of the **dividend growth model** in Chapter 2, we will just repeat the formulas here. If the annual dividend remains the same each year, then the formula for the price of a share is

$$P = D/k_e$$

where P = Present price of a share of stock

D = Annual dividend

k_e = Discount rate (the cost of equity capital)

If, however, we expect the dividend to grow at a constant rate each year, the equation becomes

$$P = D_1/(k_e - g)$$

where D_1 = Dividend to be paid in the coming year

g = Dividend's annual growth rate

Because it is the cost of equity capital that we seek, the equation can be written in terms of k_e as

$$k_e = \frac{D_1}{P} + g$$

Thus, this construction states that the cost of equity capital (k_e) equals the dividend in year one divided by today's stock price—the dividend yield—plus the expected growth rate in the dividend. This formula is often referred to as the Gordon model, so named for Myron J. Gordon, an economist who has done a great amount of work in this area and who is credited with a major role in developing this model.¹⁴

The dividend growth model just presented is generally applicable when a company reinvests the earnings that have not been paid out as dividends. If a company

¹⁴See, for instance, Myron J. Gordon, *The Investment, Financing, and Valuation of the Corporation*, Homewood, IL: Irwin, 1962.

issues new stock in the financial markets, it incurs an additional cost. The proceeds from the sale of the stock will be less than the current market price, P , due to the cost of underwriting the issue. If these costs, often referred to as *flotation costs*, f , are expressed as a percentage of P , the Gordon model converts to

$$k_e = \frac{D_1}{P(1-f)} + g$$

Obviously, the cost of external equity capital will be higher for newly issued stock than for retained earnings.¹⁵

Although the Gordon model formula appears rather simple, it requires a forecast of growth. Forecasts, as we have found out, are always tenuous. Thus, the calculation of the cost of capital can be only as good as the estimates entered into it.

The Capital Asset Pricing Model

The **capital asset pricing model (CAPM)** had its birth in the 1960s.¹⁶ It is based on the principle that there is a relationship between risk and return. The more risky the investment, the higher will be the required return. Only a brief description of this model is given here.

An important conclusion of this model is that the required rate of return¹⁷ on a stock is a function of the volatility (market risk) of its returns relative to the return on a total stock market portfolio. This volatility is referred to as *beta* (β) and is calculated by regression analysis. The variability of the individual stock's return is the dependent variable, and the variability of the market return is the independent variable. The higher the volatility of the individual stock's return compared with the market return, the higher the beta. A beta of 1.0 signifies that the stock's return is as volatile as the market's. If the beta is greater than 1.0, the stock's return is more variable and, therefore, the stock is more risky; the reverse situation holds when beta is less than 1.0.

One other item is included in this model, the riskless interest rate, usually represented by interest paid on U.S. Treasury securities. The beta coefficient is used to

¹⁵The assumption of constant growth may not always be realistic. For a variable growth rate, the growth rate must be estimated for each period and the entire equation solved for k_e . For example, in the case of a company that expects rapid growth at first, then a slowdown, and finally some "normal" constant growth, the equation becomes somewhat complicated:

$$P_0 = \sum_{t=1}^n \frac{D_t}{(1+k_e)^t} + \frac{D_{n+1}}{k_e - g} \left[\frac{1}{1+k_e} \right]^n$$

where n = Number of years of nonconstant growth

g = Expected annual growth rate during the constant growth period

D_t = Dividend expected in each year t of the nonconstant growth period

D_{n+1} = Dividend in first period of constant growth

For additional information on this construction, see any finance textbook; for instance, Michael C. Ehrhardt and Eugene F. Brigham, *Corporate Finance: A Focused Approach*, 4th ed., South-Western Cengage Learning, 2011, pp. 281–284. This method of calculation is employed later in this chapter in "The Value of a Corporation" section (Appendix 12A).

¹⁶See, for instance, William F. Sharpe, "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *Journal of Finance*, 19 (September 1964), pp. 425–42.

¹⁷The rate of return for a period is defined as

$$\frac{\text{Dividend} \pm \text{change in stock price}}{\text{Stock price at the beginning of the period}}$$

arrive at the risk premium of the individual stock relative to the difference between the average return for the market portfolio and the riskless rate.

The required rate of return on an individual stock is calculated as follows:

$$k_j = R_f + \beta(k_m - R_f)$$

where k_j = Required rate of return on stock j

R_f = Risk-free rate

k_m = Rate of return on the market portfolio

This model has experienced immense popularity—and also criticism. Probably one of the most serious objections is that the model tries to predict the present and future costs of equity capital with past data. It assumes, therefore, the past relationship between stock return and market return will continue into the future.

Another criticism of CAPM lies in the fact that betas are not always stable. They vary, depending on the time period used in making the analysis, and they are affected by the specific statistical method used. Thus, again, the cost-of-capital estimate is only as good as the data used in the computation and the method used. It is obvious that there is much work yet to be done to improve the estimates of the equity cost of capital.

The Weighted Cost of Capital



Module 12E

Despite the fact that the measurements of the components of capital costs are not entirely satisfactory, they are the best available at the present, and they are used in obtaining an overall cost of capital of the firm. This is achieved by weighting the various costs by the relative proportion of each component's value in the total capital structure.

When the **weighted average cost of capital (WACC)** is calculated, the weights should reflect the actual market value of the components rather than book values as shown on the balance sheet.

If debt makes up 20 percent and equity 80 percent of a company's financial structure, and the company's respective costs are 6 percent and 14 percent, the weighted cost of capital is

$$0.2(0.06) + 0.8(0.14) = 0.012 + 0.112 = 0.124 = 12.4 \text{ percent}$$

Because the cost of debt capital is usually lower than that of equity, this formula would indicate that a company can decrease its cost of capital (and thus increase the value of the firm) by increasing the ratio of debt to equity. This is misleading, however. As the proportion of debt increases (i.e., as leverage rises), the financial community will view the company as more risky. Consequently, the cost of both components, debt as well as equity, will rise, causing the weighted average to rise also. There is probably some point where the combination of components is optimal and the weighted cost of capital of a particular firm is at a minimum.

THE CAPITAL BUDGETING MODEL

In the arena of corporate decision making, capital budgeting is an application of the marginal revenue–marginal cost principle. Figure 12.2 illustrates this principle.

Assume a company is faced with a menu of seven independent capital budgeting proposals. The capital planning department has analyzed each of them and has calculated the internal rate of return, which is the evaluation technique the

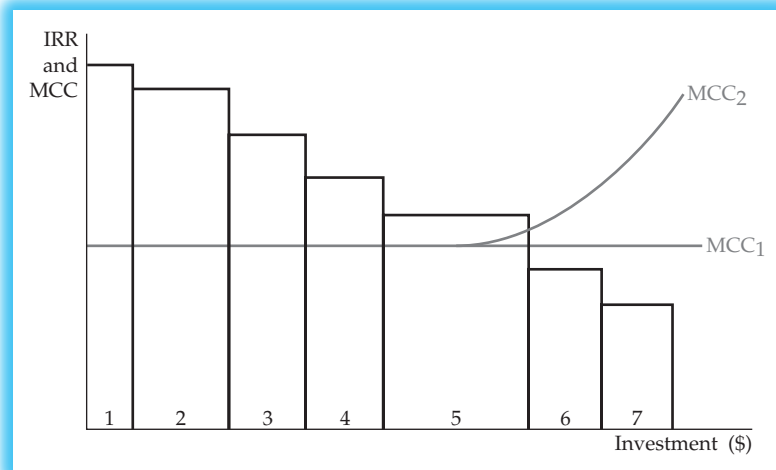


Figure 12.2 The Capital Budgeting Decision

company uses. In Figure 12.2, the projects have been ranked by IRR from highest to lowest. Each proposal is represented by a bar; its height represents the IRR, and the width indicates the size of the investment. If a line were drawn connecting the tops of the bars and then smoothed, it would be a curve representing the internal rate of return on successive doses of investment—a marginal investment opportunity curve.

If the company can obtain new funds at a constant cost of capital, then the horizontal marginal cost of capital (MCC_1) curve would be valid. However, it is more likely that as the company needs additional funds to finance its projects, it may incur higher capital costs. Thus, MCC_2 is probably the more realistic situation.

Thus, as the costs of both debt and equity rise, the marginal weighted cost of capital will show an increase as the corporation increases its demand for capital funds. Hence, it is considerably more realistic to draw the weighted marginal cost of capital as a rising curve (MCC_2), particularly after the capital budget of a certain size is reached.

The company will reach the optimal investment budget at the point where the marginal investment opportunity curve and marginal cost of capital curve intersect. This is, of course, the principle that was applied in the one-period case. In our example, the conclusion is that projects 1 through 5 should be accepted, and those with lower IRRs (projects 6 and 7) should be rejected.

There is one other important point that must be made. We have just concluded that to arrive at an optimal capital budget, a company will accept all projects with positive NPVs (an exception, capital rationing, is discussed in the next section). However, if a firm operates in a competitive market, then in the long run, all NPVs will be reduced to zero. This is consistent with our discussion of competition in Chapter 8, where we conclude that in a long-run equilibrium firms will earn only their normal profit. Of course, in reality, continuous changes in demand and supply conditions may prevent a static long-run equilibrium from ever being reached. Nevertheless, competitive forces will require companies to continue to innovate to maintain their market advantage or to attempt to create noncompetitive (monopolistic) conditions to produce economic profits—and positive NPVs—in the long run.

CAPITAL RATIONING

The marginal rule indicates that the company should invest in every project whose IRR exceeds the marginal cost of capital or whose net present value is positive. However, some corporations impose an absolute limit on capital spending during a particular interval, and not all projects that pass the IRR or NPV test will be accepted. This situation, referred to as **capital rationing**, occurs when management may not be willing to obtain external financing. There are various reasons for such a practice. For instance, there may be a reluctance to incur increasing levels of debt. Alternatively, management may not want to add to equity in fear of diluting control. Thus, under capital rationing conditions, a company may have to reject otherwise acceptable projects to remain within its expenditure limit.

Assume a corporation has evaluated five independent capital projects, with the results shown in Table 12.2. Each project has a positive net present value and therefore should be undertaken under the usual rule. To embark on these five projects, the company would have to incur cash outflows of \$250 and would obtain a net present value of \$120.

However, the management of this firm has decreed that only \$100 will be spent on capital projects at this time. Which projects should it select? To create the most value it can under the circumstances, it will choose the combination of projects that will give the highest net present value within the spending constraint. This leads to the selection of projects A, C, and D. These three projects together will require an outlay of just \$100 and will have a combined net present value of \$60. No other combination of projects within the \$100 spending constraint will achieve a net present value as high as \$60.

Although the imposition of capital rationing does not appear to be rational maximizing behavior, it nevertheless occurs quite frequently. The final judgment on capital rationing is that it does not permit a company to achieve its maximum value.

RISK VERSUS UNCERTAINTY

In economic or financial theory, the two terms **risk** and **uncertainty** have somewhat different meanings, even though they are often used interchangeably. Although no future events are known with certainty, some events can be assigned probabilities, and others cannot. Where future events can be defined and probabilities assigned, we have a case of risk. Thus, for instance, a company's sales manager estimates that next year's sales of diet cola have a 25 percent probability to be 5 million cases, a 50 percent probability to be 6 million, and a 25 percent probability to be 7 million. If there is no way to assign any probabilities to future random events, we are addressing pure uncertainty. Even though this distinction is theoretically important, many writers omit it as a matter of convenience. We follow this practice.

Table 12.2 Investment Choices Under Capital Rationing

Project	Original Investment	Net Present Value
A	\$50	\$25
B	70	30
C	20	25
D	30	10
E	80	30

Probabilities can be classified as *a priori* or *statistical*. The former can be obtained by repetition. Thus, if a true two-sided coin is flipped an infinite number of times, tails will come up on half the tosses and heads on the other half. Instead of throwing the coin an infinite number of times, we can specify the frequency based on general mathematical principles.

In everyday business, a priori probabilities cannot be specified. To assign probabilities to various outcomes, businesspeople must rely on statistical probabilities. These may be obtained empirically, based on past events. For instance, if a particular event has occurred once every ten times in the past, a 10 percent probability would be assigned to it. Alternatively, probabilities obtained from past events can be adjusted to reflect changed expectations for the future.

SOURCES OF BUSINESS RISK

What are the sources of risk faced by a businessperson?

- **Economic conditions:** Firms face rising and falling phases of the business cycle. Although forecasting can help prepare a business for changes, it cannot completely predict timing and volatility of economic activity.
- **Fluctuations in specific industries:** Such fluctuations may not always coincide with the overall economy.
- **Competition and technological change:** When competitors improve a product or new technologies are introduced, sales of particular companies or industries are affected.
- **Changes in consumer preferences:** Successful products for one year may become the discarded, unwanted items of the next. The fashion industry is a good example of changing styles.
- **Costs and expenses:** Labor and material costs are subject to change, sometimes unexpectedly. Oil prices are an excellent example of such an uncertainty.

THE MEASURES OF RISK

When outcomes are uncertain, two measures that take risk into consideration are used.

First, not just one outcome but a number of outcomes is possible. Each potential result will have a **probability** attached to it. In making estimates of a future cash inflow, for example, the analyst must decide on the probability of each possible result and construct a probability distribution.

A **probability distribution** describes the chances of all possible occurrences. When all the probabilities of the possible events are added up, they must total 1 because all

Table 12.3 Probability Distribution for Cash Inflows

Cash Inflow	Probability
\$3,000	0.1
4,000	0.2
5,000	0.4
6,000	0.2
7,000	0.1

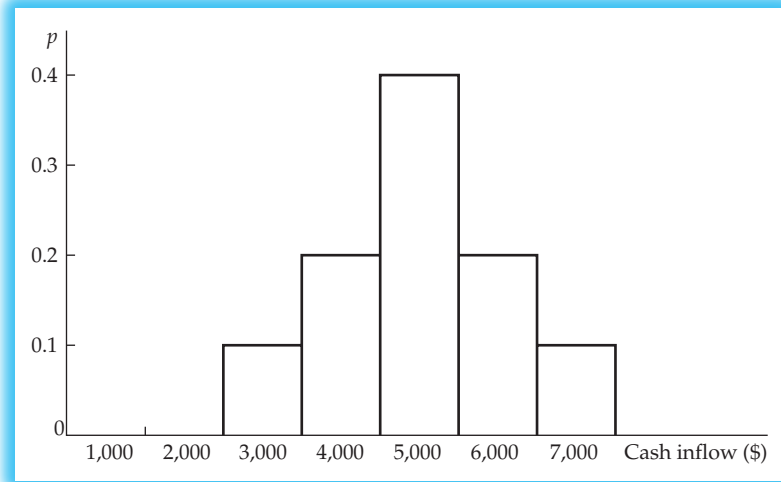


Figure 12.3 Bar Graph of Data from Table 12.3

possibilities together must equal certainty. Thus, we may assign probabilities to various possible cash flows, as shown in Table 12.3. We estimate that possible cash flows from a project during the coming year will be \$3,000, \$4,000, \$5,000, \$6,000, or \$7,000. If all conditions are favorable, the cash flow will be \$7,000; if unfavorable conditions prevail, the cash flow will be only \$3,000. Table 12.3 can also be translated into a bar chart, shown in Figure 12.3.

The assigned probabilities indicate that there is only a 10 percent chance that all the unfavorable predictions will materialize. Similarly, there is only one chance in ten (i.e., 10 percent) that all the favorable conditions will prevail. It is much more likely that some favorable and some unfavorable influences will ensue. Therefore, the probabilities for the intermediate cash flows are the highest.

Once we have established a probability distribution, we are ready to calculate the two measures used in decision making under conditions of risk.

Expected Value



Module 12F

From the numbers given in Table 12.3, we calculate the **expected value** of possible outcomes. The expected value is simply the average of all possible outcomes weighted by their respective probabilities.

$$\begin{aligned}\bar{R} &= (3,000 \times 0.1) + (4,000 \times 0.2) + (5,000 \times 0.4) + (6,000 \times 0.2) + (7,000 \times 0.1) \\ &= 300 + 800 + 2,000 + 1,200 + 700 \\ &= 5,000\end{aligned}$$

where \bar{R} represents the expected value. The generalized expression for expected value is as follows:

$$\bar{R} = \sum_{i=1}^n R_i p_i$$

where \bar{R} = Expected value

R_i = Value in case i

p_i = Probability in case i

n = Number of possible outcomes

Table 12.4 Calculation of Standard Deviation for Table 12.3

R_i	p_i	$(R_i - \bar{R})$	$(R_i - \bar{R})^2$	$(R_i - \bar{R})^2 p_i$
\$3,000	0.1	-2,000	4,000,000	400,000
4,000	0.2	-1,000	1,000,000	200,000
5,000	0.4	0	0	0
6,000	0.2	1,000	1,000,000	200,000
7,000	0.1	2,000	4,000,000	400,000
				1,200,000
$\sigma = \sqrt{1,200,000} = 1,095$				

Now that the weighted average—the expected value—has been obtained, we can determine the second measure, the one that specifies the extent of the risk.

The Standard Deviation

In economics and finance, risk is considered to be the dispersion of possible outcomes around the expected value. The greater the potential differences from the average, the greater the risk. Thus, to measure risk, we must find some yardstick that reflects the variation of possible outcomes from this average. A concept prominent in elementary statistics is used for this purpose—the standard deviation.¹⁸

The **standard deviation** is the square root of the weighted average of the squared deviations of all possible outcomes from the expected value:

$$\sigma = \sqrt{\sum_{i=1}^n (R_i - \bar{R})^2 p_i}$$

where σ is the standard deviation. For the example in Table 12.3, the standard deviation is calculated in Table 12.4.

What is the meaning of a standard deviation of 1,095? First, because our probability distribution is symmetric, there is a 50 percent chance that the outcome will be larger than the expected value and a 50 percent chance that it will be less. Standard deviation is most readily understood by reference to the normal distribution (discussed more fully in the next section). Roughly 2/3 of all outcomes are within plus or minus one standard deviation of the mean and 95 percent of all outcomes are within two standard deviations of the mean.

The combination of the expected value and the standard deviation aids in making a decision between two projects. Suppose we have to choose between project 1 (shown in Tables 12.3 and 12.4) and another, called project 2.

Project 2	
Cash Inflow	Probability
\$2,000	0.10
3,500	0.25
5,000	0.30
6,500	0.25
8,000	0.10

¹⁸If you want to review the concept of standard deviation more thoroughly, consult any college statistics textbook.

When the expected value and standard deviation are calculated, the results are 5,000 and 1,710, respectively.

Because the expected values of the two proposals are identical, the decision will be made on the basis of the standard deviation. Project 2, with the greater standard deviation (1,710), is the riskier of the two. Generally, businesspeople are averse to risk; therefore, project 1, which has the lower risk, would usually be accepted.

Discrete Versus Continuous Distributions and the Normal Curve

The examples of frequency distributions discussed so far are known as discrete.¹⁹ The potential outcomes have been limited to just five numbers for each of the two projects. However, it is quite likely that other outcomes could occur—for instance, 4,679 or 6,227. If all possible outcomes are considered, we have a continuous distribution. The possible outcomes on a continuous distribution are often described by a bell-shaped curve, referred to as a *normal curve*. The two normal curves shown in Figure 12.4 have the same expected value and standard deviation as projects 1 and 2. These distributions provide continuous counterparts to the discrete distributions described above. Both are symmetrical about the center value of 5,000. Both curves approach but do not reach zero at either end of the *X*-axis. Curves of this kind are called *asymptotic*.

With projects 1 and 2 drawn as continuous standard normal probability functions, project 1 exhibits a tighter curve, whereas project 2's curve is spread out over a much larger horizontal distance.

The probabilities for any other range of numbers can be easily obtained from a table of values of the areas under the *standard normal distribution* function (see Table A.2 in Appendix A at the end of this text). The standard normal distribution has a mean of zero and a standard deviation of one. Any distribution can be transformed into a distribution with mean zero and standard deviation of one by subtracting the mean and dividing by the standard deviation. This is called a *z*-statistic. If, in assessing project 1, we want to find the probability that the cash flow will be between, for example, 3,200 and 5,000, we can apply the following formula:

$$z = \frac{X - \bar{R}}{\sigma}$$

where z = Number of standard deviations from the mean

X = Variable in which we are interested

\bar{R} = Expected value

σ = Standard deviation

Thus,

$$z = \frac{3,200 - 5,000}{1,095} = \frac{-1,800}{1,095} = -1.64$$

The quantity 3,200 is 1.64 standard deviations below the mean. Looking up 1.64 in Table A.2, we find the value to be 0.4495. Thus the probability that the cash flow will be between 3,200 and 5,000 is 45 percent. Because the left half of the normal curve represents 50 percent of all probabilities, we can also state that the chance of the cash flow being below 3,200 is 5.0 percent. Doing the same calculation for project 2 we obtain

¹⁹They are also symmetrical, which means that the observations to the left and to the right of the mean have the same probabilities and the same deviations from the mean.

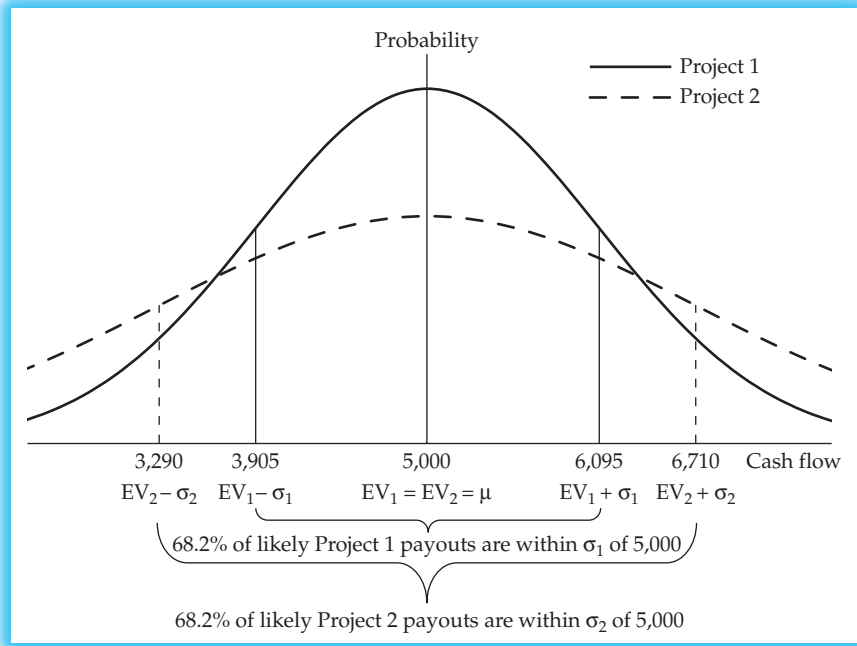


Figure 12.4 Continuous Distributions Assuming Normally Distributed Payouts for Projects 1 and 2

$z_2 = -1,800/1,710 = -1.05$, which has a corresponding value of 0.3531 in Table A.2. The chance of cash flow being less than \$3,200 for project 2 is 14.7 percent.

The Coefficient of Variation

When the expected values of two projects are equal, or at least close to one another, the standard deviation is a proper measure of risk. But because the standard deviation is an absolute measure, it may not serve our purposes if the two projects being compared have divergent expected values. For example, compare the two following hypothetical projects:

	Expected Value	Standard Deviation
Project A	\$100	30
Project B	50	20

Project A has both the larger expected value and the larger standard deviation.

Thus an absolute measure of risk may not give an adequate answer. In such cases, another concept is introduced, the **coefficient of variation**, which measures risk relative to expected value. The simple formula for the coefficient of variation is

$$CV = \sigma / \bar{R}$$

For the two projects,

$$CV_A = 30/100 = 0.30$$

$$CV_B = 20/50 = 0.40$$

The coefficient of variation is greater for project B. Because project A's expected value is greater and relative risk is smaller, project A is preferable.

The coefficient of variation will provide a satisfactory solution in most cases. However, when it does not, the businessperson will have to make a choice based on the perception of whether the risk is worth the potential return.²⁰

CAPITAL BUDGETING UNDER CONDITIONS OF RISK

So far the discussion of expected value and risk has been limited to the results of one period only. However, the assessment of risk is even more important when plans span a term of several years. So we now turn to the question of how to deal with a capital investment proposal when we include risk in the calculation.

The first task is to calculate the net present value of the expected values obtained in each year, or the expected net present value. For a 3-year project with one initial investment, we can use the following equation:

$$E \overline{\text{NPV}} = \frac{\overline{R}_1}{1 + r_f} + \frac{\overline{R}_2}{(1 + r_f)^2} + \frac{\overline{R}_3}{(1 + r_f)^3} - O_0$$

where $E \overline{\text{NPV}}$ = Expected net present value

\overline{R}_i = Expected values of the annual cash inflows

O_0 = Initial investment (outflow)

r_f = Risk-free interest rate

It should be noted that the expected cash flows are discounted at the riskless interest rate. Because risk is considered separately (in calculating the standard deviation), discounting at a rate that includes a risk premium could result in double counting risk.²¹

The standard deviation of the present value is

$$\sigma = \sqrt{\sum_{i=1}^n \frac{\sigma_i^2}{(1 + r_f)^{2i}}}$$

where σ = Standard deviation of NPV

σ_i = Standard deviation of each year's cash flow

Note that the exponent in the denominator of the expression is $2i$. Thus, the first year's σ will be discounted at $(1 + r_f)^2$, the second at $(1 + r_f)^4$, and so on.

A simple numerical example concludes this section. Table 12.5 shows that, for a 2-year project with a cash flow distribution and probabilities as specified, and with an initial investment of \$500, the expected net present value is \$430, and the standard deviation is 136.

²⁰The economic theory of the marginal utility of money can also provide a way to incorporate risk in decision making. Although such an approach is quite elegant, it is somewhat doubtful that it can be used in most practical situations. To apply this approach, one would have to know the utility function of the decision maker, or of the stockholders the decision maker represents. This appears to be a rather monumental task. However, in the final analysis, it is the decision maker's perception of risk that determines the shape of the utility function. Thus, the decision is subjective. For a more detailed explanation of utility, see, for example, Haim Levy and Marshall Sarnat, *Capital Investment and Financial Decisions*, 5th ed., Hertfordshire, England: Prentice Hall International, 1994, p. 221 ff. This exposition is based on the famous work of John Von Neumann and Oskar Morgenstern, *Theory of Games and Economic Behavior*, 2d ed., Princeton, NJ: Princeton University Press, 1953. The last problem in Chapter 11, which examines screening in insurance markets, provides an example of a Von Neumann-Morgenstern utility function that allows comparisons of risk and reward.

²¹Actually, the use of the proper discount rate is subject to some controversy. In discounting at the risk-free rate, we have followed the method recommended in several finance books. However, other books suggest that the company's (or the specific unit's or project's) risk-adjusted discount rate is more appropriate.

Table 12.5 Capital Budgeting Under Risk

		Expected Value					
Year 0		Year 1		Year 2			
		p	R	p	R		
-500		0.2	\$300	0.25	\$400		
		0.6	500	0.50	500		
		0.2	700	0.25	600		
\bar{R}			500		500		
$r_f = 0.05$ $\text{NPV} = 500/1.05 + 500/1.05^2 - 500$ $= 476 + 454 - 500 = 430$							
		Standard Deviation					
		Year 1		Year 2			
p	(R - R)	(R - R) ²	(R - R) ² p	p	(R - R)	(R - R) ²	(R - R) ² p
0.2	-200	40,000	8,000	0.25	-100	10,000	2,500
0.6	0	0	0	0.50	0	0	0
0.2	+200	40,000	8,000	0.25	+100	10,000	2,500
			\$16,000				\$5,000
$\sigma = \sqrt{16,000/1.05^2 + 5,000/1.05^4}$ $= \sqrt{14,512 + 4,114} = \sqrt{18,626} = 136$							

The preceding calculations are generally valid when the cash flows over the years are independent, that is, if in the case depicted in Table 12.5, the results in year 2 are not influenced by those of year 1.

TWO OTHER METHODS FOR INCORPORATING RISK

Two other techniques of accounting for risk are commonly used. Both make the risk adjustment within the present-value calculation (without the use of the standard deviation) so the final result is just one number: the net present value adjusted for risk. The two methods are

1. The **risk-adjusted discount rate (RADR)**, in which the risk adjustment is made in the denominator of the present-value calculation
2. The **certainty equivalent**, in which the numerator of the present-value calculation is adjusted for risk

The Risk-Adjusted Discount Rate

When we discussed the cost of capital (k) earlier in this chapter, we actually included a risk factor in our calculation. The cost of capital is composed of two components, the risk-free rate (r_f) and the risk premium (RP):

$$k = r_f + RP$$

The risk-free rate ideally represents the pure time value of money. It is usually represented by the yield on a short-term U.S. Treasury bill. The risk premium is a judgment as to the additional return needed to compensate for risk.

A company's cost of capital is the average required rate of return for all its parts combined. However, a company is composed of divisions (or even projects within a division) with different levels of risk. For example, a company with an average cost of capital of 10 percent may be composed of two same-size divisions (A and B) whose required rates of return are 8 percent (less risky division) and 12 percent (riskier division), respectively. It would be a mistake for the company to apply the 10 percent rate to all projects being considered in the two divisions. Thus a project with an IRR of 9 percent would be acceptable in division A, whereas a project with an IRR of 11 percent should be rejected in division B.

Of course, it must be recognized that developing RADRs involves a large amount of judgment. But even if such adjustments are necessarily judgmental, they are very important. Companies with staffs capable of such refinements have developed some methods for differentiating discount rates.²²

Certainty Equivalents

In calculating the RADR, the inclusion of risk in the calculation of present value is accomplished by altering the discount rate or the cost of capital, that is, the denominator of the discounting equation. Another technique for including risk in the calculation of present values is to work through the numerator of the cash flow fraction; that is, the cash flow itself is adjusted to account for risk. Basically, this is accomplished by applying a factor to the cash flow to convert a risky flow into a riskless one. To accomplish this, the risky cash flow must be reduced by some amount, or multiplied by a number smaller than 1. We refer to this adjustor as the *certainty equivalent factor*.

Just as in the case of the RADR, certainty equivalents require a significant amount of judgment. The size of the certainty equivalent factor depends on the decision maker's attitude toward risk. Thus, if he decides that an expected risky cash flow of \$100 is equivalent to a risk-free cash flow of \$95, the certainty equivalent factor equals 0.95.

So, for each risky cash flow, R_t , a certainty equivalent factor a_t is assigned. If risk increases as a function of time, the certainty equivalent factors will decrease as we move into the future. For instance, a project could have the following cash flows and certainty equivalent factors:

Period	R_t	a_t	$a_t R_t$
1	\$100	0.95	\$ 95
2	200	0.90	180
3	200	0.85	170
4	100	0.80	80

The risk-free cash flows, $a_t R_t$, are obviously smaller than the risky flows, R_t , as would be expected for a risk-averse investor. These risk-free cash flows are then discounted at the risk-free interest rate to obtain the present value of the cash flows.

²²See Grenville S. Andrews and Colin Firer, "Why Different Divisions Require Different Hurdle Rates," *Long Range Planning*, 20, 5, 1987, pp. 62–68.

Present Value Break-Even Analysis

In Appendix 8B we introduced break-even analysis. In that case, however, we used accounting data and arrived at an accounting profit. Now that we have learned about the calculation of present values, we can show a break-even analysis using this method. Present value break-even analysis can be used as an adjunct to sensitivity analysis. It tells us in terms of cash flows how much production we need to break even.

We now include the investment in equipment and depreciation in our calculations. We also have to consider taxes. Assume the following data:

Investment	\$10,000
Annual fixed costs (TFC)	3,000
Sales price per unit (P)	10
Variable cost per unit (AVC)	3
Discount rate (cost of capital)	10%
Tax rate (t)	35%
Straight-line depreciation over 5 years	

In order to put our investment on an annual basis we have to use a tool called *annual equivalent cost (AEC)*. We do that by converting the original investment into an annuity. We divide the investment by the present value factor of an annuity at 10 percent for 5 years, or $10000/3.7908 = 2638$.

The formula for the present value break-even is as follows:

$$\begin{aligned} \text{BE} &= \frac{\text{AEC} + \text{TFC}(1 - t) - \text{Depreciation} \times t}{(P - \text{AVC})(1 - t)} = \frac{2638 + 3000(0.65) - 2000(0.35)}{(10 - 3)(0.65)} \\ &= (2638 + 1950 - 700)/4.55 = 3888/4.55 = 855 \end{aligned}$$

The quantity of 855 will create a break-even situation. Had we used the accounting method, the break-even point would be lower.²³

SENSITIVITY AND SCENARIO ANALYSIS

Sensitivity analysis and scenario analysis are both pragmatic ways to estimate project risk.

Sensitivity analysis involves the changing of a key variable to evaluate the impact that the change will have on the results of the capital budgeting analysis.

A simple example illustrates this procedure. Assume we are analyzing a 10-year capital investment project that will require an investment of \$50,000. Our best estimate of net cash inflows per year is \$10,000, and our cost of capital is 10 percent.

Our best (base) estimate may, however, be incorrect. The cash flows from the project may be higher or lower, depending on such factors as the economy, inflation, and competition. So we make two more forecasts, a pessimistic one and an optimistic one. We estimate that under the worst circumstances our annual net cash inflows will be \$8,000, and if all circumstances are favorable, they will be \$12,000.

²³Using the accounting method, we would add depreciation to the fixed cost. After all, depreciation is a fixed cost. The result would have been: $(3000 + 2000)/(10 - 3) = 5000/7 = 714$.

Table 12.6 Sensitivity Analysis

Estimates	Pessimistic	Base	Optimistic
Initial investment	\$50,000	\$50,000	\$50,000
Net cash flows	8,000	10,000	12,000
Present value	\$ -843	\$11,446	\$23,735

Table 12.6 shows the results of our calculation. If the pessimistic conditions prevail, then the project will not be acceptable. At this point, it will be necessary for management to decide what is the probability of the worst case. If it is relatively small, then the project would probably be implemented. A shortcoming of sensitivity analysis is that it considers only the change in one variable.

Scenario analysis is similar to sensitivity analysis but corrects the latter's shortcoming. It takes into consideration the changes in several important variables simultaneously. Thus, in the case of "The Situation" vignette at the beginning of the chapter, we could consider possible changes in the sales quantity, sales price, production costs, and the cost of capital, to name just a few. George Kline will consider these potential variations, and we discuss their impact on the capital project in "The Solution" vignette later in this chapter.

Sensitivity analysis and scenario analysis are commonly used in business. Their outcomes can be displayed in a simple, straightforward manner. These analyses permit analysts (and their managers) to evaluate each important variable and examine the trade-offs among them. They can easily use a spreadsheet program to generate alternative results quickly.

SIMULATION

Although the sensitivity and scenario analysis techniques are popular with business, they do not make use of probability distributions. There is one method that does: simulation. In **simulation analysis**, each key variable is assigned a probability distribution. Let us say that we have estimated the sales revenue for a particular project. We can then assign it a probability distribution that will show the probabilities of the estimated value and deviations from it:

Deviation from Estimate Value (%)	Probability	Cumulative Probability
-30	0.1	0.1
-15	0.2	0.3
0	0.4	0.7
+15	0.2	0.9
+30	0.1	1.0

The column of cumulative probabilities indicates that there is a 10 percent chance that sales will be 30 percent lower than the base case. Furthermore, the probability is 30 percent that sales will be at least 15 percent below the base case, 70 percent that sales will not exceed the base estimate, and so on. Similar distributions would be estimated for the other important variables.

The device of random number is used to "simulate" a possible outcome. Suppose a random number generator with numbers from 1 to 100 is used. For this case,

we can assign numbers from 1 to 10 to represent the -30 percent case. Any number between 11 and 30 (which has a 20 percent chance of being drawn) would stand for a -15 percent sales situation. All numbers between 31 and 70 would represent 0 percent deviation (the base case estimate), and so on.

We also assign probability distributions to other key variables, such as production costs, expenses, and capital investment. The next step is to generate a random number for each key variable, obtain the appropriate values, and calculate an NPV figure. This process will then be repeated a large number of times, each time generating another NPV figure. The NPVs thus generated will form a probability distribution and also enable the analyst to calculate a standard deviation as well as a *z*-statistic.²⁴ A large number of iterations can be obtained quickly and effortlessly by using a computer.

Simulation can be a good tool for decision making. However, the illustration used here is severely simplified and may not be sufficient for the solution of complex business problems. In obtaining the preceding solution, we made at least two assumptions that may have omitted some important relationships among the variables. First, we assumed the deviations obtained with the use of random numbers remain the same in each year for which estimated cash flows were calculated. This need not be the case. A set of different random number calculations for each year may have been more appropriate. With the use of a computer, such calculations could have been taken care of quite efficiently.

Even more important, it has been assumed here that the variables are statistically independent. It is much more likely that the various factors are interrelated. For example, a shortfall in market demand may have a negative effect on the price. An unexpected increase in sales may have an effect on costs: in the short run, cost per unit may rise as plant employees work overtime at increased wages. If such interdependencies actually exist, they must be included in the simulation model. Such a model would, of course, be considerably more complex. That is the reason why, in the past, simulation analysis was not generally used by business. However, a number of software packages have been introduced that significantly facilitate such an analysis. Two of the more popular programs are @Risk and Crystal Ball. Both of these can be added to Excel programs.

Even though such models present the manager with a significant amount of useful information, the final decision, as in all cases, would still have to be based on the decision maker's judgment. In other words, no amount of data and information will substitute for mature business thinking.

DECISION TREES

One other method for making decisions under conditions of risk is the **decision tree**. This technique is especially suitable when decisions have to be made sequentially, for instance, if a decision 2 years hence depends on the outcome of an action undertaken today. The use of a tree diagram facilitates the process because it illustrates the sequence in which decisions must be made. It also compares the values (e.g., NPVs) of the various actions that can be undertaken.

The best way to explain the technique is to use a relatively uncomplicated example. A company has an opportunity to purchase a patent for the manufacture of a new product for \$200,000. It has three possible choices:

- It does not purchase the patent.
- It purchases the patent at the above price.
- It spends an additional \$50,000 on a feasibility study before purchasing the patent.

²⁴The simulation procedure discussed here is based on a technique introduced by David B. Hertz in "Risk Analysis in Capital Investment," *Harvard Business Review*, January-February 1964, pp. 95-106, and "Investment Policies That Pay Off," *Harvard Business Review*, January-February 1968, pp. 96-108.

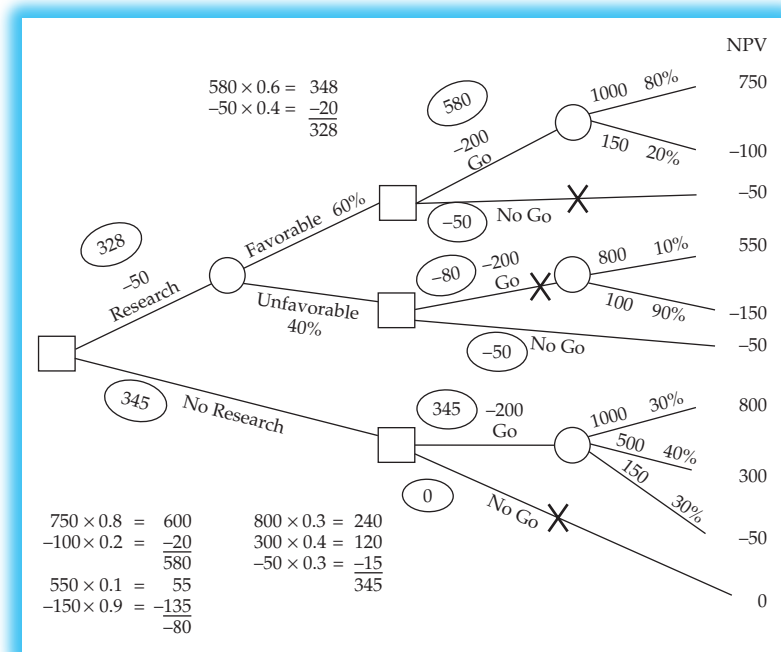
The company analysts have developed the following estimates:

- Probability that additional research will find the product to have good potential is 60 percent. If the research results are favorable, there is an 80 percent probability that the product will net the company \$1 million; a 20 percent probability that income will be only \$150,000. If the research results are unfavorable, there is a 90 percent probability that income will be \$100,000 and a 10 percent probability that it will be \$800,000.
- If the company purchases the patent without further research, the income estimates are as follows: 30 percent probability—\$1 million, 40 percent probability—\$500,000, and 30 percent probability—\$150,000.

Figure 12.5 illustrates the decision tree with the resulting calculations. The first step in the analysis is to set up all the “branches” of the decision tree. As we move from left to right on the diagram, we are faced with decision points and with chance events. On the diagram, decision points are designated with squares and chance events with circles. When the entire tree is completed, the procedure is to move back from right to left, calculate the value of each branch, and where appropriate combine or eliminate branches.

Starting with the highest branch, which represents the situation in which research is conducted before purchasing the patent and the research results are favorable, the expected income is \$830,000 ($0.8 \times \$1,000,000$, plus $0.2 \times \$150,000$). We then subtract the cost of the patent and research, \$250,000, and arrive at a final figure, \$580,000. Because not buying the patent if the market research is favorable would result in a net loss of \$50,000 (the cost of research), we eliminate that particular decision. Now, if the result of the research is unfavorable, the company would not proceed with the purchase of the patent because the loss is only \$50,000 compared with a loss of \$80,000 if the company does purchase it.

Figure 12.5 Decision Tree



Thus, if the company does the research, the resulting NPV is \$328,000 because there is a 60 percent probability of favorable results and 40 percent unfavorable ($0.6 \times 580,000 + 0.4 \times -50,000 = \$328,000$).

Next we look at the results of not doing research. Proceeding the same way as we did previously, if the company buys the patent the expected NPV will be \$345,000, whereas if it does not go ahead, the result is, of course, an NPV of \$0.

Thus, purchasing the patent without additional research appears to be the better alternative.

The preceding solution is not quite complete because the decision is being made on the basis of expected NPV alone. There has been no calculation of standard deviations, so we have, in effect, ignored the differences in risk between purchasing the patent with or without additional research. As mentioned in footnote 20 of this chapter, risk can also be measured by the use of utility analysis (this is the approach taken in the analysis of screening in insurance markets in Chapter 11, Problem 11). Each outcome could have been assigned an expected utility, and instead of maximizing NPV, as we did in this example, the objective would have been to maximize utility.

REAL OPTIONS IN CAPITAL BUDGETING

The subject of decision trees, just discussed, points to a possible event we have not addressed previously: A capital budgeting decision does not have to be set in concrete at the beginning of the project.

Up to now, in analyzing capital budgeting decisions, we made estimates of costs, cash flows, life of the project, and probabilities of outcomes, and then proceeded to calculate the NPV or the IRR. But we did not consider that there might be an opportunity for making changes in some aspects of the project while it is in progress or to make adjustments even before the project is started. The ability to make changes is commonly referred to as **real options** that are embedded in the capital project.²⁵

These real options may increase the value of a project above that resulting from a straightforward discounted cash flow calculation. The value of the option is the difference between the project's value with and without the option. We can represent this result with the following simple equation:

$$\text{Value of the project} = \text{NPV} + \text{option value}$$

Indeed, if such an option is taken into consideration, it may result in the acceptance of a project that was considered unacceptable without the option. (In other words, it can convert a project with a negative NPV to one with a positive NPV.)²⁶

There are various forms of real options:

1. **Option to vary output.** Some projects can be structured to permit operations to expand if demand rises above expectations (this is usually called a *growth option*), or to contract operations if demand falters. It may even allow for a temporary shutdown of production.
2. **Option to vary inputs—flexibility.** A plant may include the possibility of operating with different types of fuel. Although the original cost of the plant may be higher, switching from a more expensive fuel to a less expensive one may make the project more

²⁵Real options have a similarity to financial options, which are an important subject in managerial finance. An option gives the holder the right, but not the obligation, to take an action in the future.

²⁶For a more detailed introduction to real options, consult the following articles: A. K. Dixit and R. S. Pindyck, "The Options Approach to Capital Investment," *Harvard Business Review*, May–June 1995, pp. 105–15; and N. Kulatilaka and A. J. Marcus, "Project Valuation Under Uncertainty; When Does DCF Fail?" *Journal of Applied Corporate Finance*, 5, 3 (fall 1992), pp. 92–100. A leading book in this field is Lenos Trigeorgis, *Real Options*, Cambridge, MA: MIT Press, 1997. For an article on this subject, see P. Coy, "Exploiting Uncertainty," *BusinessWeek*, June 7, 1999, pp. 118–24.

favorable. Another potential flexibility is that of taking advantage of different technologies, depending on input costs. This is not restricted to fuel; for example, the same could be said of a processed-food manufacturer who substitutes various sweetening inputs based on price.

3. **Option to abandon.** If, after the project has been started, results turn sour, it is possible that abandoning the project will improve its payoff. If the project can be sold for a price higher than the present value of its expected cash flows, or if its facilities can be used more favorably in another part of the company, then abandoning the project will enhance its value.
4. **Option to postpone.** An oil company may profit by postponing the extraction of oil from its field if current prices are low but are expected to rise in the future. A company may postpone the introduction of a new product pending completion of market research to better evaluate the product's potential. The same analysis applies to any company extracting nonrenewable natural resources.

In both of these cases there is, of course, a cost involved in the postponement, whether it is the cost of delaying cash inflows or of incurring the expense of market research. However, the benefits of the postponement may exceed these costs, and thus create a positive value for the postponement option.

Another reason for postponement may be the expectation of decreases in interest rates. Lower interest rates would decrease the project's required rate of return and thus increase its present value. It must be remembered, however, that by postponing action, the company may lose the "first-mover" advantage.

5. **Option to introduce future products.** A company may be willing to launch a product with a negative NPV if doing so gives it an option to gain an advantage when later versions of the product are introduced.

Real Options in Practice

It is generally recognized that real option analysis enhances the making of investment decisions over the usual capital budgeting calculations because it allows for the flexibility of changing decisions during the project's life. Classic net present value analysis may result in an understatement of a project's value. Thus, it is rather surprising that real options have not become more popular with large corporations. A recent study of 279 corporations (taken from the 1,000 largest *Fortune* companies) showed that only 40—or 14.3 percent—employed real option analysis. Technology and energy firms were the most frequent users of real options. The companies that did not use real options gave various reasons for their nonuse. The most frequent ones were lack of top management support, discounted cash flow being a proven method, and real option analysis being too sophisticated. However, a large number of nonusers expected to begin using real options in the future.²⁷

An Abandonment Option

We conclude this brief introduction to real options with a simple example. By including the option to abandon the project, we can convert a project with a negative NPV to a positive one.

The data for this project are as follows (all dollar amounts are in thousands):

1. The life of the project is 2 years.
2. The original cash outflow is \$8,000.
3. The potential cash inflows in year 1 are \$3,000 or \$6,000, each with a probability of 50 percent.

²⁷Stanley Block, "Are 'Real Options' Actually Used in the Real World?" *The Engineering Economist*, Vol. 53, Issue 3, 2007.

4. If the first year's cash inflow is \$3,000, the second year's cash inflows will be either \$2,000 or \$4,000, each with a probability of 50 percent.
5. If the first year's cash inflow is \$6,000, the second year's cash inflows will be \$5,000 or \$7,000, each with a probability of 50 percent.
6. The project can be abandoned at the end of year 1. The expected abandonment value is \$3,500.
7. The project's discount rate is 10 percent.

The calculations are shown in Table 12.7. The upper third of the table shows the NPV calculation if abandonment is not included. The cash flows and their present value at period 0 are shown. Then the probabilities are calculated. Each potential

Table 12.7 An Abandonment Option

Expected Net Present Value Calculation								
Period 1		Period 2		Total PV ₀	P ₁	P ₂	P ₁ × P ₂	Expected PV ₀
CF	PV ₀	CF	PV ₀					
\$3,000	\$2,727	\$2,000	\$1,653	\$4,380	0.5	0.5	0.25	\$ 1,095
		4,000	3,306	6,033		0.5	0.25	1,508
6,000	5,455	5,000	4,132	9,587	0.5	0.5	0.25	2,397
		7,000	5,785	11,240		0.5	0.25	2,810
Total expected present value								\$ 7,810
Initial investment								-8,000
Expected net present value								\$ -190
Present Value of Period 2 Cash Flows at End of Period 1								
Period 2		PV ₁	P ₂	Expected PV ₁				
CF								
\$2,000		\$1,818	0.5	\$ 909				
4,000		3,636	0.5	1,818				
				\$2,727				
5,000		4,545	0.5	2,273				
7,000		6,364	0.5	3,182				
				\$5,455				
Net Present Value with Abandonment								
Period 1		PV ₀	P ₁ × P ₂	Expected PV ₀				
CF								
\$3,000		\$ 5,909*	0.5	\$ 2,955				
		9,587	0.25	2,397				
6,000								
		11,240	0.25	2,810				
Total expected present value				\$ 8,161				
Initial investment				-8,000				
Expected net present value				\$ 161				

*(3,000 + 3,500)/1.1 = 5,909

outcome has a probability of 25 percent. The present value of the cash flows is \$7,810, and the expected NPV is a negative \$190. Thus, the project would be rejected.

In the second third of the table, we show the present value at the end of period 1 of the year 2 cash inflows. We can see that if the first year's cash inflow was \$3,000, and the second year's were either \$2,000 or \$4,000, the PV of these two cash flows at the end of year 1 is \$2,727, less than the abandonment value. The expected PV_1 of the two higher cash flows is \$5,455, well above the abandonment value.

In the last section of Table 12.7, we substitute the abandonment value of \$3,500 into the calculation, and compute the expected PV_0 . The expected PV_0 is now \$8,161, which gives us a positive NPV of \$161. With the abandonment option included, this has now become an acceptable project. The value of the option to abandon is \$351.

We did not calculate the standard deviation for this problem. However, had we done so, it would have been lower when the option was considered than when we calculated a straight NPV without the option. Thus, inclusion of the option did not only increase the NPV of the project, but it also decreased its riskiness.

GLOBAL APPLICATION

Political Risk

Political risk can be defined as an action by a foreign government that is detrimental to the firm. Four types of such action can be enumerated:

1. **Regulation:** Such actions can include changes in taxes, labor law rules, minimum wages, and price controls. It should be noted that such regulation can also affect local companies in the foreign country.
2. **Discrimination:** Potential actions include restrictions on the repatriation of dividends, special labor conditions, tariff and nontariff barriers, and also imposition of administrative rules (red tape) that will make operations prohibitively expensive. Such actions are probably the most frequent.
3. **Expropriation:** A government takes over foreign property, usually with the intention of operating the business itself. Expropriation can be done with fair compensation and with inadequate or no compensation.
4. **Wars and disorders:** These can lead to destruction of a firm's property.

It is important for a company to attempt to forecast political risk using the best information and advice that it can obtain. There is also the potential of insuring through the U.S. government-owned Overseas Private Investment Corporation. This insurance can be obtained in dealing with developing countries and covers inconvertibility, expropriation, war, and political violence. Furthermore, a company can also reduce political risk by negotiating agreements with the government of the foreign country. There are also operating and financial strategies a company can adopt to decrease political risk. This brief discussion of financial risk cannot go into the details of such strategies; thorough descriptions can be obtained from books dealing with multinational corporate finance.²⁸

We discuss political and economic risk further in Chapter 13.

²⁸Two of the books on this subject, which have been mentioned previously, are David K. Eiteman, Arthur I. Stonehill, and Michael H. Moffett, *Multinational Business Finance*, 11th ed., Boston, MA: Addison-Wesley, 2007, Chapter 17, and Dennis J. O'Connor and Alberto T. Bueso, *International Dimensions of Financial Management*, New York: Macmillan, 1990, Chapters 6 and 11. Materials in these two texts have been used in this narrative.

Containing International Risk

Operating internationally entails some additional risk, but a company can manage these exposures (and possibly benefit) by diversifying its operations and finances.

Suppose that, due to a temporary disequilibrium, costs of production among different countries diverge. The company may be able to shift its production and sources of materials and components from one country to another. Or, if product prices and profitability shift among countries, so can a company's marketing efforts. If temporary deviations occur in interest or exchange rates, a company may find it possible to decrease its cost of capital by moving its financing sources. Finally, a diversification strategy may avoid some of the dangers of political risk.²⁹

Capital Budgeting in Practice

Finally we include here three examples of capital budgeting practices in countries other than the United States.

Capital Budgeting in Croatia³⁰

The authors surveyed 59 companies in both nonfinancial and financial sectors to find the most commonly used capital budgeting methods. The responses were grouped into four categories: always, often, sometimes, never. Here are the results in percentage terms:

	Always	Often or Sometimes	Never
Internal rate of return	59%	22%	19%
Payback period	56	27	17
Net present value	42	34	24
Discounted payback	27	38	36
Profitability index	22	32	46

Again, as we mentioned previously, the IRR method is preferred to NPV, contrary to the viewpoint of financial experts. Also, the payback period is still very much in use.

Forecasting Errors in Capital Budgeting in Portugal³¹

Capital budgeting analysis is always an estimate of the future. So it is quite interesting to investigate whether the actual results correspond to the estimates. The authors had access to companies' applications for investment subsidies that were presented with the data they needed to do this study. The study examined a sample of 264 companies of different sizes and different industries. The results showed that on average, forecasted sales exceeded the actual by about 9 percent. The forecasts of costs appeared to be quite accurate, with the exception of personnel costs, which were understated by about 3 percent. Thus, profits were severely affected, showing that earnings before interest and taxes (EBIT) decreased by some 42 percent.

²⁹A more complete discussion of this subject can be found in Eiteman, Stonehill, and Moffett, op cit, pp. 310–312, 436.

³⁰Lidija Dedi and Silvije Orsag, "Capital Budgeting Practices: A Survey of Croatian Firms," *South East European Journal of Economics and Business*, Vol. 2, Issue 1, April 2007, pp. 59–67.

³¹Joao Oliveria Ssoares, Maria Cristina Coutinho, and Carlos V. Martins, "Forecasting Errors in Capital Budgeting: A Multi-Firm Post-Audit Study," *Engineering Economist*, Vol. 52, Issue 1, 2007, pp. 21–39.

The authors state that there is a potential bias in the results. There is the possibility that in applying for subsidies the companies may have overstated profitability in order to improve the probability of approval. These biases are consistent with those noted by Pruitt and Gitman, which were described earlier in this chapter (see footnote 12).

Capital Budgeting Practices in South Africa³²

This article is a review of changes in capital budgeting practices in South Africa from 1972 to 2008. During this period, the use of discounted cash flow methods has increased significantly. There has been a dramatic increase in the use of NPV, while there was a decrease in the use of payback and the accounting rate of return. The internal rate of return still remains a primary method of project evaluation. In the most recent survey for the year, 2006 (the article was published in 2008), NPV was used by 82 percent of the companies surveyed, IRR by 79 percent, payback by 54 percent, and the accounting rate of return by 14 percent. For project risk analysis, sensitivity analysis was the most popular method, while scenario analysis exhibited significant growth. However, there was little growth in simulation and decision tree analyses. As could be expected, NPV and IRR methods tend to be used by large firms, while the payback method is still popular with smaller firms.



Module 12G

The Solution

George Kline is now ready to put some numbers together. He assumes the market forecast of 100 million cases pertains to Global's first year of production. Because the market research people expect, rather optimistically, that market share will grow to 5 percent after 4 years, George sets market share at 2, 3, and 4 percent for the interim years. He calculates the Modified Asset Cost Recovery System (MACRS) depreciation for the equipment according to the 1986 tax law. The depreciation percentages using this method will be 14, 25, 17, 13, 10, 9, 9, and 3 percent for years 1 to 8, respectively. Depreciation will start in year 2; thus the equipment will not be fully depreciated after 6 years. George assumes the remaining book values will be the market value at the end of year 7. He also decides that, of the original \$750,000 investment in working capital, 80 percent will be recovered as cash at the end of year 7 (there will be some uncollectible accounts receivable and obsolete inventory). The 20 percent loss is tax deductible. To simplify the worksheet, George combines product and distribution costs into one line; he also combines variable General and Administrative expenses and advertising expenses.

George now prepares his worksheet (Table 12.8). The column labeled "Constant" contains the parameters established originally. They can be changed. Most numbers in this table are self-explanatory. The startup expenses are tax deductible; thus the cash outflow is \$450.

When the cash flows are obtained and discounted at the 15 percent required rate of return, the resulting NPV is a positive \$3,552,000. The IRR is 24.2 percent—considerably above the required rate.

George could therefore recommend the acceptance of this expansion project. However, because all the cash flows are future estimates, there is probably considerable uncertainty about their accuracy. He decides to perform a risk analysis. He settles on using scenario analysis for his presentation to management. The chart he presents to management (Table 12.9) will show three summary estimates—worst case, base case, and best case. The worst case will show all important variables on the unfavorable side and the best case on the opposite. Because of the newness of the expansion and severe competition, probably the riskiest of the forecasts are sales volume and sales price, particularly on the downside. So, he decreases the base estimates by 20 percent for the worst case and increases them by only

(continued)

³²C. Correia, "Capital Budgeting Practices in South Africa: A Review," *South African Journal of Business Management*, 2012, 43(2), pp. 11–29.

Table 12.8 Expansion Project

	Constant	Sensitivity Factor	New Constant	Year 0	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7
Total market	4.0%	1.00	4.0%			\$100,000	\$104,000	\$108,160	\$112,486	\$116,986	\$121,665
Market share						1.0 %	2.0 %	3.0 %	4.0 %	5.0 %	5.0 %
Company sales	1.00	1.00	1.00			1,000	2,080	3,245	4,499	5,849	6,083
Expenditures				\$-5,000	\$-2,000						
Working capital					-750						
Startup expense	750	1.00	750		-450						
Sales	5.00	1.00	5.00			\$5,000	\$10,400	\$16,224	\$22,497	\$29,246	\$30,416
Prod. and dist. cost	3.10	1.00	3.10			3,100	6,448	10,059	13,948	18,133	18,858
Variable expenses	8.0 %	1.00	8.0 %			400	832	1,298	1,800	2,340	2,433
Fixed costs	500	1.00	500			500	500	500	500	500	500
Depreciation (plant)	3,500	1.00	3,500			111	111	111	111	111	111
Depreciation (equipment)	3,000	1.00	3,000			420	750	510	390	300	270
Total Cost and Expense						4,531	8,641	12,478	16,749	21,384	22,173
NEBT						469	1,759	3,746	5,748	7,863	8,244
Income tax	40.0 %	1.00	40.0 %			188	704	1,498	2,299	3,145	3,298
NEAT						281	1,055	2,248	3,449	4,718	4,946
Add Depreciation						531	861	621	501	411	381
Operating cash flow						812	1,916	2,869	3,950	5,129	5,327
Remaining values											
Land	\$500	1.00	\$500								500
Plant	2,833	1.00	2,833								2,833
Equipment	360	1.00	360								360
Working capital	1.00	0.80	0.80								660
Total cash flow				\$-5,000	\$-3,200	\$812	\$1,916	\$2,869	\$3,950	\$5,129	\$9,680
NPV	15.0 %										3,552
IRR											24.2%

5 percent for the best case. He further estimates that production and distribution costs and variable expenses should deviate from the base case by 10 percent.

Table 12.8 includes two columns that will help him make his calculations quickly. He enters the deviations from the base case in the "Sensitivity Factor" column (e.g., 0.8 would indicate a 20 percent decrease from the base case) to produce the "New Constant."

The worst case turns out to have a very negative NPV and an IRR close to zero. (See Table 12.9.) To further prepare his presentation, George calculates several cases for which NPV is near zero. He is thus able to show the magnitude of errors that would cause the company to be indifferent to this proposal. These are shown in Table 12.10. Many other calculations could have been made, but at this point George believes that he has a good handle on his presentation.

(continued)

Table 12.9 Three Possible Scenarios

	Worst Case	Base Case	Best Case
Percentage differences			
Sales	-20	0	+5
Sales price	-15	0	+5
Prod. and distr. costs	+10	0	-10
Variable expenses	+10	0	-10
Net present value (\$000)	-3,909	3,552	8,093
Internal rate of return	1.6 %	24.2 %	33.9 %

When the time for the presentation to top management arrived, George, accompanied by the treasurer, brought all his charts to the conference room. George proceeded with his discussion, summarizing the forecasts and estimates and arriving at the recommended solution. He reviewed with management his sensitivity analysis, and the treasurer completed the presentation by recommending that the company go ahead and begin production in the new region.

George believed that his presentation went well, and the treasurer commended him for a good job. But a few days later, the treasurer called George into his office and told him that management had, at least for the time being, decided not to go ahead with expansion. He explained that management believed that this was not the right time to expand, given the economy's somewhat cloudy future. Also, they considered the strength of their competition and decided that this was not the time to make the investment. Sensing George's disappointment, the treasurer assured him that he had done an excellent job with the data he had. Actually, the treasurer said, the vice president of finance had told him to commend George for his substantial effort. "But that's the way things go in a large corporation," he continued. "You see, decisions are not always made on the basis of figures, however thoroughly they were developed. You must remember that in business, once the data have been put together, there is still the judgment factor that must be applied by management. After all, that's what top management gets paid for. I am sure that, in the future, this or a similar project will be considered again."

Table 12.10 Additional Scenarios

	Case 1	Case 2	Case 3
Percentage differences			
Sales	-10	-5	-7
Sales price	-6	-7	-5
Prod. and distr. costs	+5	+7	+7
Variable expenses	0	+5	+4.5
Net present value (\$000)	-8	-2	0
Internal rate of return	15.0 %	15.0 %	15.0 %

SUMMARY

In this chapter, we expand the economic concept of profit maximization to multiperiod projects.

Capital budgeting involves the evaluation of projects in which initial expenditures provide streams of cash inflows over a significant period of time. The process of evaluating capital proposals includes the following:

1. Estimating all incremental cash flows resulting from the project
2. Discounting all flows to the present
3. Determining whether a proposal should be accepted

Two methods are recommended for evaluating capital budgeting proposals—NPV and IRR. The validity of these two criteria are compared. We find that, from a theoretical viewpoint, NPV is the more valid. However, there is much to recommend the use of IRR, and business, in fact, favors this technique. In most cases, both methods lead to the same answer.

The concept of the cost of capital was then developed, and methods of arriving at a weighted cost of capital are discussed. Capital rationing was also discussed.

We then turn to the second subject of this chapter, risk analysis as applied to capital budgeting. The most common measures of risk are discussed: standard deviation, coefficient of variation, and the *z*-statistic. We use the concepts of expected value and standard deviation to describe the attributes of capital budgeting for risky projects.

Two alternative calculations incorporating risk in capital budgeting are presented: risk adjusted discount rate and certainty equivalents. Although both methods can be employed, the former is simpler to use and is much more popular in business.

Additional techniques are discussed. Two similar methods are sensitivity analysis and scenario analysis. The latter is used in “The Situation” and “The Solution” vignettes. We further present a brief discussion of simulation analysis and decision trees. Simulation analysis is capable of obtaining expected value and standard deviation, whereas decision trees lend themselves to sequential decision making.

Although many techniques of accounting for risk were discussed in this chapter, none of these methods is completely satisfactory. However, the important lesson of this chapter is that risk is always present in business, and anyone engaging in business planning must be aware of the dangers of risky outcomes and be able to cope with the uncertainty of future events. Thus, the awareness of a risky situation may be more important than familiarity with any of the specific methods illustrated in this chapter.

IMPORTANT CONCEPTS

Accounting rate of return. Also known as the *return on investment* (ROI) or *return on assets* (ROA), a method for evaluating capital projects. It is obtained by dividing the average annual profit by the average investment. (p. 464)

Capital asset pricing model (CAPM). A financial model specifying relationships between risk and return. An important part of the CAPM is the development of beta, which measures the market risk of a security and is a necessary ingredient in determining a stock’s required rate of return. (p. 474)

Capital budgeting. An area of business decision making that concerns undertakings whose receipts and expenses continue over a significant period of time. (p. 463)

Capital rationing. The practice of restricting capital expenditure to a certain amount, possibly resulting in the rejection of projects that have a

positive net present value and should be accepted to maximize the company’s value. (p. 477)

Certainty equivalent. A certain (risk-free) cash flow that would be acceptable as opposed to the expected value of a risky cash flow. (p. 484)

Coefficient of variation. A measure of risk relative to expected value that is used to compare standard deviations of projects with unequal expected values. (p. 482)

Cost of capital. Also often referred to as the required rate of return, the hurdle rate, or the cutoff rate, the rate of return a company must earn on its assets to justify the using and acquiring of funds. (p. 465)

Decision tree. A method used with sequential decision making in which a diagram graphically points out the order in which decisions must be made and compares the value of the various actions that can be undertaken. (p. 488)

- Discount rate.** The rate at which cash flows are discounted. It is the required rate of return or cost of capital. (p. 465)
- Dividend growth model.** A method to arrive at the value of a security. Given the price of the security, it calculates the company's cost of equity as the dividend divided by the current stock price plus the growth rate in dividends (assumed to be constant). It is an alternative method to CAPM in calculating the equity cost of capital. (p. 473)
- Expected value.** An average of all possible outcomes weighted by their respective probabilities. (p. 479)
- Free cash flow.** Funds that are available for distribution to investors. It includes operating cash flow after taxes minus (plus) increases (decreases) in operating working capital and fixed assets investment. (See Appendix 12A.) (p. 507)
- Independent projects.** A situation in which the acceptance of one capital project does not preclude the acceptance of another project. (See *mutually exclusive projects*.) (p. 467)
- Internal rate of return (IRR).** One method of evaluating capital projects by discounting cash flows. The IRR is the interest rate that equates the present value of inflows with the present value of outflows, or, in other words, causes the net present value of the project to equal zero. (p. 464)
- Mutually exclusive projects.** A situation in which the acceptance of one project precludes the acceptance of another. (p. 467)
- Net present value (NPV).** A method of evaluating capital projects in which all cash flows are discounted at the cost of capital to the present and the present value of all outflows is subtracted from the present value of all inflows. (p. 464)
- Payback.** A method of evaluating capital projects in which the original investment is divided by the annual cash flow. It tells management how many years it will take for a project's cash inflows to repay the original investment. (p. 464)
- Probability.** An expression of the chance that a particular event will occur. (p. 478)
- Probability distribution.** A distribution indicating the chances of all possible occurrences. (p. 478)
- Profitability index (PI).** A method of evaluating capital budgeting projects by dividing the present value of cash inflows by the original investment. (p. 467)
- Real option.** An opportunity to make adjustments in a capital budgeting project in response to changing circumstances potentially resulting in improved results. (p. 490)
- Risk.** Refers to a situation in which possible future events can be defined and probabilities assigned. (p. 477)
- Risk-adjusted discount rate (RADR).** A value equal to the riskless (risk-free) interest rate plus a risk premium. The risk-free rate ideally is the pure time value of money, and the risk premium represents a judgment as to the additional return necessary to compensate for additional risk. (p. 484)
- Scenario analysis.** A method for estimating project risk. Key important variables are identified and changed simultaneously to measure the impact of their changes on results of a capital budgeting proposal. (p. 487)
- Sensitivity analysis.** A method for estimating project risk that involves identifying a key variable that affects results and then changing it to measure the impact. (p. 486)
- Simulation analysis.** A method that assigns a probability distribution to each key variable and uses random numbers to simulate a set of possible outcomes to arrive at an expected value and dispersion. (p. 487)
- Standard deviation.** The degree of dispersion of possible outcomes around the mean outcome or expected value. It is the square root of the weighted average of the squared deviations of all possible outcomes from the expected value. (p. 480)
- Time value of money.** Very basically, this means that a dollar today is worth more than a dollar tomorrow because today's dollar will earn interest and increase in value. (p. 464)
- Uncertainty.** Refers to situations in which there is no viable method of assigning probabilities to future random events. (p. 477)
- Weighted average cost of capital (WACC).** The expected average cost of funds. It is found by weighting the cost of each capital component by its proportion to the total capital structure. (p. 475)

QUESTIONS

1. What is the objective of capital budgeting?
2. Define the *time value of money*.
3. How is NPV calculated? What is the decision rule for NPV? How is the internal rate of return calculated? What is the decision rule for IRR?
4. Under which circumstances can the NPV and IRR calculations lead to conflicting results? What is the major reason for the difference? Which of the two methods is preferable? Why?

5. What are the major types of cash flows to be included in a capital budgeting analysis? Describe each.
6. Why is depreciation important in the analysis of a capital budgeting proposal?
7. You are analyzing a potential capital investment project involving a new product. As you are compiling the relevant data for your analysis, you are informed of an expenditure that occurred during the prior year. The marketing research department of your company conducted an assessment of the demand for this new product. The cost of this research was \$100,000 and was part of the company's expenses during the previous year. Is this relevant to your present analysis? Why or why not?
8. How is the company's optimal capital budget determined? Does the decision-making process in this case resemble the procedure used in determining the price and quantity of output? How?
9. How is the weighted average cost of capital determined?
10. Define *beta*. How is it used to compute the required rate of return on a company's stock (the equity cost of capital)?
11. Why is capital rationing not considered to be rational maximizing behavior?
12. Enumerate causes of business risk.
13. You are comparing two potential mutually exclusive investment projects. You have calculated the expected NPV of project A to be \$3,758 and that of project B to be \$3,114. Can you be certain that you should recommend to your managers that they should implement project A?
14. Would a risk-averse person prefer a project whose distribution of potential outcomes can be drawn as a continuous normal curve with a very high peak and a steep decline around the peak?
15. Define *coefficient of variation*.
16. "All our projects are discounted at the same interest rate," says the treasurer of a large company. Would you dispute the advisability of such a procedure?
17. Why do companies use the RADR method much more frequently than the certainty equivalent method?
18. Describe and give examples of
 - a. Sensitivity analysis
 - b. Simulation analysis
19. Why does the use of "real options" improve capital investment analysis? Under which circumstances would you recommend using real options?

PROBLEMS

1. Jay Wechsler agrees to purchase a car from a local dealer, the Con Car Co. The purchase price is \$15,000. Jay has the cash to pay the entire amount and wants to do so. Con's sales manager uses the following argument to convince him to finance the car: "All we require is a down payment of \$3,000. Then you can borrow the \$12,000 from our finance company at 12 percent. You will make monthly payments of \$266.93 for 5 years (60 months), a total of \$16,015.80. If you do that, you get to keep your \$12,000. Now suppose you keep this money in a money market account that pays you 8 percent compounded quarterly. In 5 years, the \$12,000 will grow to \$17,831.40. That means that you will be better off by \$1,815.60 than if you pay the \$12,000 in cash."

Assume all the numbers are correct. Does the offer sound too good to be true? Why? (This is an argument often used by automobile dealers. One of the authors encountered it when buying a car.)
2. Your firm has an opportunity to make an investment of \$50,000. Its cost of capital is 12 percent. It expects after-tax cash flows (including the tax shield from depreciation) for the next 5 years to be as follows:

Year 1	\$10,000
Year 2	20,000
Year 3	30,000
Year 4	20,000
Year 5	5,000

- Calculate the NPV.
 - Calculate the IRR (to the nearest percent).
 - Would you accept this project?
3. You own a large collection of fine wines. You now decide that you must liquidate this valuable asset. However, you predict that the value of your collection will rise in the next few years. The following are your estimates:

Year	Estimated Value
Today	\$70,000
1	88,000
2	104,000
3	119,000
4	132,000
5	142,000
6	150,000

If you assume your cost of capital to be 10 percent, when should you sell your collection to maximize your NPV?

- The Glendale Construction Company is considering the purchase of a new crane. Its cost would be \$500,000. If it were to make the purchase, the company would sell an old crane, which still has a book value of \$100,000 and which it could probably sell in the second-hand market for \$70,000. If the tax rate is 40 percent, what would be the actual cash investment in the new crane?
- As capital investment analyst for the Parkhurst Printing Corporation, you have been asked to evaluate the advisability of purchasing a new printing press to accommodate projected increases in demand. This new machine is expected to last 5 years, and you will be calculating the cash flows of the project for that period.

The purchase price of the press is expected to be \$140,000; in addition, it will cost \$10,000 to install it. The press will be depreciated on a straight-line basis over 5 years to a zero salvage value. However, it is expected to have a market value of \$10,000 at the end of 5 years.

The press is expected to generate the following cash revenues and cash costs and expenses:

	Year 1	Year 2	Year 3	Year 4	Year 5
Cash revenue	\$50,000	\$80,000	\$80,000	\$80,000	\$40,000
Cash cost and expense	25,000	40,000	40,000	40,000	20,000

Because of increased production, additional working capital of \$15,000 will be needed at $t = 0$ (today) and will be returned at the end of the project (5 years from now). The income tax rate is 40 percent, and the company's cost of capital is 12 percent. Calculate the net present value. Should the press be purchased?

6. The Thunderbird Quick Delivery Company needs a truck and is evaluating two alternatives for meeting its needs for the next 10 years.

Choice A: Buy a used Ford truck for \$6,000. It will be depreciated straight-line over a 5-year life to a salvage value of zero. However, it is expected that the truck will be sold after 4 years for \$800. At this point, another used Ford truck will be bought for \$12,000. It will last 6 years and will have no market value at the end of this period. This truck will also be depreciated over 5 years on a straight-line basis.

Choice B: Buy a new Chevrolet truck for \$16,000. The truck will last the entire 10 years and will have a market value of \$1,000 at the end. Again, 5-year straight-line depreciation will be used.

Other information: Income tax rate 34%
 Cost of capital 12%

Which of the two choices is preferable?

7. The Colgate Distributing Company has the choice of furnishing its sales representatives with a car or paying a mileage allowance for the use of the representatives' own cars. If the company furnishes the car, it will pay all expenses connected with it, including gasoline, for business mileage. The estimates are as follows:

Cost of car: \$15,000
Estimated life: 4 years
Depreciation method: Straight-line over 4 years (assuming no salvage value)
Expected sales value of car at end of 4 years: \$2,500
Estimated annual operating costs:

Gasoline	\$900
License and insurance	600
Garage	300
Maintenance	
Year 1	250
Year 2	350
Year 3	450
Year 4	600

If the sales representatives use their own cars, the company will reimburse them at 35 cents per mile; the company estimates that each representative will drive 18,000 miles per year for business purposes. The company's cost of capital is 10 percent, and its income tax rate is 40 percent.

Should the company buy cars for its sales representatives or pay them a mileage allowance? Use the NPV method in your calculation.

8. A company's common stock is currently selling at \$40 per share. Its most recent dividend was \$1.60, and the financial community expects that its dividend will grow at 10 percent per year in the foreseeable future. What is the company's equity cost of retained earnings? If the company sells new common stock to finance new projects and must pay \$2 per share in flotation costs, what is the cost of equity?
9. A company has 1,000,000 shares of common stock outstanding, and the current market price is \$50 per share. The company has also issued 20,000 bonds (\$1,000 maturity value each), which are presently selling in the market at \$980 each. The bonds are selling at a yield of 11 percent; the company expects to pay a dividend of \$3 per share in the coming year, and the dividend is expected to grow at 8 percent per year. The company is in the 40 percent tax bracket. What is the weighted average cost of capital?

10. A company has a beta of 1.3. The risk-free interest rate today is 8 percent, and the return on a market portfolio of stocks is 14 percent. (Therefore, the market risk premium is 6 percent, the difference between the market return and the risk-free return.)
 - a. What is the required return (equity cost) on the company's stock?
 - b. If the risk-free rate rises to 9 percent, what will be the required rate of return on the company's stock?
 - c. If the beta of this company were 0.8, what would be its required rate of return?
11. Two mutually exclusive alternatives, projects C and D, have the following investments and cash flows:

	Project C	Project D
Investment at period $t = 0$	\$40,000	\$40,000
Cash inflow at $t = 1$	10,000	20,500
Cash inflow at $t = 2$	10,000	20,500
Cash inflow at $t = 3$	47,000	20,500

- a. Calculate the NPV and IRR of each project. The company's cost of capital is 12 percent.
 - b. Which of the two projects would you accept? Explain.
 - c. Sketch the two projects' NPV profiles.
12. The Berkshire Resort Hotel has planned several improvement projects. However, it has decided to restrict its capital expenditures to \$340,000 during the next year. The following are the projects it has on its drawing board:

	Original Investment	Net Present Value
Additional tennis court	\$ 20,000	\$ 5,500
Kitchen renovation	50,000	14,000
New children's playground	60,000	12,500
New bungalows	100,000	22,500
New golf clubhouse	120,000	32,500
Olympic-size swimming pool	140,000	45,000
New theater arena	150,000	40,000

Which projects should it undertake?

13. The Quality Office Furniture Company has compiled the year's revenue expectations and their probabilities:

Sales (\$000)	Probabilities
240	0.05
280	0.10
320	0.70
360	0.10
400	0.05

Calculate

- a. The expected revenue
- b. The standard deviation
- c. The coefficient of variation

14. The Learned Book Company has a choice of publishing one of two books on the subject of Greek mythology. It expects the sales period for each to be extremely short, and it estimates profit probabilities as follows:

Book A		Book B	
Probability	Profit	Probability	Profit
0.2	\$2,000	0.1	\$1,500
0.3	2,300	0.4	1,700
0.3	2,600	0.4	1,900
0.2	2,900	0.1	2,100

Calculate the expected profit, standard deviation, and coefficient of variation for each book. If you were asked which of the two to publish, what would be your advice?

15. The Cactus Corporation is considering a 2-year project, project A, involving an initial investment of \$600 and the following cash inflows and probabilities:

Year 1		Year 2	
Probability	Cash Flow	Probability	Cash Flow
0.1	\$700	0.2	\$600
0.4	600	0.3	500
0.4	500	0.3	400
0.1	400	0.2	300

- Calculate the project's expected NPV and standard deviation, assuming the discount rate to be 8 percent.
 - The company is also considering another 2-year project, project B, which has an expected NPV of \$320 and a standard deviation of \$125. Projects A and B are mutually exclusive. Which of the two projects would you prefer? Explain.
16. The Grand Design Corporation uses the certainty equivalent approach in making capital budgeting decisions. You are given the following data for a particular project:

Year	Cash Flow	Certainty Equivalent Factor
0	\$-20,000	1.00
1	5,000	0.90
2	5,000	0.90
3	5,000	0.90
4	15,000	0.70

The risk-free discount rate is 4 percent, and the RADR is 12 percent. Calculate the NPV. Would you accept this project?

- Project A has an expected NPV of \$500 and a standard deviation of \$125. Project B has a standard deviation of \$100 and an expected NPV of \$300. Which of the two projects would you select? Explain why.
- Global Industries has calculated the return on assets (ROA) for one of its projects using the simulation method. By simulating the operations 1,000 times, they obtained an ROA of 16.7 percent and a standard deviation of 6.2. The results of the simulation conform quite closely to a normal curve.

- a. Draw a probability distribution using the given data.
 - b. The company's objective is to achieve an ROA of 12 percent. What is the probability that the project will achieve at least that level?
 - c. What is the probability of ROA being nonnegative?
19. The Great Pine Forest Corporation is analyzing an expansion project with the following information:

<i>Initial investment:</i>	\$120,000
<i>Depreciation life:</i>	5 years—straight-line
<i>Project life:</i>	5 years
<i>Additional working capital at t = 0</i>	\$20,000
<i>Working capital returned at t = 5</i>	\$20,000
<i>Expected salvage value at t = 5</i>	\$15,000
<i>Tax rate:</i>	34 percent
<i>Cost of capital:</i>	12 percent

	Year 1	Year 2	Year 3	Year 4	Year 5
Revenue	\$50,000	\$80,000	\$80,000	\$80,000	\$40,000
Cash costs	30,000	30,000	25,000	25,000	25,000

- a. Calculate the NPV of this project.
 - b. Now conduct a scenario analysis as follows:
 - 1. Assume the best case to have revenue 10 percent higher than just stated, costs 5 percent lower than given, and salvage value twice the amount given.
 - 2. Assume the worst case to have revenue 10 percent lower than given, costs 5 percent higher than given, and salvage value to be 0.

Show the results for the best case, the most likely case, and the worst case.
20. Sam Parkington has a nontransferable option to mine for gold on a certain piece of land. He has three choices of action:
- a. He can start mining immediately.
 - b. He can conduct further tests to see whether there is a good promise of finding gold.
 - c. He can drop the option.
- The cost of the test would be \$45,000, and the cost of mining would be \$150,000. If he finds gold, he expects to net \$600,000.
- He estimates the following probabilities. If he starts mining without further tests, he estimates that the probability of finding gold is 55 percent. He expects that the probability of the test being successful is 60 percent. If the test is favorable, the probability is 85 percent that there is gold in the ground, but if the test is not favorable it is only 10 percent. Using a decision tree, make a recommendation.
21. It has been your secret wish to own and operate an amusement park when you can afford to make the investment. That time has now arrived. There is a large empty lot at the outskirts of Phoenix, Arizona, owned by the city. The city is willing to lease one-half of this lot for 5 years with an option to lease both it and the other half for the next 5 years. In return for charging you a reasonable rent, the city will take ownership of your equipment at the end of the 5 or 10 years.
- You have estimated your original investment to be \$250,000. You expect your net cash flows (after lease payments, all other expenses, and taxes) to be \$55,000 for each of the first 5 years. If you exercise your option to continue with both parts of the property at the end of year 5, you will need to invest another \$150,000 (for additional equipment and a miniature golf course). Because your cash flow estimates are now far into the future, you estimate a

50 percent probability that your annual cash flows will remain the same (\$55,000) and a 50 percent probability that they will rise to \$100,000 per year for the second 5 years. Your cost of capital is 12 percent.

- a. Is the first 5-year project acceptable?
 - b. What is the value of the total project if you exercise your option? What is the value of your option?
22. You have been offered an investment, which, according to your calculations, has the following possible rates of return:

Probability	Rate of Return
0.2	-5%
0.6	+10%
0.2	+25%

You can also use your funds to buy a U.S. Treasury note, which will earn a certain 3.5 percent. In which of these two alternatives will you invest your money?

The Value of a Corporation



We have just learned how to calculate the present value of a project given cash inflows and outflows over a period of time, and discounting these to the present using a discount rate, the cost of capital.

A similar method of computation can be employed to value a corporation as a whole. Of course, if the corporation's stock trades on a stock exchange and its debt is composed of publicly owned bonds, its market value can easily be established from published data. However, assume you are one of the founders of a relatively small private corporation. You and the other owners are now planning to carry out an "initial public offering" (usually referred to as an IPO). How would you go about estimating what your company is worth?³³ To determine some kind of a reasonable value figure, you will have to estimate your cash inflows and outflows several years into the future. To do a complete valuation job, it would be necessary to project your income statements and balance sheets. We try to condense our procedure here by selecting the essential data we need to make our computation without constructing complete financial statements.

To accomplish our objective we must start with the concept of **free cash flow**.³⁴ Free cash flow consists of funds that would be available to investors from operating cash flows after we subtract investments in working capital and fixed assets needed to operate and grow the company. These terms are defined as follows:

1. Operating cash flow = $EBIT \times (1 - T) + \text{Depreciation}$

where EBIT = Earnings before interest and tax

(also frequently called operating profit)

T = Tax rate

2. New investment in operating working capital represents the annual increases in certain parts of current assets less increases in accounts payable.
3. Increases in fixed assets are increases in land, property, and equipment before subtracting depreciation (i.e., the gross increases in fixed assets).

Let us construct a simple example to illustrate the valuation method. Assume in 2012, the following were the relevant numbers (in \$000):

EBIT	\$ 60
Depreciation	20
Operating working capital	80
Gross fixed assets	120

Your company expects to have rather rapid growth during the next 3 years because of a new product you are bringing to the market. You expect, however, that after 3 years your competition will have caught up with you. Because you are not certain what new products you will introduce then, you make the conservative judgment that your company's free cash flow will increase at 6 percent per year for the foreseeable future. To calculate the free cash flow for the years 2013, 2014, and 2015, you have estimated the following:

	2013	2014	2015
EBIT	\$ 90	\$110	\$125
Depreciation	30	35	40
Operating working capital	110	135	155
Gross fixed assets	160	190	220

EBIT and depreciation give us the two components of free cash flow. To calculate the other two, we must compute the year-to-year changes to obtain the cash outflow arising from an increase in those investments. Thus,

³³The reader should notice that in calculating the value of this company we use the nonconstant growth model, discussed in note 15 of this chapter.

³⁴Free cash flow was mentioned briefly in Chapter 2. Free cash flow is essentially the same as dividends, which were used to calculate the value of the common stock in that chapter. The following discussion has benefited from Brigham, Gapenski, and Ehrhardt, *Financial Management*, Chapter 9.

	2013	2014	2015
Operating working capital	\$30	\$25	\$20
Gross fixed assets	40	30	30

We expect the income tax rate to be 40 percent in each of the years. Now we are ready to calculate the free cash flows for the 3 years:³⁵

	2013	2014	2015
EBIT $\times (1 - T)$	\$54	\$66	\$75
+ Depreciation	30	35	40
- Increase in operating working capital	30	25	20
- Increase in gross fixed assets	<u>40</u>	<u>30</u>	<u>30</u>
Free cash flow	\$14	\$46	\$65

The next step is to calculate the present value of the free cash flows. At this point we must establish the discount rate—the cost of capital. We estimate it to be 12 percent. Therefore, the present value of the above three cash flows at the end of 2012 will be the following (again using data from Appendix Table A.1c):

Year 2013	14×0.8929	\$12.5
2014	46×0.7972	36.7
2015	65×0.7118	46.3
Total		<u>\$95.5</u>

Now we must account for the cash flows after the first 3 high-growth years. We had previously assumed we would expect the normal long-term growth rate in free cash flow to be 6 percent. We can use the constant growth formula (Gordon model) to give us the answer. If free cash flow grows by 6 percent in 2016, then it will amount to $65 \times 1.06 = \$68.9$. Because the cost of capital is 12 percent and

the growth rate is 6 percent, our calculation will proceed as follows:

$$68.9 / (0.12 - 0.06) = 68.9 / (0.06) = 1,148.3$$

The answer is the present value of all future cash flows at the beginning of year 2016 (or the end of 2015). To bring this amount back to the end of 2012, we must discount the amount for 3 years, or $1148.3 \times 0.7118 = 817.4$. The last step is to add the present value of the first 3 years to the present value of the growing perpetuity: $95.5 + 817.4 = 912.9$.

We have thus estimated the total value of your business to be \$912,900. But as one of the owners of the business, you are interested in how much you could receive for your stock. We must now deduct any bank debt and long-term debt that you have on your balance sheet. Suppose you had borrowed \$300,000. Then the estimate of your stock value is \$612,900. Let us assume you and the other present owners want to retain 20 percent of the stock ownership of the firm. This means you could expect to obtain \$490,320 from your IPO.

Some cautions are in order. The result that we calculated is extremely sensitive to all the estimates we had made for cash flows, tax rates, and discount rates. The final results could differ substantially from our calculations. It would most likely be useful to do one or more additional calculations, using more optimistic and pessimistic assumptions, to help establish a range of potential outcomes.

You would perform a similar calculation if instead of an IPO you were to consider a buyout offer by a large established firm in your industry. You would then be interested in determining the price at which you would be willing to sell your company. Again it would be useful to make several calculations to arrive at a desirable selling price. Further, if the acquiring firm should be able to achieve some synergies from taking over your firm, then the value of your company could be higher. That firm certainly made calculations similar to yours, and because of the synergies, they may have arrived at a higher figure. Thus, after you have bargained with the large firm, you may end up with a larger amount. If you were lucky enough to have two large firms competing to acquire you, you may indeed strike a good bargain.

³⁵It is quite common for a new, small, and growing company to have negative free cash flows in its early years. Even though its EBIT is positive, it will have substantial cash outflows as it builds up its working capital and fixed asset base. To simplify matters, we have assumed positive free cash flows in each year.

The Multinational Corporation in a Global Setting

Learning Objectives

Upon completion of this chapter, readers should be able to:

- Understand how the factors affecting supply and demand are affected in different countries around the world.
- Define the exchange rate and identify several methods of hedging.
- Understand multinational capital budgeting and explain how it differs from capital budgeting of a domestic corporation.
- Show how changing transfer prices can benefit a corporation.

The Situation

Global Foods has ceased to be a strictly domestic corporation. It has opened production facilities that produce both its soft drinks and its bottled water in several countries in western Europe and the Pacific Rim. It is now ready to make an investment in one of the countries that was formerly part of the Soviet sphere. Management has been investigating in which of the countries to locate. It turned out that it favored the Czech Republic. This country is centrally located, has a skilled labor force, has been very active in privatizing industry, and has a fairly stable government. Its labor costs are relatively low.

After an investigation of various locations within the Czech Republic, the company found a plant not presently in use but rather well maintained just outside the town of Pelhřimov, in southeastern Bohemia, not too far from the border of Moravia. It is about 115 kilometers from Prague, the capital city. The town itself has about 17,000 inhabitants. Pelhřimov has some light industry including wood processing, food, textile, and clothing.

George Kline, manager of Global Foods' capital planning (we met George in the previous chapter), has been given the task of preparing a capital budgeting evaluation. He consults with the marketing and manufacturing departments. He discusses the Czech economic situation with the company's economist. He obtains estimates of Czech and U.S. inflation rates, and the exchange rate of the Koruna (KCZ) against the U.S. dollar. He further acquaints himself with the corporate tax rate in the Czech Republic. Armed with this information, he begins to put down the assumptions with which he will work.

He expects that the fixed investment by Global Foods, U.S.A., will be \$4 million, and that another \$400,000 will be invested in working capital. The fixed investment will be depreciated straight-line over 8 years. Because the plant is in very good condition, only the installation of machinery will be needed for the plant to begin operating rather quickly.

To be very conservative, cash flows for only 4 years will be estimated, with a terminal value estimated at the end of the 4 years. His marketing people think that first year's sales will be 400 million KCZ and that the volume will

(continued)

(continued)

increase by 6 percent each year. Production costs in the first year will be 190 million KCZ. The parent company will be providing various supplies and services to the subsidiary starting at 110 million KCZ and will grow at the same rate as volume. A license fee of 3 percent of sales will be paid by the Czech subsidiary to the parent corporation. General and administrative expenses will be 12 percent of sales revenue. The following are estimates of annual price level changes in the United States and the Czech Republic:

General price level in the United States	2.0% ¹
General price level in the Czech Republic	3.5%
Sales price in Czech Republic	4.0%
Cost of production in Czech Republic	4.0%

The exchange rate was 27 KCZ/\$1. The exchange rate is expected to change in relation to the inflation rates in the two countries.² The income tax rate in the Czech Republic is 31 percent, whereas it is 35 percent in the United States. After the initial infusion of working capital into the project, the subsidiary will require that working capital be 10 percent of sales. Fifty percent of this additional requirement will be financed internally (profits and increase in current liabilities), and the other 50 percent will be financed externally by the subsidiary.

Global Foods' cost of capital for projects of average risk is 15 percent. However, to account for the additional risk faced by operating in a foreign country that just recently embraced free trade (although still regulated), a 4 percent risk margin will be added so the discount rate for the parent company will be 19 percent. The cost of capital for similar enterprises in the Czech Republic is 22 percent.

The profit calculated on the parent's shipments of supplies to its subsidiary will be computed at 5 percent. This appears reasonable and should not be questioned by U.S. or Czech tax authorities.

The subsidiary will remit 50 percent of its aftertax profits to the parent.

With all this information, George will now input the data into an Excel spreadsheet that he specifically prepared for this project. He will have to organize his work into five separate exhibits. First, he will calculate the profit of the subsidiary. Second, he must compute the additional working capital investment made by the subsidiary. The third exhibit will be a cash flow statement for the subsidiary with a calculation of the NPV and IRR. The fourth exhibit may appear somewhat complicated. It translates the dividend by the subsidiary to the parent from KCZ into U.S. dollars; this involves a process called "grossing up" to ascertain the tax for which the parent may be liable because the Czech tax rate is lower than that in the United States.³ Last, the parent company must compute the NPV and IRR of all cash flows received from the subsidiary.

INTRODUCTION

Most of our discussion in this book does not differentiate between domestic and international businesses. We touch briefly on aspects of international business in the "Global Application" sections that are found at the end of most chapters. However, the opportunities and problems faced by American corporations in foreign countries occupy many recent headlines and articles in the business media. Therefore, we devote a chapter to this increasingly important topic—the multinational corporation.⁴

A multinational corporation (usually abbreviated as MNC or MNE for multinational enterprise) is faced with the same opportunities and problems as the domestic corporation. It must concern itself with the forces of supply and demand. It must

¹The price of supplies and services from the United States will also increase at 2 percent.

²Because inflation in the United States is 2 percent and in the Czech Republic 3.5 percent, the KCZ will depreciate each year by 1.5 percent.

³The process of grossing up is explained later in this chapter.

⁴Some of the topics discussed in this chapter may have been mentioned in some of the previous chapters. However, this chapter tries to tie the various strings together into a unified whole. This chapter benefited a great deal from Michael H. Moffett, Arthur I. Stonehill, and David K. Eiteman, *Fundamentals of Multinational Finance*, 2nd ed. Boston: Pearson Addison Wesley, 2006.

consider demand elasticity for its products and seek to achieve profit optimization by applying the marginal principle, given business and financial risks. All these concerns are discussed thoroughly in this text. One of the main reasons why a company wants to operate globally is to take advantage of new growth opportunities in the rapidly expanding countries such as China, India, and Brazil and in the emerging countries such as Vietnam, Indonesia, and Turkey. But in doing so, such companies also face new challenges and risks. These risks involve factors such as exchange rate fluctuations; different tax systems and government regulations; tariffs and other restrictions; and, of course, different languages, cultures, and ways of doing business. In this chapter, we present an overview of these opportunities and risks, starting first with the opportunities.

OPPORTUNITIES FOR MULTINATIONAL CORPORATIONS IN A GLOBAL ECONOMY

Expanding the Dressings Category Throughout the World

In Chapters 4 and 5, we talked about the key factors that affect the demand for a product and the challenges of estimating and forecasting this demand. For companies that operate in a global environment, it is also critical to understand how and why the various factors that influence demand may differ among the countries in which they sell their products. To help illustrate this point, we asked Lisa Vortsman, product manager of a real “Global Foods Company,” to share some of her experiences.⁵ This section is based on what she told us about her efforts to increase the sales of the dressings category in the top ten countries of the world based on per capita consumption of dressings. These countries include the United States, the United Kingdom, France, Brazil, and Russia.

In the food business, the dressings category consists primarily of mayonnaise, ketchup, mustard, and salad dressings.⁶ All food companies find it very challenging to grow revenue and profit in this category. Strong price competition exists in all of the top ten countries. Unit costs are increasing due to the rising price of key ingredients such as soybean oil. But the key factors that Lisa and her team knew they had to address relate to the nature of the market demand for dressings in the top ten countries. More specifically, she provided the following information:

1. **Household penetration is already very high:** The dressings category is mature and developed. For example, there is a 75 percent household penetration in western Europe and a 98 percent penetration rate in Russia.
2. **Most consumers use dressings on limited occasions:** Many consumers use condiments such as mayonnaise, ketchup, and salad dressings on a limited basis because they tie it to only a few dishes. In large dressings markets like the United Kingdom and the United States, mayonnaise and ketchup are perceived as condiments that cannot be eaten with too many dishes. For example, ketchup is used primarily with hamburgers and fries. Consumer habits are difficult to break and eating habits are no exception.
3. **Health is a significant concern:** Consumer perception of the category tends to be negative across the globe. For example, because of its high fat and calorie content, mayonnaise is considered to be unhealthy in top markets such as the United Kingdom, the United States, and Brazil. Many Russian consumers are concerned about using mayonnaise because they believe it is made with artificial ingredients; they associate ketchup with junk food.

⁵We are grateful to Lisa Vortsman for her contribution to this chapter. She is currently the product manager in a major global food company. This section is based on her actual experience as the product manager of the dressings category. However, she has requested that we do not publish the name of her company.

⁶For a good overview of the sauces and dressings industry in the United States, see <http://www.preparedfoods.com/articles/print/110990-cooking-sauces-and-dressings>.

Lisa and her team realized that, to grow this category's revenue and profits, they would have to overcome these barriers. Could they get the 25 percent of the households in western Europe to at least try their company's condiments? Could they get those who use these dressings to try it with different foods or in different recipes? Could they get people who have negative views of dressings because of health concerns to see that these dressings actually contained ingredients that were not so unhealthy and in fact were "good for you"? This was their challenge.

Before taking any action, Lisa commissioned a Triggers and Barriers study to find out as much as possible about the current tastes and preferences of the consumers of dressings in the world's top ten markets. What really would get consumers to use more dressings? What actually prevents them from buying it or using it more often? These were the kinds of questions that the study sought to answer.

The Triggers and Barriers study identified the top requirements that consumers have when using dressings products. As suspected, the study revealed that in the United States, Brazil, and the United Kingdom, consumers looked for dressings that contained healthier ingredients, such as certain types of oils. They also looked for mayonnaise that contained fewer calories and fats. In France and Russia, consumers wanted ketchup that contained less sugar. In the United States, the United Kingdom, Brazil, and Russia, consumers looked for dressings that provided them with cooking solutions. The lack of healthier ingredients was the main barrier that kept consumers from buying dressings or from using them more frequently.

Based on the study's findings, Lisa decided that an effective way to overcome these barriers was to educate the consumers on the actual content of her company's dressings. For example, consumers in Brazil thought that a spoonful of mayonnaise contained anywhere between 200 and 600 calories when, in reality, it has only 40 calories. The marketing message in Brazil made sure that this fact was made clear to consumers. The demand by American consumers for foods that contain omega-3 oils has been steadily increasing. An important source of omega-3 is soybean oil, which also happens to be a major ingredient of mayonnaise. Hence, the company's marketing message in the United States was that mayonnaise is a good source of omega-3 oils. In the United Kingdom, consumers were particularly keen on using dressings with less saturated fat. Therefore the message in this market stressed the lower saturated fat content of the company's line of dressings. In the United States, the United Kingdom, and Brazil, the communication messages focused on stressing that mayonnaise is made with real ingredients such as eggs, oil, and vinegar. In France (as well as Argentina and Colombia), the communication message highlighted the amount of tomatoes that are in each ketchup bottle. The main objective of the messages in all of these countries was to tell consumers about the benefits of using products in the dressings category.

The communication's focus on positive (or less negative) health factors effectively dealt with the barriers to buying and using food dressings. However, two challenges still remained: Dressings is a mature category, and, in many countries, the use of the major products such as ketchup and mayonnaise is limited to being an accompaniment to only certain foods or dishes.

To overcome the barrier of lack of versatility and, even better, to offer a trigger to buying and using dressings, Lisa and her team came up with a marketing strategy based on the theme of "mom's home cooking." They conducted a study in the top ten countries to learn about the following:

1. What are mom's favorite dishes to prepare?
2. Which of these dishes use any of the main products in the dressings category (i.e., ketchup, mayonnaise, and mustard)?
3. How could these dishes be improved by using one or more of the products from the dressings category in their recipe?

Once they got their answers, the team then launched a marketing campaign that used the theme of “make something new today.” They also offered suggestions for recipes whose ingredients included one or more of the products. They based these recipes on favorite recipes of mothers around the world, keeping in mind the differences in cooking preferences and behaviors in different countries. For example, Russia has the highest consumption per capita of mayonnaise. Russians use it on almost everything, including salads, chicken, soup, and potatoes. French consumers use a lot of dressings, but they like to make their own from scratch whenever possible. The British like to keep things simple. So if they are given a “secret ingredient” that will add a new taste to their everyday dishes, they tend to try it.

To summarize this case in relation to Chapters 3, 4 and 5, we can say that Lisa’s studies of the demand for dressings in the top ten country markets led to a marketing communications strategy that helped to increase demand for her company’s dressing in each of these countries without their having to cut prices. In effect, we can say that all of their efforts caused a rightward shift in the demand for food dressings. To continue this section’s theme of opportunities for multinational countries, we now discuss the challenges of doing business in India, the country with the world’s second largest population: 1.2 billion people.

DOING BUSINESS IN INDIA⁷

Brief Introduction to Some of India’s Key Industries

Back in 2001, Jim O’Neill, chief economist for Goldman Sachs, believed that the combined economies of four countries—Brazil, Russia, India, and China—would eclipse the total output of the six largest western economic powers in about forty years.⁸ Interestingly enough, at the time, O’Neill had visited China only. In hindsight, not only did he have a keen sense of global economic trends, he showed a streak of marketing genius by referring to this group of four as the BRIC countries. Countless articles and books have been written about the BRICs. Our aim in this section is to give readers a brief idea of what doing business in the BRIC countries is like by using India as an example. We start first with a general overview of the Indian economy and some of its major industries. We then conclude with the challenges that McDonald’s faced when it chose to operate in India. In terms of managerial economics, McDonald’s challenges and eventual success in its India operations clearly indicate the importance of understanding the differences in consumer tastes and preferences, incomes, and government regulation (three key factors affecting demand or supply) in different countries of the world.

India is on the Asian subcontinent. Its land mass is roughly one-third the size of the United States, and it represents 2.4 percent of the world’s total. However its population of approximately 1.2 billion people is about four times that of the United States, and it represents 17.5 percent of the world’s population. Thus, India is the world’s second most populous country (after China, of course), but it is the most populous democracy in the world.⁹ India has a legacy of long-term colonial stagnation and economic backwardness. But in 1947, after India regained independence from Britain, it began its journey toward self-reliance.¹⁰ The aspiration of India’s leaders during

⁷This section was prepared by Dr. Navin Punjabi, Professor of Business at H.R. College of Business and Economics, Mumbai, India. We are grateful for his assistance.

⁸<http://www.ft.com/intl/cms/s/0/112ca932-00ab-11df-ae8d-00144feabdc0.html#axzz23u2w1vFn>.

⁹<http://www.india.gov.in/knowindia/profile.php> accessed on July 22, 2012.

¹⁰Kapoor, A. (2006). Part One Overviews: 1.3. An Economic Overview of India. In *Doing Business with India* (pp. 17–23). Blue Ibex Ltd.

independence was to turn India into a vibrant, self-reliant national economy and to do away with the abject poverty so prevalent among its people.

It appears that India is on its way to becoming economically developed. From 2000 until mid-2012, the annual growth rate of India's gross domestic product (GDP) averaged 7.4 percent.¹¹ This is more than twice the long-term growth rate of 3.3 percent for the United States.¹² As of this writing, the Indian economy has slowed somewhat from its rapid growth over the past decade. There are concerns about whether this is a long-term trend or just a cyclical dip. However, the optimistic view is that India's economy will resume its fast growth.¹³

In 2011, agriculture and related activities accounted for 14.7 percent of the GDP. It also made up about 58 percent of the nation's employment. The industrial sector, including construction, made up 28 percent of the GDP. Within this sector, manufacturing accounted for about 15 percent of the total GDP. Note that this share was considerably less than China, where manufacturing accounted for over 40 percent, and other East Asian countries, where manufacturing's share of GDP was over 30 percent.¹⁴

China is sometimes called the world's factory, while India is sometimes called the world's back office. And the economic data clearly supports these catchphrases. The service sector share in India's GDP is close to 57 percent. As expected, the information technology (IT) sector dominates the service sector in India. IT and business processes outsourcing (BPO) generates employment for close to 2.8 million persons and indirect employment for another 8.9 million people. The share of IT in India's exports was close to 62 percent in 2011.¹⁵

Although India is known for its services sector, it has also become a growing presence in the global pharmaceutical industry. For almost four decades prior to 2005, Indian pharmaceutical companies (also known as pharmas) manufactured foreign patent-protected drugs that they had reverse-engineered and sold their copies in the Indian home market. They were able to do this because the Indian government did not abide by international patent laws, and it protected these companies by putting very high tariffs on foreign patent products. But in 2005, the government agreed to adhere to international patent laws. Under the new patent regime, Indian pharmaceutical companies could no longer make products protected by foreign patents, nor could they expect tariff protection. At first, these actions obviously hurt their revenue and profits, but Indian pharmaceutical companies learned to adapt quickly to the new competitive environment. They began manufacturing generic versions of the foreign products whose patents had expired. They then started exporting these generics as well as selling in the domestic market.¹⁶ They have also begun to invest in the research and development required to invent their own patented products. Some industry analysts estimate that India pharmas have the potential to capture up to 10 percent of the global market for patented drugs.¹⁷

The Indian retail industry contributes 14 percent to India's GDP and employs 7 percent of India's population. The retail industry comprises organized and unorganized stores. Organized retailing refers to trading activities undertaken by licensed

¹¹<http://www.tradingeconomics.com/india/gdp-growth>.

¹²<http://www.tradingeconomics.com/united-states/gdp-growth>.

¹³Punj Shweta, (March 16, 2012), Budget 2012: Economic Survey exudes hope, optimism and caution, *Business Today*. <http://businesstoday.intoday.in/story/budget-2012-economic-survey-union-budget-2012-13/1/23243.html>, accessed on July 23, 2012.

¹⁴Government of India (2012) Economic Survey of India 2012-13 retrieved from <http://www.indiabudget.nic.in/es2011-12/echap-09.pdf> July 23, 2012.

¹⁵Government of India (2012) Economic Survey of India 2012-13 retrieved from <http://www.indiabudget.nic.in/es2011-12/echap-10.pdf>.

¹⁶http://www.usitc.gov/publications/332/working_papers/EC200705A.pdf.

¹⁷Ernst and Young (2011), *Doing Business in India*, retrieved from <http://www.ey.com/IN/en/Services/Tax/Tax—Budget-Plus-2011>.

retailers who are registered for sales and income taxes. These retailers are typically incorporated or privately owned hypermarkets and retail chains. Unorganized retailers engage in the more traditional forms of buying and selling. One of the most important types in this retail category is the Kirana: the small, owner-operated neighborhood shop similar to what Americans call a mom and pop store. Another common type of unorganized retailer is the street vendor. Unorganized retailers account for more than 95 percent of total retail sales in India. As India continues to develop, the growth potential for organized retailing is tremendous. According to the AT Kearney Global Retail Development Index (GRDI),¹⁸ India ranks first among thirty emerging market economies in retail attractiveness.

The Indian government currently allows only single-brand retail chains to operate in India. Thus, companies like Nike, the Gap, or Abercrombie and Fitch can own and operate retail shops in India, but Wal-Mart and Carrefour are prohibited from doing business in India. The situation may change, but as of 2012, these global chains can only operate in India as 50 percent owners of wholesale operations.

The Business and Investment Climate

The World Bank and the International Finance Corporation (IFC) rank countries in order of ease of doing business. The ranking is based on factors such as starting a business, obtaining construction permits, getting electric power, registering property, obtaining credit, protecting investors, paying taxes, trading across borders, enforcing contracts, and resolving insolvency. India ranks 132nd among the 183 countries in the World Bank and IFC report. A higher rank indicates an environment more conducive to business. Singapore ranks number one and the African country of Chad ranks 183rd.¹⁹

Despite the relative difficulty of doing business in India, the country's huge growth potential continues to attract foreign investment. According to a 2012 report prepared by the United Nations Conference on Trade and Development (UNCTAD), India attracted \$31.6 billion in foreign direct investment in 2011. This makes India the third most favored destination for foreign direct investment after the United States and China.²⁰

Socioeconomic and Political Factors

Undoubtedly the most important socioeconomic problem facing India today is its pervasive and abject poverty. Roughly one-third of its 1.2 billion people live below the international poverty line, which is the equivalent of about \$1.25 per day. Almost 70 percent live on about \$2.00 per day.²¹ Unfortunately, this degree of poverty also characterizes the other three BRICs: Brazil, Russia, and China. Over the past decade, however, there has been rapid growth and development in India and the other three BRICs, accompanied by a rising middle class. This offers businesses a market of hundreds of millions of households with more discretionary income to spend on consumer durables and nondurables.

India is culturally rich and very diverse. India's national language is Hindi, but it also has twenty-one other official languages. Its four main religions are Hinduism,

¹⁸<http://www.atkearney.com/index.php/Publications/global-retail-development-index.html>.

¹⁹World Bank and IFCI, 2012, *Doing Business in a More Transparent World*, <http://www.doingbusiness.org/~media/fpdkm/doing%20business/documents/annual-reports/english/db12-fullreport.pdf>, retrieved on July 23, 2012.

²⁰E. T. Bureau, July 9, 2012, "India seen as an attractive destination for foreign direct investment" http://articles.economicstimes.indiatimes.com/2012-07-09/news/32605031_1_fdi-flows-foreign-direct-investment-world-investment-report, retrieved on July 23, 2012.

²¹An article in Wikipedia provides a plethora of government data detailing the extent and seriousness of this problem. See http://en.wikipedia.org/wiki/Poverty_in_India.

Islam, Christianity, and Sikhism, and other religions such as Buddhism, Jainism, Judaism, and Zoroastrianism also have many followers. Because there are so many languages spoken throughout the country, English is used by many Indians as the common language. English is taught at an early age in many schools throughout the country. Of course, this widespread command of English has given the country its competitive advantage in many global services such as call center operations, business process outsourcing, and IT consulting.

India's government is a parliamentary democracy, established in 1950 by the Constitution of India. Its political system has three main branches: executive, legislative, and judiciary. The president of India is the head of the executive branch, followed by the vice president and then the prime minister. But as is the case in all parliamentary forms of government, the prime minister holds the most political power, while the president serves as the titular head of state. The legislative system includes the Rajya Sabha (upper house/Council of States) and the Lok Sabha (lower house/House of Commons). The judiciary consists of the Supreme Court, followed by high courts and district courts.²²

More than 5 percent of India's population is below the age of twenty-five. The educational system in India has produced a very wide range of results. On one hand, about 25 percent of the population is illiterate. On the other hand, India's higher education system is extremely demanding and has produced many highly qualified professionals in all of the major fields, such as science, engineering, and medicine. The institutions of higher education that comprise its world-renowned Indian Institute of Technology (IIT) are more difficult to gain entry to than all of the other top universities around the world, including the U.S. Ivy League colleges and Oxford and Cambridge in the United Kingdom.

In summary, the rapid growth rate in GDP, a rising middle class, a youthful population, large numbers of highly skilled workers and professionals, and widespread use of English all make India a very attractive market in which to invest and operate. However, India still remains a rather difficult country in which to do business because of its lack of business transparency and legal safeguards. To illustrate some of the specific challenges of doing business, let us look at the Indian operations of one of the world's most recognizable global companies: McDonald's.

McDonald's in India

McDonald's is the world's leading food service retailer. It is difficult to stay current with McDonald's statistics. As of 2011, this global food chain had 33,000 restaurants in 118 countries serving more than 67 million customers each day. We found an interesting source that listed ten facts in a highly entertaining way. For example, consider the fact that more people eat at McDonald's per day than the total population of Great Britain!²³ McDonald's business in India is particularly fascinating because one might wonder how a company that is known for selling hamburgers is able to succeed in a country where the majority of its population does not eat beef. About 80 percent of India's 1.2 billion people do not eat beef because the cow is a sacred animal in the Hindu religion. Another 13 percent or so of the country's Muslims do not eat pork. In fact, the majority of India's population is vegetarian.

McDonald's entered the Indian market in 1996 as part of its global expansion strategy.²⁴ Starting with only a few stores at that time, McDonald's India had a network

²²http://en.wikipedia.org/wiki/Indian_government.

²³<http://www.businessinsider.com/amazing-facts-mcdonalds-2010-12?op=1>.

²⁴An interesting history of McDonald's in India is provided in the following PowerPoint presentation by an Indian business professor. See <http://www.scribd.com/doc/2303069/Case-McDonalds-India-Launch>.

of over 250 restaurants across the country serving about 650,000 customers per day.²⁵ In the beginning, McDonald's faced two key challenges in India: (1) how to avoid hurting religious sentiments, and (2) how to avoid political confrontations with the Indian government.²⁶ To help with these problems, McDonald's secured the help of Indian consultants, managers, and licensing partners.

McDonald's then developed a menu especially for India, with vegetarian selections to suit the Indian palate. It also reengineered its operations to address the special requirements of vegetarians. It takes special care to ensure that all vegetable products are prepared separately, using dedicated equipment and utensils. This separation of vegetarian and nonvegetarian food products is maintained throughout the various stages of procurement, cooking, and serving. McDonald's uses only vegetable oil as a cooking medium in India. Even the mayonnaise and condiments are 100 percent vegetarian. Instead of the Big Mac, McDonald's launched the Maharaja Mac (made of two lamb patties, special sauce, cheese, pickles, and onion on a sesame seed bun). As it has done everywhere else in the world, McDonald's has positioned its restaurants in India as a place for families and children. Therefore, Ronald McDonald is a well-known figure to many Indian children, as he is to children around the world.

RISKS FACED BY A MULTINATIONAL CORPORATION

All companies, whether domestic or international, face risk. Risk can be defined as the chance that business results will be different from those that are expected. Because most businesspeople are risk averse, they are mostly concerned with downside risk. A multinational corporation (MNC) must deal with risks faced by domestic industries; however, it is faced with risks that are present only because it transacts business across national borders, that is, **multinational corporation risk**.

One of the important concerns for a corporation is **exchange rate risk** resulting from changes in exchange rates. We discuss these concerns later in this chapter. Here we touch on various other risks faced by the MNC:

Blockage of funds and capital controls: A country may prohibit the repatriation of funds. Such action can lead to a corporation investing its funds less than optimally.

Differences in cultural and religious philosophies: These may create potentially hostile attitudes toward foreign corporations.

Ownership restrictions: Complete ownership of subsidiaries may not be permitted. Thus, the MNC may be restricted to minority participation.

Human resources restrictions: Rules and regulations regarding employment practices may differ from those in the home country. Companies may be required to hire only local workers. In some countries, labor laws and union contracts make hiring and firing more rigid. In some countries, there may be restrictions regarding the hiring and promotion of female employees.

Intellectual property: In some countries, it may be difficult to enforce protection of intellectual property, such as software, textbooks, and films.

Discrimination: Governments may impose special regulations and taxes on foreign corporations, while subsidizing domestic business.

Red tape and corruption: Such problems exist in most countries, in some to a greater extent than in others. As an example, Transparency International publishes a

²⁵Retrieved from <http://www.mcdonaldsindia.com/aboutus.html> on July 23, 2012.

²⁶For examples of some of McDonald's challenges dealing with the Indian government, see Dash Kishore, 2007, "McDonald's in India, Thunderbird, The Garvin School of International Management," <http://www.csumba.org/mba602/McDonald's%20in%20India.pdf>, retrieved on July 23, 2012.

“Corruption Perceptions Index” ranking 183 countries. This index is based on the perception of the degree of corruption as seen by businesspeople, academics, and risk analysts. The scores can range from 0 (most corrupt) to 10 (least corrupt). In the 2011 index, New Zealand ranked the highest (with a score of 9.5), followed by Denmark (9.4), Finland (9.4), and Sweden (9.3). Somalia and North Korea came in last, with scores of 1.0. The United States ranked twenty-fourth, with a score of 7.1.²⁷

Internal and external wars: These can occur at any time.

Changes in government: A friendly government may be replaced by a hostile one. At the extreme, the new government may expropriate foreign property.

These are just some of the risks that an MNC must take into consideration. No listing can be entirely complete.

The PRS Group, Inc., publishes the “International Country Risk Guide.” Its composite index combines three measures of risk: political, economic, and financial. Each component rates countries on a large number of risk categories. The publication has ranked 140 countries for many years. The highest achievable rating (i.e., the lowest risk) is 100, the lowest is 0. In a recent issue, Norway (89.0), Brunei (88.8), and Switzerland (88.5) topped the list, and Somalia (41.3), Zimbabwe (44.3), and Sudan (49.3) were last. The United States was thirtieth, with a score of 76.²⁸

EXCHANGE RATES

One important risk that an MNC faces has to do with exchange rates. Companies that sell in or buy from foreign countries and receive or make payments later in foreign currencies must always consider the possibility that exchange rates may turn unfavorable during the period between the setting of the price and the time when payment is made.

An **exchange rate** is the price of one country’s currency in terms of another country’s. An exchange rate may be quoted in terms of the domestic or foreign currency. If we want to measure the euro in terms of the U.S. dollar, and at a given time the euro is worth \$1.40, then we write the exchange rate as €1/\$1.40. If we want to express the exchange rate in terms of the euro, then we write it as \$1/€0.714.

Assume a U.S. exporter sells goods worth €100,000 to an importer in Germany when the exchange rate is €1/\$1.40. At today’s exchange rate, the shipment would be worth \$140,000. When the payment becomes due in 60 days, the exchange rate is quoted at €1/\$1.35. The euro has weakened during this interval. The U.S. exporter will receive only \$135,000.

Now let us look at the exchange rate risk faced by an importer. The U.S. company purchases goods from a German exporter in the amount of €100,000, at a time when the exchange rate is €1/\$1.40. The shipment is worth \$140,000. Sixty days later, when the payment becomes due, the exchange rate is quoted at €1/\$1.45; the dollar has weakened, and now the U.S. importer will have to pay \$145,000. Of course, if the exchange rate movement had been opposite, the U.S. company would have gained.

EXCHANGE RATE HEDGING

Exchange rate hedging refers to various ways that companies can protect themselves from such a potential loss. Of course, a U.S. company could attempt to have all transactions denominated in U.S. dollars. This, however, is not always possible.

²⁷Transparency International, *Corruption Perception Index 2011*, <http://www.transparency.org>.

²⁸The PRS Group, Inc., *International Country Risk Guide*, February 2012.

Offsetting Transactions

A U.S. company could export goods of the same amount to the same country from which it had imported goods spanning the same period of time. Thus a change in the exchange rate would be offset. Although such transactions may be the easiest way to hedge against possible exchange rate losses, this is not always possible or practical.

The Forward Market

A company wanting to protect itself against adverse exchange rate fluctuations can enter the forward market. Although the spot exchange rate is the price of currency today for immediate delivery, a forward contract permits a company to buy or sell currency at a specific rate at a specific time.

Assume a U.S. importer owes €1 million to a French exporter 60 days from now. If the importer does not want to be affected by a sudden strengthening of the euro, it may execute a contract to buy €1 million 60 days forward. Thus it assures itself of a specific price that it will have to pay in 60 days.²⁹ In contrast, a U.S. exporter who expects to receive a payment of €1 million 60 days from now may contract to sell the euros 60 days from now, thus protecting itself against an unexpected weakening of the euro.

Both London and New York have extensive forward markets. Forward transactions are usually executed by commercial banks, and sometimes by currency brokers. Contracts are usually available for amounts of \$1 million or greater and can be made for 30 to 360 days, and sometimes longer.

The Futures Market

Futures contracts are similar to forwards. However, there are several differences.

Futures contracts are made for standard amounts, and they mature on certain days (on Wednesdays of specific months). When a future is purchased, a certain amount—called the margin—must be deposited. The contracts are marked to market each day and transaction fees are negotiated. The largest futures market in the United States is the International Monetary Market of the Chicago Mercantile Exchange. Again, as in the case with forwards, if a person or a company expects the foreign currency to rise, the futures will be purchased. If a weakening of the foreign currency is expected, the futures will be sold.

Currency Options

A currency option is a contract that gives the buyer the right to buy or sell a certain amount of currency at a specified price during a certain period of time. A call option gives the purchaser the right to buy the foreign currency, whereas a put option represents the right to sell. The price of the option for each unit of currency is called the premium, and the exercise (or strike) price is the price that must be paid when the option is exercised. Each contract specifies the expiration date of the option—the last day on which the option can be exercised.³⁰

A company will purchase a call option if it has an obligation to pay in a foreign currency in the future. If during the period before expiration the price of the foreign currency rises above the strike price, the company will exercise the option. If, however,

²⁹Of course, if the euro were to weaken during the next 60 days, the importer could have made a gain. But again, we must remember that most people are risk averse, and thus prefer to assure themselves that they will not incur a loss.

³⁰An American option can be exercised at any time prior to the expiration date, whereas a European option can be exercised only on the expiration date.

the foreign currency remains below the strike price, the company will let the option expire. As mentioned previously, the purchase of an option gives the buyer the right to exercise the option; however, it has no obligation to do so.

Here is a brief example of a U.S. company that will have to pay a French supplier in 90 days. The present exchange rate is \$1.40 for €1. Because the company wants to obtain euros to pay the French company, it buys a 90-day call option with a strike price of \$1.42/€ at a premium of \$0.015 (1.5 cents) per option. If at the expiration date the exchange rate is \$1.44/€, the company will exercise the option. It will gain 2 cents per option, and after deducting the cost of the premium it will have a net gain of \$0.005 per option. If a €1 million option had been purchased, the net gain would have been \$5,000. Even if the exchange rate 3 months hence is only \$1.43/€, the company would still exercise the option. It will only lose \$0.005 per option because it would recover a part of the premium. If the exchange rate is less than \$1.42/€ the company would let the option lapse. In this case it will lose only what it originally paid for the option.

Similarly, when a company expects to receive funds in the future, it can protect itself against a decline in the value of a foreign currency by purchasing a put. Options are traded on various exchanges, for instance, the Chicago Mercantile Exchange.

Currency Swaps

A currency swap can occur when two companies in two different countries expect a cash flow from other companies in their respective countries. Assume U.S. company A sells its products to customers in the United Kingdom and will receive payment in pounds sterling (£), whereas company B, located in the United Kingdom sells its product to companies in the United States and expects payment in U.S. dollars. However, each company wants to receive its own currency. In that case, the two companies can arrange a currency swap to receive payment in their own currencies. Such transactions can be negotiated directly or through a bank. Thus, both companies can be sure to receive payments denominated in their own currencies and avoid a possible exchange rate loss.

FOREIGN DIRECT INVESTMENT

As corporations move into their multinational phase, they expand their operations into foreign countries. These operations may involve the licensing of foreign companies or joint ventures. But in creating a permanent presence abroad, they will begin to invest in real assets. They may do this by acquiring fixed assets, buying out existing firms, or establishing foreign subsidiaries with their own infrastructure. This is called **foreign direct investment (FDI)**.

The main reason for a company's foreign investment is, of course, to increase its earnings and increase the value of the company. Companies that experience heavy competition in their domestic market may seek a competitive advantage in foreign markets. Their domestic market may have become saturated and their growth potential can only be realized by entering new previously unexplored markets. In contrast, a corporation may have a competitive advantage in its home country and by moving abroad will be able to exploit this advantage worldwide. Another reason for FDI may be that a foreign country imposes restrictions on imports; by establishing itself in that country, the company will avoid these restrictions and may even be welcomed there if it contributes to the country's balance of payments by exporting its products to another country. The company may also be able to take advantage of economies of scale, as well as lower production and transportation costs.

Before making a foreign investment, whether it is to purchase a company, establish a subsidiary, or undertake a large project, the company will have to conduct a capital budgeting analysis to ascertain whether the project will be profitable (i.e., whether

it will have a positive net present value (NPV) or whether its return on the investment will exceed the cost of capital). We turn to this subject next.

MULTINATIONAL CAPITAL BUDGETING

The method of capital budgeting for an MNC proceeds along the same lines as for a domestic corporation. We must identify the investment made in a project, estimate the cash flows, establish a discount rate, and then solve for NPV and/or internal rate of return (IRR). However, in addition to the complications encountered in doing a capital budgeting analysis for a domestic corporation, there are several other variables that must be taken into consideration when we deal with **multinational capital budgeting**.

Intercompany Fund Flows

There will be cash flows going from the parent to the subsidiary, and vice versa. The parent will tend to finance, at least partially, large projects in a subsidiary. It may provide equity investment or loans. As the project gets under way, the subsidiary will most likely begin to remit cash payments—dividends—to the parent. Further, the subsidiary may also pay license fees or royalties (e.g., for the use of the parent's patents used in the products made by the subsidiary). The parent may also ship certain products, such as components, to be assembled and marketed by the subsidiary for which the subsidiary must pay the parent.³¹

Inflation Rates

The rates of inflation in the country of the parent corporation and of the subsidiary may differ. The sales price and the cost of production in the subsidiary country may change at a different rate from those of the home country. Therefore, expected inflation must be included in the calculation.

Exchange Rates

The exchange rate between the parent and subsidiary country will change during the project period. This change may be based on the differential inflation between the two countries. Because the subsidiary may remit its profits, or part of the profits, and other payments to the parent, it is extremely important that expected changes in exchange rates be considered in the analysis. Of course, unexpected changes cannot be predicted. The company may engage in hedging to avoid losses from unexpected changes in the exchange rate. However, because a capital project will span over many years, long-term hedges are needed. They are available for certain leading currencies (euro, yen, etc.); however, many of the world's currencies cannot be hedged for longer periods of time.

Tax Differences

Tax rates differ between the two locations and this has to be taken into account when a capital project is under consideration. But differences in income tax rates are not the only problems that may be encountered. A country may not only have a tax on profits made locally, but may also charge an additional tax on remittances to the parent's country. This withholding tax can be levied on remitted profits (dividends), as well as license fees, royalties, or interest payments made by the subsidiary to the parent; the tax rates can differ among the various payments.

³¹This brings up the question of the price the parent should charge the subsidiary. Transfer pricing is an important aspect of the parent–subsidiary relationship and is discussed later in this chapter.

Another complication arises. When the parent receives payments from the subsidiary that are taxable to it, it could be open to double taxation because taxes have already been paid in the subsidiary's country. To avoid double taxation, most countries permit the parent to credit tax payments in the subsidiary's home country. How the tax credits are treated may differ among countries. Generally, if the tax in the parent country turns out to be higher than in the subsidiary country, the parent will have to pay the difference. As a rule, the tax credit is limited to the amount of taxes that would have been paid in the parent's country.³²

To calculate the foreign tax credit, a calculation called "gross up" must be performed. The dividend by the subsidiary must be "grossed up" to a before-tax amount. This is done by dividing the dividend received by $1 - t$ (1 minus the tax rate in the subsidiary's country). Then the tax to be paid by the parent using the parent's tax rate must be calculated. The calculation is illustrated in "The Solution" vignette of this chapter, using the example posed in "The Situation" vignette.

Differences in Cash Flows

Because of the various differences mentioned previously, the cash flows received and recorded by the parent may differ substantially from those in the subsidiary's country. Thus, capital budgeting calculations must be made at both locations to determine whether the project meets the company's criteria.

Cost of Capital

When projects are evaluated both from the parent's and the subsidiary's viewpoint, they must be discounted at the cost of capital. But here again, there could be a difference between the two. Because the project may be financed partially in the subsidiary's country and because of different tax rates, as well as the possibility of obtaining subsidized loans locally, the subsidiary's cost of capital may differ substantially from that applied by the parent to the cash flows from the subsidiary. The parent company presumably uses a specific cost of capital for its investments and adjusts this rate up or down to take account of project risks. Thus, because of certain uncertainties that exist in making a cross-border investment, the company may adjust the capital cost of such a project higher than the company's weighted average cost of capital. Political risk is certainly one of those that must be considered in setting that rate.

The Final Project Valuation

As mentioned previously, the cash flows and discount rates used may differ substantially between those of the parent and those of the subsidiary. It can easily happen that the differences are so significant that the project may be deemed acceptable (i.e., positive NPV, or IRR exceeding the cost of capital) in one calculation and not acceptable in the other. Which of the two measures should prevail, the parent's or the subsidiary's?

Most financial practitioners would agree that the results from the parent's point of view should dominate. The cash flows from the subsidiary are those that add to or detract from the value of the company and the company's goal of maximizing shareholder value. Although the results from the parent's viewpoint appear to be theoretically more correct, this may not be the case in practice. Many companies appear to prefer to determine acceptability of a project based on the results in the subsidiary.

³²In the United States, an excess foreign tax credit can be offset against tax liabilities owed by the parent on income from other subsidiaries, as long as the income is of a similar nature.

THE REPOSITIONING OF FUNDS

For an MNC, the movement of funds among countries may contribute to greater profitability and an enhancement of the company's value. The positioning of funds may affect a company's tax liabilities, earnings on excess funds, and its cost of capital. Several methods of achieving this are outlined briefly. One of the methods, transfer pricing, is discussed at greater length.

Royalties and license fees can be used to channel funds to those areas of the company where they may be used most profitably.

Dividend payments to the parent are another method of controlling the positioning of funds. Because the parent company pays dividends to stockholders, dividends paid by subsidiaries to the parent are influenced by the size of the parent's dividend payout. However, other considerations enter. First, there are tax implications. Some countries apply different tax rates to distributed and undistributed earnings. Second, some countries levy a tax on dividends that are transmitted to the parent.

However, if a particular subsidiary has the opportunity to reinvest its earnings profitably (i.e., more profitably than other parts of the company), it may be asked to remit a smaller portion of its earnings, whereas subsidiaries with less favorable investment opportunities may be required to remit larger percentages of their earnings. In addition, a corporation must consider political risk and potential changes in exchange rates in shaping its dividend remittance policies. If a subsidiary must make payments to its parent or other subsidiaries (e.g., for goods purchased from the parent or another subsidiary), they can be postponed if the subsidiary that is to pay can earn higher yields than the parent or subsidiary that is supposed to receive the payment. Borrowing (interest) rates may also determine when the payment will be made.

Reinvoicing centers can be established by the company to manage the cash flows among the parent and its subsidiaries. They can direct funds to where they would earn the most.³³

MULTINATIONAL TRANSFER PRICING



Module 13A

The theory of transfer pricing is discussed in Chapter 10. Ensuring a company sets prices that will optimize profits when products are transferred from one division to another within the same company and within the same country is difficult enough. When products are transferred across borders, additional complications arise. **Multinational transfer pricing** involves products or services that are transferred from the parent company to a subsidiary or among subsidiaries located in different countries.

An MNC can affect a transfer of funds from one unit to another by charging high or low prices. If a transfer is made at low prices, this will basically shift funds to the country receiving the product. This may be profitable for a company if investment opportunities are higher in the receiving country than they are in the shipping country.

However, charging specifically high or low transfer prices may have an effect on a corporation's tax liability. The reason for this is that different countries, as we already mentioned, have different tax systems and different tax rates. Thus a company can try to take advantage of these differentials by setting its transfer price at a level to minimize its taxes.

Let us use a simple example to show how this can happen.

³³However, caution must be exercised lest the company could be accused by a government of profit shifting.

Multinational Transfer Pricing Example

An MNC has a components plant in its parent country, P. It ships these components to its subsidiary in country S, where other components are produced and then assembled into a final piece of electronic equipment that will then be sold in S.³⁴ Assume that during one year P ships 100,000 units with a production cost of 2 per unit to S. Operating expenses in P are 35,000 per year, while in S they are 25,000. Additional costs to complete the product and bring it to market in S are 1 per unit. P sets a transfer price of 3.20 per unit and S will sell the final product at 5.00 per unit. The income tax rate in P is 25 percent and in S 35 percent. In Table 13.1 we calculate the profit in each country, the combined profit and the combined income tax.

As shown in the first half of Table 13.1, the cost to the subsidiary is the revenue received by the parent (320,000) plus 100,000 of its own production cost. The combined profit for the two entities is 99,500, and their combined tax liability is 40,500. However, because the tax rate in S is higher than in P, the company would gain if the parent retained a higher portion of the profit at the expense of the subsidiary. In the second half of Table 13.1, the transfer price per unit has been increased to 3.60. The parent's profit now rises to 93,750 while the subsidiary's profit declines to 9,750. Because of the lower tax rate in S, the combined profit has risen by 4,000 to 103,500, and the tax liability has decreased by the same amount to 36,500.

If the income tax rate had been higher in the parent country than in the subsidiary, the corporation could have increased its profits (and cut its tax liability) by lowering the transfer price.³⁵ If a tariff is imposed by the importing country, this will offset, at least to some extent, the advantage of a high transfer price.

Table 13.1 The Effect of Taxes on Profit in Transfer Pricing

	Parent	Subsidiary
Income tax rate	25.0 %	35.0 %
Revenue	320,000	500,000
Cost of production	<u>200,000</u>	<u>420,000</u>
Gross profit	120,000	80,000
Operating expenses	<u>35,000</u>	<u>25,000</u>
Profit before taxes	85,000	55,000
Income taxes	<u>21,250</u>	<u>19,250</u>
Profit after taxes	63,750	35,750
Combined result:		
	Taxes	40,500
	Profit	99,500
Revenue	360,000	500,000
Cost of production	<u>200,000</u>	<u>460,000</u>
Gross profit	160,000	40,000
Operating expenses	<u>35,000</u>	<u>25,000</u>
Profit before taxes	125,000	15,000
Income taxes	<u>31,250</u>	<u>5,250</u>
Profit after taxes	93,750	9,750
Combined result:		
	Taxes	36,500
	Profit	103,500

³⁴To make this example simple, let us assume that the exchange rate between the currencies of the two countries is $1P = 1S$. This will permit us to explain the example without the complication of translating the values of the currencies.

³⁵Readers can calculate similar results by turning to Module 13A. The module includes the effects of tariffs.

A simple equation can be used to ascertain the effect on the tax liability due to a change in the transfer price:

$$\Delta T = (Q \times \Delta P \times t_e) - (Q \times \Delta P \times t_m)$$

where ΔT = Change in the total tax bill

Q = Quantity of products shipped by E (exporter) to M (importer)

ΔP = Change in the price of the product

t_e and t_m = Tax rate in the exporting and importing countries, respectively

If $t_e > t_m$, combined income would increase if transfer prices are lowered, and if $t_e < t_m$, combined incomes would rise if transfer prices are increased.

Transfer Pricing in Practice

The potential for decreasing the tax liability is a subject that is attracting a great deal of attention by tax authorities in many countries.

In the United States, Section 482 of the U.S. Internal Revenue Code gives the Internal Revenue Service (IRS) the authority to “shift around income and expense figures to arrive at what the government considers a more equitable result.”³⁶ In such cases, the burden of proof is on the taxpayer to prove that the IRS has been incorrect in reallocating income.

The IRS requires that transfer pricing be done on an “arm’s length” relationship. There are several methods of calculating such a relationship; however, the best evidence is that prices resemble those that would be established between two independent companies.

Although regulations regarding transfer pricing have been present in industrialized countries for many years, as globalization becomes even more prominent, developing countries are now becoming active in this area. For instance, the government of Vietnam amended its tax laws on January 1, 2004, to include specific rules regarding transfer pricing. The Income Tax Department of India issued its first transfer pricing orders in 2004, asking some MNCs in India to pay additional taxes.³⁷ Following the recent economic and financial crises, many governments face budgetary shortages and have thus strengthened their transfer pricing regulations and enforcement in order to increase revenues. In the United States, the Internal Revenue Service added 100 specialists to tighten the enforcement of transfer pricing laws.³⁸

Here are just a few examples of transfer pricing actions taken by governments. In January 2004, the IRS notified GlaxoSmithKline (GSK), the prominent pharmaceutical company, that it owes the government \$5.2 billion (\$2.7 billion in back taxes and \$2.5 billion in interest charges). The IRS claims that GSK has “used development and production of drugs outside the U.S. to improperly avoid taxes due” since 1989. In 2006 GSK agreed to pay the Internal Revenue Service \$3.4 billion to settle the case.³⁹ This settlement was the largest in the IRS’s history.

In February 2007, Merck & Co. agreed to settle a transfer pricing case with the Internal Revenue Service for about \$2.3 billion. An agreement dating back to 1993

³⁶C. Carroll, “IRS Target Local Company in ‘Transfer Pricing’ Case,” *Houston Business Journal*, February 3, 1997.

³⁷“Finance—Vietnam Tightens Management on Multinationals’ Transfer Pricing,” *Vietnam News Brief Service*, January 16, 2004; “I-T’s First Transfer Pricing Order to Demand More Tax from MNCs,” *Economic Times*, February 18, 2004, p. 5.

³⁸Brian Tully, “Transfer Pricing Strategies and the Impact on Organizations,” *Financial Executive*, July/August 2012, pp. 34–37.

³⁹David Firm, “GSK Faces \$5.2 Claim for Tax and Interest,” *Financial Times*, January 7, 2004; David S. Hilzenrath, “Glaxo to Pay IRS \$3.4 Billion; Tax Settlement Is Biggest in Agency’s History,” *Washington Post*, September 12, 2006.

allowed Merck to shift taxable income to a subsidiary of a United Kingdom bank.⁴⁰ More recently AstraZeneca, an Anglo-Swedish drug company, was involved in two transfer pricing cases. In 2010, it agreed to pay £505 million to British tax authorities in a case that spanned a 15-year period, from 1996 to 2010. In 2011, the company settled a case with the United States for \$1.1 billion.⁴¹ In 2006, Symantec Corporation, a large antivirus software producer, agreed to settle one dispute with the Internal Revenue Service for \$36 million. In another claim the IRS has charged Symantec about \$1 billion. This amount relates to transactions by Veritas Software Corporation (which Symantec acquired in 2005) with its subsidiary in Ireland.⁴²

Several large technology firms have recently reported disputes with the IRS regarding transfer pricing. Among these firms are AOL, Inc.; Hewlett-Packard Company; Microsoft Corporation; and Yahoo, Inc. In a case involving Amazon.com, Inc., the IRS has proposed a \$1.5 billion tax increase relating to transfer pricing over a period starting in 2006. Amazon is planning to contest the IRS's proposal.⁴³

A 2010 Global Transfer Pricing Survey of 877 multinational companies across 25 countries by Ernst & Young reported that in 2010, 67 percent of the companies responding had experienced a transfer pricing audit, compared to 52 percent in 2007.⁴⁴

GLOBAL APPLICATION: TONY THE TIGER MEETS A BENGAL TIGER⁴⁵

The Kellogg Company is one of the authentic and venerable global food companies around the world today. It has been in business for over 100 years. It has manufacturing facilities in eighteen different countries and sells its products in more than 180 countries around the world.⁴⁶ With all of Kellogg's experience and success as a global food company, we think it is interesting to note the risks that it had to deal with in its early efforts to enter the markets of developing economies. A case in point is Kellogg's initial efforts to get into the Indian market in 1994.

In the early 1990s, Kellogg saw that the U.S. market was stagnating and revenues were being generated only by small price increases. It wanted a new market, and the Indian market held great significance for the Kellogg Company because of the sheer size of its young population. (Remember that India at that time had not even begun to grow at the rate it has over the past decade.) Unfortunately, Kellogg's initial offerings in India included cornflakes, wheat flakes, and Basmati rice flakes. Despite the high quality of its products and the backing of corporate technical, managerial, and financial resources, Kellogg's products failed in the Indian market.

In assessing the causes of its failure, Kellogg realized that it was not the quality of its products but a misunderstanding of Indian tastes and preferences that stymied their efforts. A distinguishing feature of Kellogg's breakfast cereals is their crispiness. But Indians prefer to use warm milk in their cereal, and pouring warm milk on Kellogg's corn flakes, no matter how crispy they are inside the wax liner in the box, made them soggy. Indians also prefer hot or warm food for breakfast. And cold milk

⁴⁰"Merck Settles IRS Probe of 1993–2001 Taxes," *Dow Jones News Service*, February 14, 2007.

⁴¹Richard Wachman, "AstraZeneca to pay pounds 505m tax bill after 15-year transfer pricing row," *The Guardian*, February 24, 2010; Graham Ruddick, "UK loses out to US in AstraZeneca tax row," *The Daily Telegraph*, March 29, 2011.

⁴²"Symantec Faces \$1B Tax Demand," *Reuters Technology News*, June 30, 2006; "IRS Hits Symantec with \$1 bn Tax Bill for Irish Arm," *Irish Independent*, April 19, 2006.

⁴³"Transfer Pricing Brings Tax Troubles to Tech Giants," *Dow Jones News Service*, April 2, 2012.

⁴⁴Ruan Joste, "Transfer pricing in numbers," *Financial Mail*, January 28, 2011.

⁴⁵We are grateful to Navin Punjabi for suggesting this topic and for providing us with content references.

⁴⁶<http://www.kelloggcompany.com/company.aspx?id=32>.

over a cold bowl of corn flakes will not taste great, no matter what Tony the Tiger says (“Tastes greaaaat!”).⁴⁷

The challenge for Kellogg was to persuade Indian consumers to eat corn flakes and its other products. It began by launching Kellogg’s Breakfast Week,⁴⁸ a community service initiative to generate awareness about eating breakfast. The program focused on the prevention of anemia and offered a series of nutrition workshops and activities for both individuals and families. As summarized by an expert in marketing in India, “Some product categories are more conducive to local interpretation. If you don’t honor that, you’re likely to remain on a slow growth curve. Food is perhaps the most culturally sensitive category.”⁴⁹

As evidenced by the discussion about building the dressings category earlier in this chapter, global food companies today are well aware of the need to “do their homework” in studying local tastes and preferences among the different countries of the world in which they operate. They know that one of the best ways to mitigate the risk of launching products in different countries is to make sure that their products and marketing efforts take these local factors into account.

The Solution



Module 13B



With all the data at hand, George Kline begins to input the information into the Excel spreadsheet he had prepared (Table 13.2).

At the top of the sheet, he enters the percent volume change and the price level changes. He could have entered different numbers for each year, but he chose not to.

He then enters today’s exchange rate in year 0. The new exchange rates are calculated automatically based on the relationship between U.S. and Czech inflation.

He enters the fixed asset investment and the working capital in the year 0 column. Next he enters the various factors in the “Constant” column. In the “Depreciation” line he obtains the annual depreciation by dividing I by the depreciation life of the property. He sets the terminal value at 100 million KCZ. Any of these inputs can be easily changed if his assumptions change and new iterations of the solution are desired.

He also enters the sales revenue, production cost, and imports in year 1. George now goes over the results of his calculations. The NPV for the subsidiary is positive and the IRR exceeds the cost of capital. Thus, it would appear that the project is acceptable from the subsidiary’s viewpoint. George notes, however, that the IRR is just barely above the cost of capital and NPV is relatively small. Any small adverse change in his assumptions would probably cause the NPV to turn negative.

He next looks at the results from the parent company’s viewpoint. Here the NPV is negative (and the IRR substantially below the cost of capital). George is very familiar with the opinion of most of the experts in the financial area—that the parent’s results are the ones that should be considered primarily.

George arranges for a meeting with the company’s executives, including the vice presidents (VPs) of finance, marketing, and manufacturing, to present his results. He recommends that the project should be shelved at this time. However, he points out that the various assumptions he made may be subject to change, and proposes to consult the various departments in 3 months and repeat the analysis. The VP of manufacturing suggests that better results could possibly be obtained if the company were to build a new, more efficient plant (i.e., to make a greenfield investment) and thus improve the profitability of the project. He asks George to make another calculation with new assumptions. However, the VPs of marketing and finance argue that this would probably only make the results worse. They have several reasons for this. From the marketing viewpoint, the VP points out that this is not a new product in Europe and that the

(continued)

⁴⁷<http://www.icmrimdia.org/casestudies/catalogue/Marketing/Kellogg's%20Indian%20Experience.htm>, retrieved on July 24, 2012.

⁴⁸Business Standard, April 17, 1997, Kellogg’s Breakfast week in Calcutta, <http://www.business-standard.com/india/news/kelloggs-breakfast-week-in-calcutta/41852/> retrieved on July 24, 2012.

⁴⁹Knowledge@Wharton, March 12, 2009, Made for India: Succeeding in a Market Where One Size Won’t Fit All, <http://knowledge.wharton.upenn.edu/india/article.cfm?articleid=4358> retrieved on July 24, 2012.

(continued)

company would face stiff competition. Because the demand for the product is probably quite price elastic, the company would have to be very careful in its pricing. The VP of finance adds that even though a brand new plant may be more efficient, the advantage of purchasing the existing plant is the relatively low price, resulting in lower depreciation charges.

The final decision is to hold off on this project. However, George is asked to review his numbers with the various staffs and report back to the executives in about 2 months.⁵⁰

Table 13.2 Capital Project Evaluation

	Constant	Year 0	Year 1	Year 2	Year 3	Year 4
<i>Assumptions</i>						
Volume change			6.0%	6.0%	6.0%	6.0%
Price level %Δ in United States			2.0%	2.0%	2.0%	2.0%
Price level %Δ in Czech Republic			3.5%	3.5%	3.5%	3.5%
Sales price %Δ in Czech Republic			4.0%	4.0%	4.0%	4.0%
Cost %Δ in Czech Republic			4.0%	4.0%	4.0%	4.0%
Exchange rate—KCZ/\$		27.00	27.41	27.82	28.23	28.66
Fixed-asset investment (\$000)		4,000				
Working capital (\$000)		400				
Cost of capital United States	19.0%					
Cost of capital Czech Republic	22.0%					
<i>Income Statement for Czech Subsidiary (Thousands of KCZ)</i>						
			Year 1	Year 2	Year 3	Year 4
1. Sales revenue			400,000	440,960	486,114	535,892
2. Production costs— Czech Republic			190,000	209,456	230,904	254,549
3. Imports from United States			110,000	120,716	132,476	145,381
4. License fees	3.0%		12,000	13,229	14,583	16,077
5. G&A expense	12.0%		48,000	52,915	58,334	64,307
6. Depreciation	12.5%		13,500	13,500	13,500	13,500
7. EBIT			26,500	31,144	36,317	42,078
8. Income taxes	31.0%		8,215	9,655	11,258	13,044
9. NEAT			18,285	21,489	25,059	29,034
10. Cash dividend	50.0%		9,143	10,745	12,529	14,517

(continued)

⁵⁰In 2012, when this edition was being prepared, the exchange rate for the Czech Koruna was about 19 to the U.S. dollar. However, we decided to keep the exchange rate at 27 to show that a project can be financially acceptable from the subsidiary's viewpoint and not acceptable from the parent's viewpoint. Readers may want to insert the new exchange rate in Module 13B to see how the results are affected.

(continued)

(continued)

Working Capital (WC) Requirements
(Thousands of KCZ)

	Constant	Year 0	Year 1	Year 2	Year 3	Year 4
1. Required at $t = 0$		10,800				
2. Year-end requirement (as percent of sales)	10.0%		40,000	44,096	48,611	53,589
3. Required additions			29,200	4,096	4,515	4,978
4. Financed internally in Czech Republic	50.0%		14,600	2,048	2,258	2,489
5. New WC required			14,600	2,048	2,258	2,489

Cash Flows for Czech Subsidiary
(Thousands of KCZ)

1. NEAT			18,285	21,489	25,059	29,034
2. Depreciation			13,500	13,500	13,500	13,500
3. Terminal value	100,000					100,000
4. WC additions			-14,600	-2,048	-2,258	-2,489
5. Original investment by U.S. parent		-118,800				
6. Net cash flow		-118,800	17,185	32,941	36,301	140,045
7. NPV		626				
8. IRR		22.2%				

**Calculation of Dividends Received
by Parent from Subsidiary**

(KCZ in 000)						
1. Cash dividend paid		9,143	10,745	12,529	14,517	
2. Dividend percent of tax paid		4,108	4,827	5,629	6,522	
3. Grossed-up dividend (\$ in 000)		13,250	15,572	18,158	21,039	
4. Grossed-up dividend		483	560	643	734	
5. Income tax in United States	35.0%	169	196	225	257	
6. Credit for sub's taxes		150	174	199	228	
7. Additional tax due in United States		19	22	26	29	
8. Cash dividend paid to parent		334	386	444	507	
9. Aftertax div. rec. by parent		314	364	418	477	

(continued)

(continued)

(continued)

(continued)

<i>Cash Flows Received by U.S. Parent</i>						
	Constant	Year 0	Year 1	Year 2	Year 3	Year 4
(KCZ in 000)						
1. License fees from Czech subsidiary			12,000	13,229	14,583	16,077
2. Profit on exports to Czech subsidiary	5.0%		5,500	6,036	6,624	7,269
3. Total receipts from Czech subsidiary (\$ in 000)			17,500	19,265	21,207	23,346
4. Total receipts from Czech subsidiary			639	693	751	815
5. Income tax in United States			223	242	263	285
6. Aftertax receipts			415	450	488	530
7. Aftertax dividend			314	364	418	477
8. Project cost		-4,400				
9. Terminal value						3,490
10. Net cash flows		-4,400	729	814	906	4,496
11. NPV						-432
12. IRR						15.1%

SUMMARY

A multinational corporation must compete not just domestically but worldwide. As we discuss in previous chapters, managers must always be concerned with the economic environment facing them. They must inform themselves about the demand for their products, the cost of supplies and components, the productivity of their machinery and labor force, and changes in technology. They must digest this information to try to maximize the corporation's profitability (in a broader definition, the corporation's market value). We have learned that, to achieve best results, corporate managers must be aware that any increase in revenue must be greater than costs when the corporation grows; in other words, they must apply the concepts of marginal revenue and marginal cost.

An MNC must consider these factors and many others:

- **Economic factors:** Exchange rates and exchange rate changes, differences in cost of capital, economic stability of the foreign country, production costs in foreign country influencing the choice of sourcing locations
- **Political factors:** Stability of government institutions, different tax systems, restrictions on foreign ownership, blockage of fund transfers, laws regarding employment and wages, bureaucracy and corruption, attitude of government toward multinational corporations, expropriation, war, terrorism
- **Social and cultural factors:** Religious differences, differences in the hiring of and promotion of female employees, different attitudes toward profit maximization

We describe in detail the differences in capital budgeting between a domestic and a multinational corporation. We specifically direct our attention at differences in taxation,

consideration of differential rates of inflation and their influence on the exchange rate, and differences in the cost of capital. We also discuss the question of whether the decision to make an investment should be based on the results in the subsidiary or the parent.

Finally, we addressed the extremely important topic of transfer pricing. A corporation can achieve higher profits (and lower taxes) by manipulating the prices charged for products or services flowing from the parent to a subsidiary or among subsidiaries. Governments are becoming very wary of these practices and, in many cases, corporations have been required to pay additional taxes.

IMPORTANT CONCEPTS

Exchange rate. The price of one country's currency in terms of another. (p. 518)

Exchange rate hedging. Actions that a company can employ to protect itself from exchange rate fluctuations. Among these are forward contracts, futures, currency options, and currency swaps. (p. 518)

Exchange rate risk. The possibility that exchange rates will change to the detriment of an MNC. (p. 517)

Foreign direct investment (FDI). Investment in real assets by an MNC, including the acquisition of fixed assets, purchase of a foreign plant, or establishment of a subsidiary. (p. 520)

Multinational capital budgeting. Decision making by an MNC when it seeks to undertake a long-term project in a foreign country. (p. 521)

Multinational corporation risk. Risks that are faced by a multinational corporation not directly faced by a domestic corporation. These include economic, political, and social and cultural risks. (p. 517)

Multinational transfer pricing. The pricing of goods and services when these are transferred across borders. These decisions are affected by different tax rates between countries and import tariffs. (p. 523)

QUESTIONS

1. What is the difference between a futures contract and a forward contract?
2. A U.S. importer who owes a Belgian company €500,000 payable 30 days from today expects that the US\$ will weaken during this period. What would you advise the importer to do? What would happen if the US\$ were to strengthen during this period?
3. A U.S. importer purchases a currency option. If the foreign currency does not rise to the strike price, what should the importer do?
4. Describe the additional complications facing an MNC compared with a domestic corporation when it is evaluating a capital budgeting project.
5. Why should an MNC's capital budgeting decision be based on the parent's results rather than those of the subsidiary?
6. Is an MNC generally faced with incurring double taxation on its profits in the subsidiary's country? Why or why not?
7. Why should a government be concerned with the pricing of products that a company transfers to an affiliate in another country?

PROBLEMS

1. The Great Computer Company, a U.S. corporation, has a subsidiary in the Netherlands. It is deciding whether to invest \$2 million of its (the parent's) funds in a 3-year project in the Netherlands.

The aftertax cash flows to the subsidiary are estimated to be as follows (in euros):

Year 1	€500,000
2	800,000
3	900,000

The entire cash flows of the subsidiary are remitted to the parent annually. There is no additional tax (nor credit) in the parent country.

The exchange rate today is €1/\$1.20. The exchange rate forecast for the next 3 years is the following:

Year 1	€1/\$1.15
2	€1/\$1.10
3	€1/\$1.05

The cost of capital for both the parent and the subsidiary is 13 percent.

- a. What is the NPV of this project to the Netherlands' subsidiary?
 - b. What is the NPV of this project to the U.S. parent?
 - c. Should the project be accepted?
2. The XYZ Multinational Corporation has manufacturing facilities in country A and an assembly plant in country B. The company ships manufactured units from its plant in A to its assembly plant in B.
 - a. In April 2013, the company will ship 1,000 units with a production cost of 65 per unit to its plant in country B. Its operating expenses in A are 15,000 for the month. The income tax rate in A is 20 percent and in B 40 percent. The company plans to have a transfer price of 100 per unit. The final product can be sold in B for 140. B's operating expenses are 10,000 during the month. How much will the combined profits be of the two operations in April 2013?
 - b. Could the company benefit by changing the transfer price to 120?
 - c. Now, suppose the income tax rate in A is 40 percent, while in B it is 20 percent. What will the combined profit be if all other numbers are the same as in a?
 - d. What would be the result in c if the company decreased its transfer price to 90?
 3. Today the XYZ Corporation shipped goods valued at €1 million to a customer in Belgium. Payment is due in 90 days, and the Belgian firm will make the payment in euros. Today's spot rate is €1/\$1.40. The 90-day forward rate is €1/\$1.38.
 - a. How many dollars would XYZ receive if payment were made today?
 - b. If XYZ sells €1 million forward for 90 days, how much is it assured to receive 90 days from now?
 - c. If XYZ had not hedged in the forward market and the spot rate 90 days from now is €1/\$1.39, how much would XYZ receive (in U.S. dollars)?
 - d. If the U.S. dollar were to weaken in the 90 days and XYZ did not hedge, would it benefit or lose?
 4. The ABC Company expects to receive payment in euros from a German company in 60 days. To protect itself from a decline in the value of the euro, it purchases a put option with a strike price of \$1.42/€ at a premium of \$0.015 (1.5 cents) per option.
 - a. If the exchange rate 60 days from now is \$1.40/€, should ABC exercise the option? How much will it gain or lose on the transaction?
 - b. If the exchange rate is \$1.41/€, what should the company do and why?
 - c. At what exchange rate will the company break even?

Government and Industry: Challenges and Opportunities for Today's Manager

Learning Objectives

Upon completion of this chapter, readers should be able to:

- Cite the five major functions of government in a market economy.
- Understand the meaning of market externalities and the reasoning of the Coase theorem in its contention that government involvement may not be necessary to deal with market externalities.
- Understand the causes and consequences of the 2007–2009 financial crisis.
- Explain why firms merge and why, in particular, firms have chosen to merge in markets that have experienced government deregulation.
- Briefly explain the concept of intellectual property (IP) and the role of government in protecting IP rights.

The Situation

In an offsite retreat, Bob Burns was leading a discussion for all of the members of his C-level team. Present at the meeting were the chief operating officer (COO); chief financial officer (CFO); chief technology officer (CTO); chief marketing officer (CMO); and, given the topic of discussion, the chief legal counsel (CLC). In these types of corporate retreats, senior leaders try to talk about the big picture and the company's long-term strategy.

Bob Burns told the group, "About 20 years ago, I decided that Global Foods should enter the carbonated soft drink market. I did this because I felt that this was a fast-growing market with very attractive profit margins. I'm starting to see that this market is no longer as fast growing or profitable as it was back then. What really got me thinking about shifting our focus to new or different product categories within the beverage industry is the recent effort by the mayor of New York to put legal limits on the size of the serving of carbonated soft drinks."¹

(continued)

¹See http://www.cbsnews.com/8301-504763_162-57451372-10391704/mayor-bloombergs-soda-ban-proposal-to-be-submitted-to-nyc-health-board-today/.

(continued)

“Oh, I wouldn't let that worry you, Bob,” said Samantha Brown, the company's chief legal counsel. Our legal team and those of our competitors can combine to fight such moves.”

“Yes, and however it turns out, our research shows that the publicity regarding the issue of consumer freedom will actually be good publicity for us,” added Allison Weitz, the company's chief marketing officer.

“I appreciate your words of support for our flagship product. But regardless of the legal environment, I think we should look behind the move by New York City to try to limit the size of our servings. There is no question in my mind that there is a long-term trend toward healthier foods and drinks. Look how well our bottled-water business has done. I want to form a special task force—led by you, Allison—to come up with a strategic plan to move aggressively and quickly into healthier products.”

INTRODUCTION

The primary objective of this chapter is to discuss the impact of government policies on managerial decision making. When the government is involved in the market economy, it generally controls the behavior of buyers and sellers through a process of “indirect command.” That is, rather than ordering buyers and sellers to allocate resources in a particular way, the government uses market incentives or disincentives. This “visible hand” of the government can take such forms as price controls, rules and regulations, taxes, and subsidies. In using the incentive of profit or the disincentive of loss, the government does not change the basic system of rewards and punishments used in the market. Instead, it simply alters the reward structure of a laissez-faire market so resources are allocated more in accordance with government policy than with the actions of individual buyers and sellers.

In the following section, we elaborate on the different roles that government plays in the market economy and the justification for such roles. We begin with a discussion of the rationale for government involvement in a market economy. We then go into more detail about the government's policies to stabilize the macroeconomy. We conclude with sections on government oversight on mergers and acquisitions and intellectual property.

THE RATIONALE FOR GOVERNMENT INVOLVEMENT IN A MARKET ECONOMY

There are five major functions the government can perform in a market economy such as that of the United States. First, it provides a legal and social framework within which market participants buy and sell the goods and services produced by the economy's scarce resources. For example, the Food and Drug Administration (FDA) seeks to ensure that food and pharmaceutical companies sell products that meet certain standards of safety and quality.

Second, the government strives to maintain competition in markets for goods and services by trying to ensure no one seller dominates the market in an unfair manner. Third, the government may decide to play a role in the redistribution of income and wealth. It can do so through the tax system (particularly through income taxes), and also through various types of government subsidies and grants for special-interest groups. For example, one of the major subsidies for middle and upper income in the United States is the ability to deduct interest payments on residential mortgages.

The fourth market-related function of government is the reallocation of resources. According to economic theory, a misallocation of resources results whenever a market has certain externalities or spillovers. That is, some of the benefits or costs

associated with the production or consumption of a particular product accrue to parties other than the buyers or sellers of a product.

The fifth major function of government in a market economy is the stabilization of the aggregate economy. The market economy is prone to periodic upswings and downswings in economic activity. As you have probably already studied in a course in macroeconomics, governments can employ monetary and fiscal policy to deal with the problems of unemployment and inflation, which usually occur at different stages of the cycle.

One government function that does not quite fit into any of the five main categories is the regulation of natural monopolies. The economic definition of a **natural monopoly** is an industry in which a single firm can serve customers more efficiently than many smaller competing firms because of the predominance of economies of scale. Examples of natural monopolies are electricity and gas utilities, and until recently, telephone services. However, over the past decade or so, there has been a movement by governments throughout the world to reduce the government ownership or regulation of these natural monopolies. Let us elaborate a bit further on three of the five main functions of government: providing a competitive framework for market participants, the reallocation of resources in the presence of market externalities, and the stabilization of the aggregate economy.

Providing a Legal Framework for Competition: The Antitrust Laws

A thorough interpretation of antitrust laws would carry us beyond the scope of this text.² Thus, we provide a brief review of the subject. The beginning of antitrust legislation in the United States dates back to the late nineteenth century in response to the formation of large corporations and the so-called merger-to-monopoly wave. While under common law various anticompetitive actions (e.g., price fixing) were not enforceable; therefore, stronger positive legal action appeared necessary. The first law to be passed was the Sherman Anti-Trust Act, followed by several others.

The Sherman Anti-Trust Act (1890)

The two important sections of this law were the following:

Section 1: . . . every contract, combination in the form of trust or otherwise, or conspiracy, in restraint of trade or commerce among the several states, or with foreign nations, is declared to be illegal . . .

Section 2: . . . every person who shall monopolize or attempt to monopolize any part of the trade or commerce among the several states, or with foreign nations, shall be deemed guilty of a felony . . .

Section 1 outlawed explicit cartels. Section 2, which appeared to outlaw monopoly, was interpreted by the courts as a prohibition of “bad acts” and not monopoly *per se*.

The Clayton Act (1914)

The Clayton Act was intended to strengthen and widen the application of antitrust enforcement, and specifically enumerated four forbidden practices:

Section 2: prohibits price discrimination that would “substantially lessen competition . . .”
[It exempted price differentiation due to differences in quality, quantity, and costs of selling and transportation.]

²A large number of texts on industrial organization treat this topic thoroughly. A recent example is Don E. Waldman and Elizabeth J. Jensen, *Industrial Organization: Theory and Practice*, 4th ed., Upper Saddle River, NJ: Prentice Hall, 2013.

Section 3: Prohibits the use of tying or exclusive contracts that would lessen competition.

Section 7: Prohibits the acquisition of other companies through purchase of stock if this lessens competition.

Section 8: Restricts interlocking directorates.

The Clayton Act also provided for the recovery of treble damages by the injured party.

The Federal Trade Commission Act (1914)

This law set up a new agency, the Federal Trade Commission (FTC), to investigate violations, enforce antitrust laws, and determine what actions constitute “unfair method of competition,” which the law declared unlawful. The Wheeler-Lea Act (1938) extended the FTC’s authority to the protection of consumers.

Subsequent Acts

The Robinson-Patman Act (1936) strengthened the price discrimination provisions of the Clayton Act to include not only the lessening of competition, but also injury to or prevention of competition “with any person who either grants or knowingly receives the benefits of such discrimination, or with customers of either of them.” This act was the result of political pressure by small independent grocery stores to make it difficult for large grocery chains to buy and sell at lower prices.

The Celler-Kefauver Act (1950) corrected some omissions of the Clayton Act. Specifically, it applied the law to mergers accomplished through asset acquisition (the Clayton Act applied only to acquisition of stock).

The Hart-Scott-Rodino Act (1976) imposed a premerger notification requirement on large firms (where the acquiring firm has at least \$100 million in assets and \$10 million in annual sales).

Government Antitrust Policy in Action: The Case of Microsoft

Over time, the interpretation and enforcement of antitrust laws has varied a great deal. Before leaving this subject, we briefly discuss the two main schools of thought regarding the purpose of antitrust laws and comment on perhaps one of the most publicized antitrust cases in recent years: the U.S. government versus Microsoft Corporation.

One school of thought suggests that the main purpose of antitrust laws is economic efficiency. The laws act against practices that restrict output and raise prices. But not all mergers and agreements will create economic inefficiencies; instead, they may lead to greater efficiency and lowering of costs. Thus courts should interpret such actions case by case. The final goal of policy should be to enhance consumer welfare.

The opposing school of thought argues that the purpose of antitrust laws is actually to limit the power of large firms and protect smaller independent firms, regardless of the effects on efficiency. This seemed to be particularly evident in the case of the Robinson-Patman Act, which was explicitly aimed at the protection of small grocery businesses. Obviously, this is an extremely complex issue that has not yet been solved, and probably will not be for many years to come. An excellent contrast between these two schools of thought is illustrated in the case of the failed attempt by GE to take over Honeywell in 2001. This move was approved by American regulators but rejected by their European counterparts. (See the “Global Application” section at the end of this chapter for details.)

One of the highest-profile antitrust cases in recent years is the U.S. government’s case against Microsoft Corporation. In a nutshell, the U.S. government has

charged that Microsoft has used its monopolistic powers to compete unfairly against the Netscape Communications Corporation in the Web browser market. (Microsoft was also accused of unfair practices against other companies; however, we confine ourselves here to the government's case regarding Netscape.) After Microsoft bundled its own product, Internet Explorer, with its Windows operating system, the government charged that the company gained an illegal competitive advantage against Netscape's Navigator. The thrust of the government's arguments rested on two points: (1) Microsoft indeed has monopolistic powers that violate current antitrust laws, and (2) Microsoft used these powers to unfairly restrict Netscape's ability to compete.

Readers can well expect that top economists were among the cast of expert witnesses called in to testify on behalf of either side. In a dramatic display of "you've got your expert and I've got mine," the government called in prominent economist Franklin M. Fisher of the Massachusetts Institute of Technology (MIT), and Microsoft countered with Richard L. Schmalensee, also of MIT. By coincidence (or perhaps not), Professor Fisher was Professor Schmalensee's Ph.D. thesis advisor.

Professor Fisher challenged the fairness of Microsoft's decision to imbed its browser into its Windows operating system, which has more than a 90 percent share of the market. Professor Schmalensee countered by saying that this is like saying that "consumers will be made better off if they are deprived, by court order, of Web-browsing functionality that Microsoft wishes to provide at zero marginal cost and that consumers are free to ignore or replace."³ Professor Fisher argued, saying that the threat is not merely that Microsoft will wrest the browser market away from the current leader, Netscape, but worse, that Microsoft might well use browser domination to eliminate competition in operating systems. He further charged that Microsoft's decision to give away its browser constituted an extreme form of "predatory pricing," designed to drive a competitor out of business.⁴

In late 2001, Microsoft settled with the Justice Department. This settlement is very controversial. Antitrust enforcers from nine states are not satisfied and have decided to pursue the case on their own. U.S. District Judge Coleen Kollar-Kotelly suggested in late 2001 that the case could proceed on parallel tracks. Under the requirements of the Tunney Act, the Justice Department settlement is subject to additional review. Meanwhile, the states could continue preparing for a final round in court.⁵ The concerns raised by the state antitrust enforcers and many other critics are quite involved. Essentially, these critics believe that the settlement will not prevent Microsoft from continuing to use the same type of monopolistic tactics that brought the suit on it in the first place.⁶

There is also an international aspect to government antitrust action against Microsoft. In 1999, following a complaint by Sun Microsystems, the European Commission began an investigation of Microsoft's business practices. The investigation revealed two questionable practices:

- Nondisclosure of interface documentation precluding rivals from making their software work well with Microsoft Windows
- The bundling of Microsoft's Windows Media Player with the Windows 2000 operating system

³"Trial's War of Economists Pits Student Against Teacher," *New York Times*, January 25, 1999.

⁴Readers should have no trouble finding material in the popular press about this case. However, two good sources are "As Microsoft Struggles with Antitrust Case, Tactical Errors Emerge," *Wall Street Journal*, February 18, 1999, and "For Microsoft, Humbled May Not Mean Defeated," *New York Times*, February 28, 1999.

⁵"States Find Flaws in Deal with Microsoft," *Wall Street Journal*, November 5, 2001.

⁶There are numerous articles about the controversial settlement besides the one cited previously. For example, see "Commentary: Settlement or Sellout?" *BusinessWeek*, November 19, 2001.

After five years, in March 2004, the Commission concluded that Microsoft broke the European Union competition law and fined the company €497 million, at that time the Commission's largest penalty. It gave Microsoft 90 days to offer Windows without the Media Player and 120 days to release "complete and accurate" information to rivals in the server market. The U.S. Department of Justice's antitrust department reacted very negatively, stating that "sound antitrust policy must avoid chilling innovation and competition by 'dominant' companies. A contrary approach risks protecting competitors, not competition, in ways that may ultimately harm innovation and consumers that benefit from it." In 2005 the Commission threatened additional fines because Microsoft had missed its deadlines and in the summer of 2006 it levied an additional fine of €280 million. Microsoft paid the fine but appealed both fines. The decision was then left to the European Union Court of First Instance.

In 2007 the Commission stated that Microsoft had continued to obtain a larger proportion of the server market, increasing from 35–40 percent in 1999 to 75–80 percent. Also, it said that Microsoft was charging royalties that were too high for the information it revealed to its competitors. The Commission wanted the information released at little or no cost.

On September 17, 2007, the Court upheld the Commission's decision on both parts of the case, saying that Microsoft unfairly withheld code from competitors at a reasonable price. Microsoft considered appealing to the European Union's highest court, the European Court of Justice. However, on October 22, Microsoft decided not to appeal, agreed to obey key parts of the 2004 antitrust ruling, and agreed to slash the charges for revealing the interface information.⁷ The results of this case seem to confirm the different interpretation of antitrust laws between the European Union and the United States. The decision here again appeared to favor the protection of competitors.

Dealing with Market Externalities: Another Key Function of Government in the Market Economy

In the microeconomic theory of the firm, on which much of managerial economics is based, it can be shown that under conditions of perfect competition, an economy's scarce resources will be most efficiently allocated and social welfare will be maximized. This situation prevails when all costs are fully accounted for in the price of the product. However, there are frequent cases when not all costs are included in the price or when not all costs are compensated. Such situations give rise to externalities and result in market failure.

A **benefit externality** is one where not all costs are compensated. In other words, certain benefits accrue to third parties. For instance, a beautiful private garden will benefit people who walk by, but this will not compensate the owner. Benefit externalities (also called *external economies* or *positive externalities*) also arise in the case of information (several people can read the same newspaper) or from innovations that benefit many people. Because producers of these products cannot appropriate all the revenue, too little may be produced.

But **cost externalities** are the ones with which economists are mostly concerned. These are the cases where a producer does not pay all the costs generated by the product. The most popular example of cost externalities (also called *external diseconomies* or *negative externalities*) in economic literature is pollution. A factory will produce both the product and pollution. But if it does not have to pay for the cost of the pollution it has created, the product will be priced based only on the "private" production

⁷The above information was obtained from a large number of newspaper articles, items from the Associated Press, AFX International Focus, and several other sources.

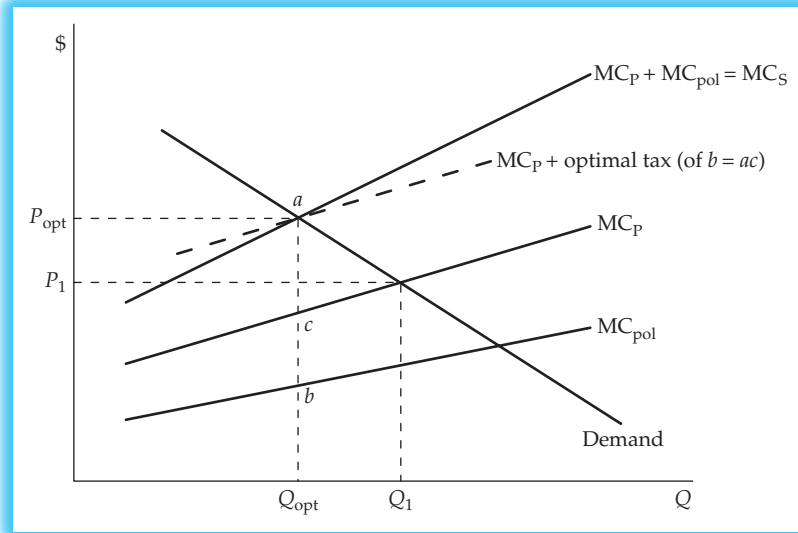


Figure 14.1 The Impact of a Cost Externality on Price: The Optimal Per-Unit Tax Is the Marginal Cost (MC) of Pollution at the Optimal Output Level

costs. Thus, the product's price will be lower than if it had been fully costed, and too much of the product will be produced. This can be shown on a simple graph for a competitive industry, shown in Figure 14.1.

The private marginal cost curve of the product is shown as MC_P . The price of the product will be P_1 and the quantity produced will be Q_1 . The cost of pollution is not considered in this result. The marginal cost of pollution is shown as MC_{pol} .

The total marginal cost of the product to the economy, the *social marginal cost*, is the vertical summation of MC_P and MC_{pol} , or MC_S . If this cost had been included in the pricing of the product, then equilibrium would have been at P_{opt} and Q_{opt} . Obviously, there will always be some pollution. If pollution were to be completely eliminated, then this industry would have to shut down and produce nothing. But this would not be an ideal situation because there is demand for the product. The **socially optimal price** will occur at P_{opt} , where the price of the product is equal to marginal social cost. Note that at this point less pollution will be produced (and less of the product also) than under strictly competitive conditions where only MC_P was priced into the product.

How can the optimal equilibrium be attained? The usual prescription is for government to act in one of two ways.⁸ The first would be to restrict production to Q_{opt} . The other method is taxation. Ideally, this tax would be equal to the MC_{pol} at the optimal output level, or height b in Figure 14.1. If such a tax were levied, then the marginal cost of production to the firm becomes $MC_P + b$. This marginal cost schedule is represented by the dashed line parallel to MC_P passing through the socially optimal production level (at a in Figure 14.1 because, by construction, $ac = b$), thereby causing the externality to be internalized. A third method is to create a market for the right to pollute. As part of the 1990 Clean Air Act, the government granted companies a certain number of allowances for generating carbon pollution. The total amount of allowances was capped

⁸The entire subject of externalities and government intervention was first treated fully by a famous English economist, A. C. Pigou. See *The Economics of Welfare*, 4th ed., London: Macmillan, 1932.

and has then been reduced each year since the act was passed. Each firm then has the option of using its pollution allowances or selling to another party, who, in effect, has determined that the market price of the allowance is less than the cost of reducing its carbon emission. The total amount of emissions is presumably reduced over time as the total amount of allowances is reduced. Allowing the trading of these allowances creates a kind of market mechanism that makes the total reduction as efficient as possible. It is also the method espoused by supporters of cap-and-trade legislation.⁹

The Coase Theorem: An Alternative Treatment of Externalities

In 1960, Ronald Coase advanced the idea that government intervention to eliminate the effect of externalities is not necessary if property rights are correctly and clearly defined.¹⁰ This is often referred to as the **Coase theorem**. Who has the right to pollute, and who has the right to keep pollution low? If it were originally decided who owns the property in question, then bargaining between the two parties would result in an optimal solution, without government intervention, at the level that would have been brought about with a correct taxation policy.

Various examples of such a situation can be given. Let us assume a stream on which we find a chemical plant and a fishing establishment. If the property rights to this stream go to the chemical plant, the fishing company will have to pay the plant to reduce the dumping of its waste into the stream. It will pay just up to the point where the marginal benefit to the fishermen (from harvesting more and healthier fish) will be equal to the cost incurred by the chemical plant (in dumping its waste at a less convenient and more expensive location, or actually decreasing its production). In contrast, if the fishing company has the property rights, the plant will have to pay the fishermen to permit the plant to dump waste. In both cases, the final result will be the same and will take place at the optimal point that would have been achieved if a correct tax policy (i.e., taxes equal to the marginal cost of polluting) had been imposed.

Other examples of this kind abound. Coase, in his famous article, described the case of farmers growing grain in unfenced fields until a cattle rancher, whose cattle tend to damage the farmers' crops, arrives in the area. In addition, in his article, Coase cited a number of legal cases spanning back to the nineteenth century where the assignment of property rights was crucial in the decision.

The Coase theorem poses interesting questions regarding the need for government interference, but it is important to look at its limitations. For instance, although optimal production levels are reached regardless of which party is assigned property rights, income distribution will be affected. In our example, if the chemical plant was awarded property rights, it will gain at the expense of the fishing industry. Thus, normative issues are not addressed by Coase.

But there are additional limitations. First, if transaction costs—i.e., the cost of obtaining an agreement between the interested parties—are high, then bargaining between interested parties may not lead to an optimal solution. In the above example, if there are a number of individual fishermen rather than one fishing company who wish to use a stream being polluted by a chemical plant, the possibility of optimal bargain may fail to materialize. The transaction costs involved in achieving an optimal solution between interested parties may well be too high in this instance. Second, if one company

⁹For a useful summary of this cap-and-trade approach to reducing carbon emissions, see <http://people.ucsc.edu/~nuclear/econ1/hotnews/Carbon%20markets.htm>.

¹⁰Ronald Coase, "The Problem of Social Cost," *Journal of Law and Economics*, 3 (October 1960), pp. 1–44. Coase received the Nobel Prize for his work on social cost as well as for his work (cited in Chapter 2) on the nature of the firm. The social cost article is widely viewed as creating, almost singlehandedly, the field of law and economics. Coase's article is long and complex. For an excellent brief summary, see George J. Stigler, *The Theory of Price*, 4th ed., New York: Macmillan, 1987.

engages in unfair bargaining, no reasonable agreement may be reached. Third, neither side may have complete information regarding the costs and benefits of arriving at the optimal solution.¹¹ Most basically, Coase forced economists and legal scholars to redirect their efforts toward a deeper understanding of property rights and the analysis of transactions costs as an impediment to the efficient internalization of externalities.

STABILIZATION OF THE AGGREGATE ECONOMY: MONETARY AND FISCAL POLICY

Monetary and fiscal policy are mentioned previously as government activities designed to stabilize the economy. In light of the “Great Recession” which began in 2008 and the difficult time that the American economy has had in making a full recovery, this function has become a much more important part of the role of government in the market economy. We discuss each of these briefly.

Monetary Policy

Monetary policy is carried out by the Federal Reserve System (the Fed) through its control of the quantity of money in the economic system and/or interest rates. It is directed at attaining certain goals such as economic growth or price stability. Over the years, the specific methods used by the Fed have differed. At one time, its primary concern was the quantity of money. More recently, monetary policy has concentrated on influencing interest rates, specifically the Federal Funds rate.

Banks are required to keep certain cash reserves with the Fed. On any given day, some banks' reserves may exceed requirements. Federal Funds is the name given to these excess reserves, which can be borrowed by banks whose reserve balances are deficient. Most of these loans are extremely short term, mostly overnight. It is the rate on these loans that the Federal Reserve targets. Depending on the condition of the economy, the Federal Funds rate target can be lowered, raised, or left alone. This determination is made by the Fed's Open Market Committee, which meets eight times each year.

Fiscal Policy

The U.S. government collects taxes to finance expenditures for goods and services, as well as transfer payments that have been legislated by the U.S. Congress. Such expenditures are generally based on long-term goals. However, fiscal policy—the changes in the level of taxation and spending—is designed to achieve macroeconomic goals relating to output (gross domestic product) and employment.

By manipulating receipts or expenditures, a surplus or deficit is created. To stimulate the economy, the government—through Congress—will either decrease taxes or increase expenditures (or both) and create a deficit in the federal budget. Although both actions may stimulate an economy that has descended into a recession, the result of an increase in spending is to enlarge the role of the U.S. government in the economy, which may be considered unfavorable by the part of the population that believes in limited government. However, another part of the population looks at such actions favorably, believing that increased government actions may be beneficial. Again, the actions in the year 2001 present us with good examples of changes in fiscal policy.

One of the new administration's primary goals was to lower the level of taxes in the long run, thus decreasing the government's role in our lives. During 2001, Congress passed a law that decreased income taxes gradually over the next 10 years (other

¹¹A. M. Polinsky, “Controlling Externalities and Protecting Entitlements: Property Right, Liability Rule, and Tax-Subsidy Approaches,” *Journal of Legal Studies*, 8, 1979, pp. 1–48.

provisions were also included, such as the phasing out of the estate tax). However, because the economy appeared to be turning downward, Congress included a provision that gave immediate refunds of up to \$600 to most taxpayers. This was, of course, intended to give the economy a quick shot in the arm. As the economy appeared to continue its decline toward the end of the year, the administration and Congress embarked on trying to pass an “economic stimulus” package composed of certain tax cuts and spending increases. It was passed by Congress and signed by the president in March 2002. Congress enacted further tax cuts in the Jobs and Growth Tax Relief Reconciliation Act of 2003.

Lags

Although both policies can affect the movement of the economy, it must be recognized that in both cases the effect is not immediate. In fact, there are significant lags in the consequences of both monetary and fiscal actions. First, there is usually a delay in recognizing that a problem exists. Although both the Fed and the U.S. government have a large amount of economic data available, they may not be timely, or different data may give conflicting signals. (As already mentioned in Chapter 5, the onset of the most recent recession was dated March 2001, but this was not announced until December 2001.) Second, after the problem is recognized, there is a delay in implementation. This interval is particularly present in fiscal policy where the action must go through a political process that may be excruciatingly slow. The action can be realized more quickly in the case of monetary policy. Third, even when action is taken, there is a substantial lag for it to work itself into the economy and become effective.

Various economists have tried to estimate these lags. Although there is no agreement on the precise size of the lag, we know that it can last more than 6 months and may at times be as long as 1 year or more. Thus, it is quite possible that either monetary or fiscal action will actually turn out to be counterproductive. It may become effective when the economy has actually righted itself on its own, and the policy action may either add fuel to an already growing economy or aggravate a decline. The stimulus package passed in 2002 may be just such a case. Even without it, the economy appeared to have bottomed out in early 2002, and any stimulus that it may provide could be too late and could become effective at a time when the economy no longer needs it.

SUBPRIME LOAN FINANCIAL CRISIS OF 2007 TO 2009¹²

Background

A good example of the impact that government regulation (and deregulation) can have on the macroeconomic environment in which companies do business can be found in the subprime loan financial crisis of 2007 to 2009. Therefore, we devote this section to a discussion of the role of the federal government in this crisis.

During the 1980s, the gap between low- and high-income earners in the United States began to widen, and the real income levels of the low-income group fell. Congress passed legislation that permitted easier mortgage lending terms to be accepted on subprime loans (loans to people considered below prime or creditworthy borrowers). Mortgage loans to low-income people would help raise their standard of living by allowing easier conditions on mortgage loans made to them. The loans did not require income verification, did not require home appraisals, and allowed very low teaser rates during the first 1 to 5 years of the loan. The result was an increase in the normal amount of mortgage loans outstanding, adding about 20 percent to total mortgage

¹²We are grateful to Dr. F. John Mathis, Professor of Global Economics and Finance and Director of the Global Financial Services Office, Thunderbird School of Global Management, for writing this section.

loans. Because subprime mortgage loans had not been made before, a default rate history on this liability did not exist.

While this new subprime loan demand for homes was growing, upward pressure on home prices grew because interest rates were low and people began investing in second homes. They would then rent them out or renovate them, and after a period of two years (to lower their capital gain tax rate by meeting the requirements for long-term capital gains), sell them. The home price bubble continued to expand until the Federal Reserve Bank began to raise interest rates and reduce liquidity, which had previously increased significantly.

On February 1, 2006, President Bush appointed Dr. Ben Bernanke to a fourteen-year term as a member of the Federal Reserve Board of Governors, and to a four-year term as chairman. Bernanke was very concerned about the potential for inflation because of all the liquidity in the economy and because of rising home prices and stock market prices. As Bernanke began raising interest rates, liquidity began to become less readily available, which in turn curtailed the demand for housing. As home prices peaked and began to decline rapidly, owners tried to sell their homes. It was not long before many home prices had fallen so much that the prices of the homes were below the amount of their respective mortgages. This was a problem for the subprime home buyers and also for regular home buyers. To make matters worse, many of the subprime home buyers and home speculators had financed their purchases with very low teaser interest rates; as the interest rates began to rise, borrowers whose homes were worth less than their mortgages could not sell their homes and were not able to carry the sharply increasing monthly mortgage payments. Defaults on the loans began, and the share of mortgages in foreclosure grew rapidly. Many states, particularly California, Arizona, Nevada, and Florida, experienced 50 percent to 70 percent declines in housing prices accompanied by massive foreclosures.

Securitization of Mortgaged-Backed Securities

This new potential for expanding mortgage loans to low-income lenders resulted in a rapid increase in the expansion of loan origination mortgage brokers, many of whom were limited liability companies of major mortgage lenders. Unlicensed and untrained mortgage brokers and lenders recognized a great opportunity for making substantial amounts of money depending on their ability to write and process a large number of loans each day. These loans were then sold to banks that had a securitization operation at discount, and in a few days, the loans were converted into a cash payment to the mortgage brokers or some other originating and underwriting entity. This is exactly what happened to regular mortgages for years, but subprime loans were different because there were no historical data on their default record. Therefore, there was no default probability that the banks could use to set aside reserves or price the risk of default. Many subprime mortgage loans were particularly vulnerable because the home buyer did not understand the risks associated with adjustable rate mortgages and because these low-income families were often the first hit by layoffs as the economy slowed down.

Securitization is the process of executing a "true sale" of the loan to a bank, which then converts the associated cash flow (principal and interest) into an investment asset according to the investor's specifications (maturity, credit rating, etc.). It is credit-enhanced with an insurance policy to protect the investor from default, and it has several financial reserve pools set aside to protect the investor from being adversely affected by one or several mortgage defaults. The advantage to a bank for participating in securitization is that it freed up the bank's capital to make more loans by getting loans off its balance sheet. It should also be mentioned that all sorts of cash flows from auto loans, student loans, and so on, have been securitized. Securitization also supported the expansion of liquidity beyond what the Federal Reserve would have sanctioned because the Fed did not have

good data on the total amount of securitization in which banks had participated. For this reason, securitization has been referred to as a shadow money supply.

Another development related to securitization added to the financial risk faced by the United States and other countries: Increasingly complex securitization practices and associated financial derivative products were developed to expand liquidity and provide risk protection of investors that were mostly pension funds, insurance companies, and hedge funds. Among the best known derivative products were collateralized debt obligations (CDOs) and credit default swaps (CDSs). Even more complex instruments were developed in an attempt to reduce the correlation and default risk of the underlying assets, which were known as CDO² and CDO³. These instruments involved different ways of shuffling the pool of assets in an attempt to reduce the amount of correlation risk of the underlying pool of assets. These activities also earned the financial institutions additional fees, which were believed to help reduce systemic risk. For example, if a pool of mortgages came from one low-income neighborhood, they were all vulnerable to the possibility of a collapse of a company that employed most of the people in that community. So the more communities, cities, and states, and the more financial products (mortgages, auto loans, student loans, credit card receivables, etc.) that could be added to the underlying pool of assets, the greater the diversification and the less exposure to risk from a single event resulting in defaults. All of this is true, but it does not protect against the risk of the entire economy going into a recession and nearing depression. When this happens, the correlation coefficient among all of the underlying assets moves toward one, and the entire economy is adversely affected, which is exactly what happened in 2008 and 2009.

Disappearing Liquidity Challenges the Financial System

Several suddenly emerging conditions threatened the financial system in the United States and the global economy. Banks that had been involved in securitization of subprime loans now faced the loss of their reputation if defaults occurred on the instruments that had been sold to investors based on securitization, even though these banks were supposed to have been protected from bankruptcy. Consequently, banks put back these securitizations on their balance sheets and recognized their responsibility and liability to support the implied cash flow to the investors. The result was that not all banks that had participated in the securitization process were in strong financial condition because some were more leveraged than others. Bear Stearns and Lehman Brothers were the weakest links among the major financial institutions. At this point in early to mid-2007, the U.S. Federal Reserve and U.S. Treasury had to make a decision about which banks they would save or let fail. Bear Stearns was the first to fail: It could not get access to sufficient liquidity to fund its operations, and no other bank was interested in acquiring it. Its stock price fell and the value of its capital declined; other major banks would not lend it overnight funds. At this point, the financial system in the United States stopped functioning and the velocity of money approached zero (the velocity of money is the rate at which the money supply turns over in the economy during a 1-year period). Banks no longer trusted one another and therefore would not lend overnight to each other. Once banks could not get the amount of funding they needed to support the loans they had made, and if they did not have sufficient capital to draw down and fund the loans, they went out of business.

To avoid a U.S. and global financial crisis, the Federal Reserve and U.S. Treasury injected liquidity into the U.S. and international economy. They met with all of the major banks in the United States and increased the banks' capital in return for preferred stock in the banks. The banks would have to buy back this preferred stock at a later date and free themselves from government control of their operations.

The negative impact on U.S. economic activity did not stop with the banks. Many of the securitized loans had been purchased by Fannie Mae and Freddie Mac, which are government sponsored enterprises (GSEs), and investors to whom these GSEs had sold investment assets were now in danger of not being paid as the delinquency rate on subprime loans jumped to over 25 percent by 2008. Consequently, the U.S. government was also required to recapitalize these institutions to prevent their failure. Insurance companies that had provided credit enhancement insurance as protection against default were also hit with major losses and faced bankruptcy; some of these companies failed. The U.S. government considered one major insurer, AIG, so critical to global finance that it was rescued with sizable cash injections. (AIG had been very active in not only providing credit enhancements to securitization, but also in issuing credit default swaps.) Countrywide was a major mortgage lender that was considered one of the predatory lenders to subprime borrowers; the U.S. government negotiated its closure and sale to Bank of America. But many smaller predatory mortgage lenders declared bankruptcy and were later investigated by the Justice Department for fraud.

As of 2012, the U.S. economy is still struggling to recover from the deepest recession since the Great Depression. Real growth in the gross domestic product (GDP) has been averaging below 2 percent (compared to a historic average around 3.5 percent), and consequently unemployment remains around 8 percent. The unemployment rate would be much higher except that an estimated 5 million workers have dropped out of the labor force because they could not find a job at their skill level. The U.S. banking system is still struggling to work off bad loans and deleverage, and it has not become a significant provider of new lending to the struggling economy. The result has been continued dependence on fiscal policy to keep the economy growing, which in turn has resulted in a U.S. national debt that has more than doubled and in continuous battles in Congress between the Democrats and Republicans. The Democrats and President Obama want to increase the role of government spending and thus the U.S. debt level. The Republicans and Republican presidential candidate Mitt Romney favor cutting spending and reducing what they believe to be an unsustainable national debt level.

Changing Bank Regulations to Avoid Future Crises

While the subprime loan problem was the immediate global financial crisis trigger, many analysts identify other changes in the banking system prior to the crisis as contributing factors. Some of these structural changes in the banking system include the following changes in bank regulations. In 1999, Congress passed the Gramm-Leach-Bliley Act, which allowed banks, in addition to conducting traditional banking, to establish holding companies under which they could engage in investment banking; provide insurance; and undertake more leveraged or risky types of activities, including proprietary trading, and hedge fund and alternative asset investment operations. During this same period, the banking system expanded to account for 8 percent of GDP, up from 4 percent in the 1980s. And the chairman of the Federal Reserve Board of Governors, Alan Greenspan, used neo-Keynesian macroeconomic quantitative theory of economic activity to manage monetary policy. However, during testimony about the financial crisis and the rapid emergence of asset price bubbles, he admitted that these models did not pay enough attention to the “behavioral” side of economics. Thus, they did not warn that the excessive growth in liquidity that the Fed was providing would set the stage for the economic crisis. Low interest rates and narrow risk differentials between financial assets were not accurately pricing the degree of financial risk that was rising in the asset price bubbles in real estate and in stock prices.

Thus, while some regulatory changes in the banking system may be appropriate, the timing of the Dodd–Frank Wall Street Reform and Consumer Protection Act,

signed into law by President Obama in July 2010, could not be more ill-timed in terms of hindering the U.S. economic recovery from the crisis. Many of the restrictions that would be placed on U.S. banks would adversely affect the competitiveness of the U.S. financial system. The legislation does support the raising of bank liquidity requirements and how liquidity should be measured by the Bank for International Settlements (located in Basel, Switzerland) and raising the level of capital required to back more risky types of lending activities.

Global Financial Deregulation and Growth in International Capital Flows

During the period from 2002 to 2006, when the U.S. economy was rapidly expanding and the subprime loan crisis was growing, U.S. imports also grew much faster than U.S. exports. Thus, the U.S. current account showed a series of record deficits exceeding 6 percent of GDP, or almost \$850 billion. The total increase in U.S. dollar liquidity was in excess of \$3 trillion from 2002 to 2006. For at least a decade earlier, developing countries had been liberalizing their financial markets to attract more capital inflows, first in the form of foreign direct investment and later in the form of foreign portfolio investments as their stock markets became more developed and accessible to foreign financial inflows. This resulted in a huge savings glut in countries in Asia and Latin America, which is typically not the situation in emerging markets. A substantial portion of these dollar financial flows out of the United States actually ended up back in the United States in the form of foreign purchases of U.S. government securities or investments in securitized investments and equity markets. This increase in liquidity in the United States helped to keep interest rates low and to fund the expansion of the shadow banking system resulting from the securitization of mortgages, which was undetected by the U.S. Federal Reserve and thus kept the money supply expanding to support domestic growth.

The result of this increased level of global finance was the transmission of the financial crisis to both advanced and developing countries. For example, when the United States suffered from the subprime loan crisis, so did investors in Europe. In an attempt to prevent the great recession in the United States from transmitting negative growth and employment effects, European countries borrowed heavily, primarily in short-term financial markets at low interest rates, to stimulate economic growth in their home economies. This helped many countries avoid suffering a prolonged serious recession like the one the U.S. economy had experienced. However, when the world economy did not immediately recover and these increased foreign borrowings had to be repaid, growth was insufficient to generate a large enough government surplus to repay the increased debt serving obligations. Thus, the less developed countries of Europe, which rely more heavily on increased export earnings to repay foreign borrowing, ended up in their own financial crises. These crises required a bailout by the International Monetary Fund (IMF) and the European Central Bank, the latter supported by the German government. This European debt problem continues to block global recovery and is forcing harsh government cuts in spending to prevent further budget crises and a potential breakup of the euro common currency area.

GOVERNMENT DEREGULATION, MERGERS, AND ACQUISITIONS

For various historic, political, and economic reasons, certain major industries in the United States were subject to considerable government regulation. Telecommunications, electric and gas utilities, airlines, and commercial banks are perhaps the best examples. Beginning in the late 1970s and continuing on through the 1990s, the U.S. government in effect eliminated most if not all regulatory control that it had exercised

over firms in these industries. Regardless of the original reasons for their regulation, the government has assumed that by deregulating these industries, consumers will be better served by lower prices, better service, or more rapid introduction of technology that would ostensibly stem from a more competitive environment.

However, in recent years, the deregulation has apparently resulted in such a fiercely competitive environment that a number of companies in these industries have sought to merge or acquire other companies in order to survive and grow. This is somewhat ironic because one of the great concerns of antitrust policy in the United States has been the level of merger activity. As companies merge, is there a danger that their greater size makes them monopolistic and a threat to free competition? If so, then the very actions of the government to make certain industries more competitive may require it to assert its role as the arbiter of what constitutes "competition."

In commercial banking, the mergers (the combination of two firms or the acquisition of one firm by another) that have grabbed most of the headlines have been among the leading financial institutions. Citicorp merged with Traveler's Insurance to form Citigroup. Chemical Bank and Chase Manhattan merged to become the Chase Bank, which then merged with JP Morgan to become JP Morgan Chase. Bank of America merged with Nations Bank, and then later acquired Fleet Bank. In 2004, Bank One (which had earlier merged with First Chicago) merged with JP Morgan Chase.

The government's hand in regulating, deregulating, and then regulating an industry again can be seen in the case of the U.S. telecommunications industry. For many years, AT&T was a publicly traded company that was owned by shareholders, just like any other corporation. But it was also subject to government regulation regarding its services and its pricing of services. In other words, it was known as a regulated monopoly. This was not the model followed by most other countries in the world. For example, the old British Telecom (BT) was actually part of the Royal Mail Service. Then, in the 1980s, there was a movement to deregulate or privatize the telecommunications industries around the world. In the case of the United States, in 1984, AT&T was split into seven regional companies, each responsible mainly for telecommunications service within its region. Meanwhile AT&T kept the long-distance service (i.e., the service among the different regions).

Over the years, however, the seven regional companies gradually began to merge with one another. Today's Verizon is a result of the merging of Nynex, Bell Atlantic, MCI, and GTE. Today's (new) AT&T is a result of the merger of SBC, Pacific Telesis, Ameritech, Bell South, and the old AT&T. The old AT&T was added to this group, mainly for its brand name. Because the telecommunications industry was deregulated, the federal government did not move to block any of these mergers. In 2011, AT&T announced its intention of buying T-Mobile, a subsidiary of the German company Deutsche Telecom. But the federal government once again stepped in by blocking AT&T's move to buy T-Mobile.¹³

The airline industry in the United States also has experienced some merger and acquisition activity, although not quite as much as in the financial and telecommunications industries. In 2000, United Airlines tried to buy U.S. Airways but was prevented from doing so by the government. However, in 2001, the government allowed American Airlines to acquire TWA. After the approval of this acquisition, two additional mergers occurred. In 2006, U.S. Airways merged with America West. In 2008, Delta joined forces with Northwest. In 2012, United Airlines merged with Continental Airlines.

¹³http://en.wikipedia.org/wiki/Attempted_purchase_of_T-Mobile_USA_by_AT%26T.

Why Firms Merge

The recent increase in mergers in the deregulated industries is nothing new. The U.S. economy has passed through a number of merger waves since the end of the nineteenth century. During the 1980s and early 1990s, many mergers were motivated by low valuations of the target firms, which had been underperforming. In addition to mergers, this period also saw a large number of leveraged buyouts (LBOs), where public companies were taken private by groups of investors, using large amounts of debt to finance the transaction.¹⁴ In the mid- and late 1990s, mergers have been motivated primarily by the necessity to obtain greater efficiencies—economies of scale or scope—to be able to compete in a global economy. This impetus for merging with or acquiring companies is relevant for any industry where the intensity of competition has increased.

Besides cost synergies, companies acquire or merge for revenue synergies. Experience shows, however, that it is easier for consolidations to reduce costs than to increase revenues on a sustained basis. One of the most well-known examples of this point was the purchase by Daimler Benz of Chrysler back in 1998. The German company bought the U.S. company to fill in its product line at the lower end of the market. This plan did not work out, and Daimler Benz had to sell Chrysler to a private equity group nine years later. It bought Chrysler for \$37 billion and ended up selling it for about \$7.4 billion. After all the terms and conditions for the deal were satisfied, Daimler ended up losing about \$650 million.¹⁵

Regardless of the specific reasons for a merger or acquisition, the basic motivation for mergers is to increase the value of the combined firms compared with their separate valuations. This simple idea is usually expressed with the following equation:

$$V_{A+B} > (V_A + V_B)$$

where V stands for total market value, and A and B are the two companies involved in the merger.

Among the incentives to merge, some result in increased economic efficiency, but others do not. A partial list of incentives follows:

1. *Synergies in production.* If synergies exist, then the value of the combined companies should exceed the value of the two separately. Among the synergistic results, we would find the following:
 - a. *Revenue enhancements.* For example, a better distribution system for products when two companies combine may increase sales.
 - b. *Operating economies.* These results would include economies of scale and/or scope. Economies could also be achieved through improved research and development (because of complementarities in technical skills resident in the two companies) or “management meshing,” where the skills of the management of the companies complement each other. For instance, one company has strong marketing management while the other’s managers have technical superiority.
 - c. *Financial economies.* The combined company may be able to lower its cost of capital.
2. *Improved management.* In some cases, the acquired firm may lack good management skills, while the acquiring firm has a relative abundance of skilled managers. The merger will create an opportunity for improving the overall management level of the new company by eliminating poor managers.

¹⁴The ultimate goal was, of course, to streamline these companies, make them more efficient (possibly by disposing of some parts of the company), and eventually bring them back as publicly owned companies.

¹⁵http://money.cnn.com/2007/05/14/news/companies/chrysler_sale/index.htm.

3. *Tax consequences.* Although a merger may not result in increased economic efficiency, it may reduce the tax bill of the two combined companies. If a company has been incurring losses, it will not be paying taxes. It will carry tax losses forward and may reduce its tax liabilities in the future when it becomes profitable. By merging with a profitable company, it will be able to save taxes immediately and increase the combined companies' cash flow, even in the absence of any synergies.

Taxes will also be decreased for a company with much cash and no great investment opportunities. If this company were to pay a large dividend, then its stockholders would pay taxes immediately; or if the company were to buy back its stock, stockholders could be liable for capital gains taxes. By using its cash to acquire another company, it will avoid creating tax liabilities.

4. *Managerial power.* Mergers may occur when the acquiring company's managers are seeking to increase their span of authority. Although the acquisition of another company may expand the power of managers, it will, in many cases, not result in enhanced efficiency.
5. *Diversification.* During the 1960s, diversification was the ostensible motivation for a large number of mergers. Mergers among companies in unrelated fields of activity predominated. It was said that diversification could decrease the variability of sales and earnings, and thus be of benefit to stockholders, even if no synergistic effects were present. This is a flawed argument. There is no reason why stockholders cannot achieve their own diversification by investing in both companies. This would be accomplished much more cheaply because it would avoid the significant costs that would be incurred in completing the merger.
6. *Market power.* The combination of two or more powerful firms could lead to a decrease of competition in the industry. The result could be lower production, higher prices, and a negative effect on the efficient functioning of the economy. It is with these potential effects that U.S. antitrust laws are concerned.

A large number of studies have investigated the effects of mergers on stockholders and the economy. We summarize their results as follows:

1. There is general agreement that stockholders of the target companies are the big winners, gaining between 20 and 30 percent when their company is acquired. In contrast, the stockholders of the acquiring firms gain very little because their stock prices, on average, remain constant. Overall, there appears to be an increase in the value of the combined companies.
2. The evidence regarding increased profitability of merged firms is rather mixed.
3. Merger activity does not appear to have increased the level of industry concentration.
4. There appears to be no decrease in research and development activity of merged firms, contrary to the opinion of some commentators.¹⁶

Two recent studies have generally reinforced the summary of findings above. One, by Robert F. Bruner, surveyed over 120 studies that have been conducted over the last three decades.¹⁷ Essentially, Bruner examined the following outcomes: value conserved, value created and value destroyed, using three forms of measurement:

- **Weak form:** Did share prices rise?
- **Semi-strong:** Did the firm's returns exceed a benchmark (such as S&P 500 Index)?
- **Strong:** Are shareholders better off after the deal than they would have been if the deal had not occurred?

¹⁶For an excellent review of recent mergers and acquisitions, including those in industries that we discuss in this chapter, see "Oligopolies Are on the Rise as the Urge to Merge Grows," *Wall Street Journal*, February 25, 2002.

¹⁷Robert F. Bruner, *Applied Mergers and Acquisitions*, John Wiley, 2004, Chapter 3.

While the strong form would have been the most valid measurement, such results are not observable. Thus, the semi-strong form is used in summarizing results. The result of a large number of market-based return studies confirmed that the returns to stockholders of target firms are significantly positive. However, when the results for the acquirer are investigated, the results are not clear. Bruner summarized at least fifty studies and concluded that “abnormal returns to buyer shareholders from M&A activity are essentially zero,” and that buyers “essentially break even.” When the returns to the buyer and target firms are combined, most of the studies report positive returns.

However, Bruner points out that, from an economist's viewpoint, breaking even does not represent failure. Even though no new value is created, such mergers have a present value of zero, which means that stockholders are earning at least their opportunity cost of capital. An economist would consider failure only if value is destroyed.

Several studies investigated longer term effects. A majority of these studies reported negative returns. However, in these cases there is a possibility that the outcomes are not caused by the transaction itself but that there are other reasons for these results. Bruner mentions two possible plausible explanations:

- The buyer's shares were overvalued at the time of the merger.
- Industry conditions after the merger have influenced results.

The studies that investigated the results of mergers have identified a number of factors that are influential in determining the success (or lack of success) of mergers and acquisitions. Among those that are instrumental in enhancing value are:

- Expected synergies
- Mergers that look for value
- Restructuring that includes divestitures of underperforming businesses
- Tender offers (as compared to friendly mergers)

Those that do not create value include:

- Glamour acquisitions (based on book-to-market ratios)
- Mergers to build market power
- Mergers to use excess cash

The other study examined the performance of over 12,000 publicly listed U.S. acquiring firms during the 1990s.¹⁸ These acquiring firms appear to have outperformed the market by 50 percent. While most prior studies have analyzed the performance of companies after acquisition, this paper appears to find that companies that undertake acquisitions show a significant stock price increase in the year before the acquisitions. These increases are greater when the acquisition target is a nonpublic rather than a public company. Also, the pre-announcement run-up is greater when stock rather than cash is used to complete the acquisition. After the acquisition is completed, stock prices tend to decline; however, this decrease does not erase the pre-announcement increases. The authors find that the findings of previous studies that “targets gain and acquirers break even” are too simplistic. Most acquisitions appear to bring about positive gains. The only category that confirms the previous findings is the acquisition with stock of large, public target companies.

¹⁸Michael Bradley and Anant Sundaram, “Do Acquisitions Drive Performance or Does Performance Drive Acquisitions?” Paper presented at a seminar at Thunderbird, the Garvin School of International Management, in October 2004. An earlier version was presented at the 2004 European Finance Association meeting in Maastricht, Netherlands.

GOVERNMENT PROTECTION OF INTELLECTUAL PROPERTY (IP)¹⁹

At the beginning of this chapter, we listed the five major functions of the government in a market economy. These functions essentially involve government intervention in the market economy to redistribute or reallocate resources, stabilize the macroeconomy, and ensure that market power is not dominated by monopolistic power. Some or all of these activities might be looked at as ways that might possibly hinder or limit business activity. However, government can perform another function, in addition to the five mentioned previously, that is clearly intended to expand rather than limit commercial enterprises. This function is the protection and safeguarding of private property. We use the term *property* not only in reference to real estate. We use it here in the more general sense of being any financial asset that is intended to generate future revenue and profit for a firm. Property can include tangible assets such as factories, office buildings, private residences, apartment buildings, machinery, equipment, and inventory. It can also include intangible assets such as patents and trademarks. These types of intangible assets are called intellectual property (IP).

In the United States, the government agency responsible for the issuing of patents and the granting of trademarks is the United States Patent and Trademark Office (USPTO). This office is an agency of the Department of Commerce. As stated on the homepage of its website, the USPTO's mission is "to ensure that the Intellectual Property system contributes to a strong global economy, encourages investment in innovation, and fosters entrepreneurial spirit. Intellectual Property is an invention or creation embodied in the form of a patent, trademark, trade secret, or copyright."²⁰ In today's global economy, IP (i.e., patents, trademarks, trade secrets, and copyrights) has become such an important class of assets that the term *knowledge-based economy* was coined in reference to this phenomenon.

To illustrate some of the rights attributed to intellectual property, let us begin by discussing the basic rights of those who own real property such as land, houses, and buildings. Consider a homeowner, who has the following rights: the right to keep everyone else out (no trespassing), the right to invite friends and family to visit or perhaps even use the property while the owner is away (free use), the right to rent a room of the home, several rooms, or even the whole house (leasing), and the right to sell the house outright (transferring the rights to someone else). Within the confines of our legal system, there is the right to prosecute those who have damaged or stolen one's property.

Intellectual property has the same types of rights attributed to real property. For example, a patent provides a "no trespassing" right. Formally speaking, a patent is a property right granted by the government to an owner to exclude others from practicing the claimed invention in exchange for donating the invention to the public after a given period of time. The patent owner has the right to exclude others from using the intellectual capital described in the claims of the patent (usually up to about 20 years from the date on which a patent application was filed). The patent claims are analogous to a fence around one's house. The fence clearly defines the property boundary and makes a nice place to hang a No Trespassing sign. What happens after the 20 years is up? The no trespassing rights go away. What kind of property legally allows trespassing? Public property. A public park is a good example. The park is free for anyone to use. Most important, no one can claim any portion of the park to be his

¹⁹The idea for including this section, as well as most of its content, was provided by Riyon Harding, Executive IP Strategist at a major global information technology (IT) company who advises other companies on matters relating to the commercialization of their intellectual property. Reprinted with permission from Riyon Harding and IBM.

²⁰http://www.uspto.gov/about/stratplan/ar/2004/0401_mission_org.jsp.

or her own private property or exclude anyone else from enjoying the amenities the park has to offer. So when the rights expire for a patent, the invention goes into what is called the public domain—or what could be called the park—free for anyone and everyone to use.

Copyrights and trademarks similarly provide “no trespassing” rights to their owners. Copyrights are typically enforceable for up to 70 years after the author’s death. In some cases copyrights are one of the most important product lines a company has. Sony Publishing Company, for example, makes most of its income from royalties and the distribution of music for which it owns the copyrights, including rights to over 250 Beatles songs. The Disney Corporation is another example of a company that makes a significant amount of its bottom line from “renting” its characters and movies.

Trademarks and service marks differ slightly from copyrights because they are typically not rented for income but rather are important aspects of building brand and goodwill. Trademarks last as long as the owner takes precautions to ensure that no one else is trespassing. The owner is also required to ensure that the name continues to designate its own products only and does not become a household name. The process known as genericide has befallen many companies, including Otis Elevator Company, which lost the trademark “escalator.” “Thermos” was originally a trademark of Thermos GmbH, and “zipper” was originally a trademark of B.F. Goodrich. Other companies have come dangerously close to losing their trademarks but have averted this by marketing efforts to change public perception. Examples include the trademarks for Xerox (photocopies), Lego (building blocks), and Kleenex (brand facial tissues). Currently, Google is one of the most recognized trademarks in the world, and the company has fought its own trademark battles. At one point in time, the *New Oxford Dictionary* planned to incorporate the word *Google* as a verb, meaning “search the Internet.” Google convinced the editors to use lowercase *g* (*google*), thereby preserving its uppercase trademark.

Trade secrets comprise the fourth category of IP. Although, they have no “no trespassing” rights per se, they still offer a company a kind of insurance policy. Trade secrets do not expire as long as the owner keeps them secret. Trade secrets are especially useful for products a company is planning to sell for more than 20 years. Imagine if the Coca-Cola Company had not patented its formula when it started producing the drink in the late 1800s. It would have meant that anyone could have been making and selling the exact Coca-Cola formula for the last 100 or so years! Here is where the insurance policy comes in. Suppose Coca-Cola somehow lost its trade secret either by innocent or devious means. Coca-Cola may still be able to take the offending party to court and thereby force the offending party to cease and desist. It may even be possible to sue the offending party for damages based on loss of business. Courts ultimately rule whether something is afforded trade secret status. The burden is on the trade secret holder to prove that it took reasonable steps to ensure the product’s safety.

In a knowledge-based economy, IP can be considered a separate product whose value is based on supply and demand, just like any other tangible product or service. For example, in August 2011, Google purchased Motorola Mobility for \$12.5 billion. Google’s CEO stated that one of the major reasons for this purchase was to build up Google’s patent portfolio, thereby helping the company to “better protect Android [Google’s mobile operating system platform] from anti-competitive threats from Microsoft, Apple and other companies.”²¹

About a month prior to Google’s announcement of its intent to purchase Motorola Mobility, a consortium of top technology companies led by Apple, Microsoft, and

²¹<http://techcrunch.com/2011/08/15/breaking-google-buys-motorola-for-12-5-billion/>.

RIM announced its intention of buying the patent portfolio of Nortel for \$4.5 billion. This was about five times the \$900 million that Google was prepared to pay in the bidding process leading up to the final sale price.²²

GLOBAL APPLICATION: THE FAILED ATTEMPT TO MERGE BY GENERAL ELECTRIC AND HONEYWELL

The failed attempt by General Electric to merge with Honeywell in the first half of 2001 provides an excellent example of the international dimension of the role of government in a market economy. What it really illustrates is two opposite philosophies of antitrust held by American and European regulators and government officials. In October 2000, the General Electric Company proposed to merge with Honeywell International, Inc. Prior to this move, Honeywell was in the process of merging with the United Technologies Corporation. After GE topped United's bid, Honeywell and GE prepared for their merger.

The U.S. Department of Justice approved the merger in May 2001. During the past 20 years, the American interpretation of the antitrust laws has been moving in the direction of making decisions on the basis of a merger's effect on consumer welfare. The GE–Honeywell merger appeared to promise production and distribution efficiencies. However, because the merger would affect business in Europe, it had to be approved by the European Commission. After investigation and conferences during the first half of 2001, the European Commission ruled against the merger, despite various concessions made by GE. The basis for the rejection was the merger's potential threat to competitors.

Two major issues appeared to influence the commission's decision:

1. GE and Honeywell, if combined, could “bundle” their products. GE is a major producer of aircraft engines, whereas Honeywell produces avionics and aerospace products. The merged company would now be able to offer discounts to “customers who bought a package of components and services from GE–Honeywell.”²³ Thus, the merged company's competitors would be injured. Several competing companies actively opposed the merger; among them were United Technologies, Rolls-Royce, and Rockwell.
2. GE Capital Aviation Services (GECAS) is one of the world's largest buyers of commercial aircraft, which it sells or leases to airlines. The commission feared that GECAS could require aircraft manufacturers to install GE–Honeywell products on the aircraft ordered by it.²⁴

Based on this case, we can say that U.S. government officials appear to favor the demand side of the market (i.e., the buyers of aerospace and related products), whereas their European counterparts tend to favor the supply side (i.e., the firms that produce these products). We think that the concern about the loss in jobs that inevitably follows such mergers is probably what tilts the Europeans toward the supply side.

²²<http://www.bloomberg.com/news/2011-07-01/nortel-sells-patent-portfolio-for-4-5-billion-to-group.html>.

²³Andrew Hill and Deborah Hargreaves, “How Monti Turned GE–Honeywell into a Flight of Fancy,” *Financial Times*, July 6, 2001.

²⁴In addition to the article quoted in footnote 24, several other articles appeared in the press on this subject. Among these were Mike France, “Europe: a Different Take on Antitrust,” *BusinessWeek*, June 25, 2001, p. 40; Gary Becker, “What U.S. Courts Could Teach Europe's Trustbuster,” *BusinessWeek*, August 6, 2001, p. 20; Bryan M. Carney, “Blame the EU's Antitrust Rules—Not Monti,” *Wall Street Journal*, July 7, 2001; Philip Shishkin, “Barred Merger Signals U.S.–EU Divergence,” *Wall Street Journal*, July 5, 2001.

The Solution

Several weeks after the executive retreat, Allison and her team presented their findings to the C-Level committee.

Allison opened by saying, "Essentially, we found a lot of data supporting Bob's view that consumers are starting to prefer healthier drinks rather than carbonated soft drinks. We're not going to bore you with PowerPoint slides showing everything we found. We have a short executive summary for anyone who is interesting in the numbers. However, one thing makes us hesitant to support Bob's proposal. Have you noticed what has happened to PepsiCo recently? Over the past five or six years, their senior leadership tried moving more into healthier foods and drinks. Now they're switching back to their old strategy of focusing on Pepsi-Cola."

"Why?" asked Arun Mehta, the CFO.

"You would really appreciate this, Arun." Allison responded. "It's a simple matter of profit and share price. Over the past five or six years, Coca-Cola's shares have gone way up, while PepsiCo's shares have barely increased. Also, PepsiCo's return on capital has decreased over this time. And I believe these financial results start with a simple marketing statistic: Diet Coke has replaced Pepsi in the number 2 spot in total U.S. sales. Coca-Cola now has the Number 1 and 2 best-selling carbonated soft drinks in the country."²⁵

Then Bob chimed in. "Maybe we can make an offer for the healthy food and beverage division of PepsiCo. In that way, we can grow our healthier beverage and food business by acquisition rather than through organic growth."

Of course, Samantha, the CLC, quickly jumped in. "I'm not so sure. There may be antitrust issues involved."

"Don't worry, I was only trying to introduce a bit of humor. We've got to figure out how we can achieve profitable organic growth. There are too many risks in making a major acquisition. Look what happened to Quaker Oats and Snapple."²⁶

SUMMARY

We illustrate specifically how various business decisions can be influenced by government involvement in the market economy. As we discuss in Chapter 1, the primary advantage of the market process over the command and traditional processes is the efficient manner in which market participants allocate a country's scarce resources. Throughout this text, we try to show how managers, equipped with an understanding of the major factors of the market process (supply, demand, production, cost, and competition) and various quantitative tools of analysis are able to make optimal decisions to help their firms maximize economic profit.

However, managers must often take government involvement into account in the making of an optimal decision. This is particularly true when managers operate on a global basis and must deal with the laws and regulations of different governments. A case in point is the failed attempt by GE to merge with Honeywell discussed in this chapter's "Global Application." Government laws and regulations can reduce a firm's profits. But at the same time, as we show, the government itself is a major customer and so businesses can profit by being suppliers to the government's demand for various goods and services. Today's manager must be equally versed in matters of government and private industry.

IMPORTANT CONCEPTS

Benefit externalities. Benefits that accrue to individuals other than those who have paid for a particular good or services. Also referred to as *external economy*, *positive benefits*, *spillover benefits*, *third-party benefits*, and *social benefits*. The demand

for products with external benefits tends to be understated in the market. (p. 538)

Coase theorem. The idea, developed by Ronald Coase, that government intervention to eliminate the effect of externalities is not necessary if

²⁵A good summary of the actual situation facing PepsiCo in 2012 can be found at http://www.nytimes.com/2012/03/13/business/pepsico-executives-line-up-behind-ceo.html?pagewanted=2&_r=1. For an in-depth report, go to <http://management.fortune.cnn.com/2012/05/29/pepsi-indra-nooyi-2/>.

²⁶In 1994, Quaker Oats bought Snapple for \$1.7 billion. It failed badly with this product and ended up selling it to a private company for \$300 million. Ironically, PepsiCo bought Quaker Oats in 2001 mainly to gain ownership of Gatorade.

property rights are correctly and clearly defined and transactions costs are small. (p. 540)

Cost externalities. Costs incurred by individuals other than those who produce a particular good or service. Also referred to as *external diseconomy*, *negative costs*, *spillover costs*, and *social costs*. The supply of goods whose production involves cost externalities tends to be overstated in the market. A good example of cost externalities is environmental pollution. (p. 538)

Natural monopoly. An industry in which a single large firm can serve customers more efficiently than many smaller ones because of economies of scale. (p. 535)

Socially optimal price. The price of a good or service that is equal to its total marginal cost of production. (p. 539)

QUESTIONS

1. What is the rationale for government involvement in the market economy? (Cite the five points presented at the outset of this chapter.)
2. Define *benefit* and *cost externalities*. Explain why situations involving benefit externalities tend to result in an underallocation of society's scarce resources, and why situations involving cost externalities tend to result in an overallocation of society's scarce resources.
3. What is the role of government in dealing with benefit externalities? With cost externalities?
4. Suppose a chemical company was fined for violating certain antipollution laws. As the spokesperson for the Environmental Protection Agency, how would you explain the economic reasons for these actions to angry customers of this company who were forced to pay more for the chemicals as a result of this government action?
5. "The reason the government has to step in and 'internalize' benefit and cost externalities is because people are basically selfish." Do you agree with this statement? Explain.
6. Briefly discuss the Coase theorem. What does this theory imply about the role of government in dealing with market externalities?
7. Based on the competitive challenges that Microsoft faces from Google, do you think it still has "monopoly powers"? Explain.
8. Discuss the economic justification for a merger. In particular, how might these reasons apply to companies now merging in the following industries: oil, automobiles, telecommunications, electric power, and commercial banks?

The Global Soft Drink Industry¹

INTRODUCTION

Throughout the text, we have provided a number of hypothetical vignettes: Situations and Solutions involving our hypothetical Global Foods Company and its production of different types of soft drinks. In this concluding chapter, we illustrate how some of the key concepts in managerial economics can be applied to the actual global soft drinks industry. The major product categories of the soft drink industry are packaged water, carbonates, concentrates, 100% fruit juices, nectars, still drinks, ready-to-drink (RTD) tea and coffee, and energy and sports drinks. We start with key factors and trends in this industry that are affecting market demand. We then discuss the impact that these trends and factors have on supply. We close with a brief look at the energy drink market, the segment of the soft drink market that our hypothetical Global Foods Company decided to enter in Chapter 2.

The global soft drinks market had a total consumption of over 700 billion liters in 2011. The percentage of the total held by each product category is as follows: carbonated soft drinks (30.83 percent), sports and energy drinks (2.68 percent), RTD tea and coffee (5.96 percent), packaged water (25.51 percent), bulk water (17.11 percent), 100% juice (3.13 percent), nectars (2.36 percent), concentrates (2.63 percent), fruit powders (2.86 percent), and still drinks (6.94 percent). Coca-Cola claimed the first place in the global company ranking, while PepsiCo, Nestle, and Danone Group claimed the second to fourth positions respectively.²

FACTORS AND TRENDS INFLUENCING DEMAND

In Chapter 3, we discussed the key nonprice factors that affect the demand for a good or service: (1) tastes and preferences, (2) income, (3) number of buyers, (4) price of substitutes and complements, and (5) future expectations. Following are some of the major trends in the global beverage industry. As you can see, these trends pertain directly to many of the factors listed above.

1. **Shift in Demographics:** As of October 2012, two billion people over the age of 60 are living on this planet.³ More importantly, this age group is growing five times as quickly as the population as a whole.⁴ With some of the biggest markets for soft drinks, i.e. North America and Europe, being on the forefront of this trend, it is logical to expect a

¹This chapter was written by Farshad Sarmad. Farshad has over ten years of experience in operations management, project management, strategic marketing, and consulting within the food and beverage industry. He is currently the president of Bev Canada Enterprises Co., a product development and marketing consulting company, located in Vancouver, British Columbia. Farshad is also an adjunct faculty member at Fairleigh Dickinson University -Vancouver Campus, where he teaches courses in International Business and Business Ethics.

²Canadean, Global Soft Drinks Report, 2012. (<http://www.canadean.com>).

³Planet of the aged, 2012. *Discover*, 33, 38. Retrieved from ProQuest Database.

⁴Planet of the aged.

significant change in the dynamics of these markets in the coming years. Living longer could translate into a higher demand for healthier products that potentially maintain, or even improve the health state of the consumers. In addition, formation of a large senior consumer segment with a considerable disposable income could further boost the growth of premium products and services, particularly in the developed nations. Overall, shifts in demographics, i.e. the aging of population, are expected to play a key role in shaping the future of the beverage industry across the globe.

2. **Integration of Markets:** Research findings show ample evidence that the world of business is increasingly becoming global.⁵ Globalization has been accelerating the social, political, technological, and economic integration of societies.⁶ This trend, in effect, has had a profound impact on the way businesses and markets are being managed globally. Companies and customers alike have encountered challenges and opportunities that result from complexities of global integration. Consumers have enjoyed a wider range of product choices, while producers have gained from a larger market size as well as a greater cultural diversity in the demand for their offerings.

Food and beverage products have increasingly shown a relatively strong need for local responsiveness. Nevertheless, as the markets become ever more integrated, we see a rise in the level of cross-cultural availability of products in the global soft drinks industry. The integration of markets leads to integration of consumer trends, which could spur cross-category innovations. The emergence of Smoothies as a successful product category is an example of integrating different consumer trends in one category. Another emerging cross-category example is combination of juice drinks and energy drinks.

3. **Going Back to Basics:** Health, wellness, convenience, and premium status continue to be the main drivers of growth and global trends in the beverage industry. However, due to socio-economic changes across the globe, the way customers engage with these trends is expected to be different. For instance, simplicity and the *return-to-basics* approach to eating and drinking have begun to form an underlying trend in the beverage industry. This trend has been mostly apparent in the developed countries so far. Nevertheless, its impacts on various aspects of the business are expected to be felt more profoundly across the globe. To remain competitive, industry players including; manufacturers, distributors, bottlers, ingredient houses, and retailers, all would have to modify some of their strategic decisions throughout their value chain. Such modifications could, in effect translate into a need for transforming important strategic elements such as: packaging, label design, branding, product formulation, choice of product ingredients, and other marketing mix elements, in order to remain in alignment with the evolving market trends.
4. **Demand from Emerging Markets:** Economic projections suggest that the economic order of the world will change dramatically over the next few decades. By mid-century China is likely to be the biggest economy, followed by the United States and India.⁷ Unquestionably, the BRICs—Brazil, Russia, India, and China—will be the key to future growth of the beverage industry. Moreover, it is expected to see a shift of focus from highly saturated North American and Western European beverage markets, as well as the fast-maturing countries of Eastern Europe and Latin America, toward the untapped emerging markets of Asia, Middle East, and Africa. In 2009, consumption of branded commercial packaged beverages accounted for only one in every 10 litres of recommended liquid intake in Asia-Pacific, Middle East, and Africa, with a combined population of some 5 billion people.⁸ The BRIC countries as well as the untapped regions of Asia-Pacific, Middle East, and Africa present an immense growth potential for the beverage industry. Consequently, they are likely to become primary targets of multinationals in the next decade.

⁵A. Inkpen, and Ramaswamy, K., *Global Strategy: Creating and Sustaining Advantage Across Borders*, 2006. Oxford, UK: Oxford University Press.

⁶R. M. Hodgetts, Luthans, F., & Doh, J. P., *International Management-Culture, Strategy, and Behavior* (6th ed.), 2006. New York, NY: McGraw-Hill.

⁷J. D. Daniels, Radebaugh, L. H., & Sullivan, D. P., *International Business* (13 ed.), 2012. New Jersey: Prentice Hall.

⁸Euromonitor International, 2009. *How to Tap into the Untapped Potential in Drinks*.

FACTORS AND TRENDS INFLUENCING SUPPLY

As we noted in Chapter 3, the nonprice factors affecting the market supply of a product are: (1) costs and technology, (2) prices of substitute goods and services, (3) future expectations, (4) number of sellers, and (5) weather conditions. These factors can be seen to work best in perfectly competitive markets, where it is very difficult for a competing firm to differentiate its product from the other firms. However, the market for beverages is far from being perfect, in the economic sense of the term. It would be more accurately described as imperfect competition. As a global market, it might be considered monopolistic competition. In the market for specific types of products such as carbonated soft drinks, it could be described as an oligopoly. Even more specifically, if we consider the market for the cola carbonated soft drinks in the United States, then for all intents and purposes, it could be described as a duopoly. So when we look at the supply of beverages, we have to look at the supply of specific types of products rather than the general category of beverage. With this in mind, let us look at some key trends and factors.

- 1. Mass Customization:** Despite the fact that there has been a surge in the number and type of functionality claims of products in recent years, historically, the beverage industry has not been strong in terms of product differentiation on the basis of product *functionality*. This becomes even more evident by comparing common differentiating strategies of the beverage industry with another industry such as the IT. Soft drink producers have shown a preference toward creating a point of difference for their products based on aspirational attributes and *image*. Coca-Cola's "open happiness" slogan is a case in point for this type of strategy. Nevertheless, mass customization has started to play a more important role as a differentiating tool for soft drink manufacturers in recent years. Advocates of this strategy consider mass customization as a viable source of competitive advantage that allows companies to build strong brands by effectively engaging with customers. Here are three examples of how companies have used mass customization:
 - ▶ Jones Soda Co., a U.S. based beverage company, has been using a customization strategy by constantly changing its product labels that are generated and submitted by its consumers.⁹
 - ▶ Starbucks lets its customers customize their Frappuccino drinks.¹⁰
 - ▶ Nespresso, part of Nestle Group head quartered in Lausanne, Switzerland, has been utilizing mass customization as a way of forging long lasting consumer relationships. One of the company's core strategic objectives is to enable anyone to create the perfect cup of espresso coffee at home—just like a skilled barista.¹¹ Mass customization has helped Nespresso achieve this goal by shaping the purchase and consumption of coffee into a unique and engaging experience. Convenient and easy-to-use Nespresso machines, as well as availability of a wide range of coffee choices in individual capsule formats, enable consumers to customize their drinks to their likings, based on coffee flavor, strength, aroma, caffeine level, and more.
- 2. Private Label:** The emergence of private labeling as an important trend has been supported by two main forces; (a) suppliers and retailers' efforts to lower the price of products for driving store traffic and (b) pressuring the well-established brands in the battle for market share. Large retailers have also used this strategy as a critical source of competitive advantage by forming strong relationships with their customers. However, this trend began to show strong signs of growth on the global basis, only after the economic crisis of 2008. Increasing number of value conscious consumers raised the price elasticity of demand in major markets, which in effect generated a higher demand for private label

⁹<http://www.jonessoda.com>.

¹⁰<http://www.starbucks.com/blog/how-do-you-customize-your-frappuccino174-blended-beverage/1031>.

¹¹<http://nestle-nespresso.com>.

and value products. With the economic slow-down starting to improve, we may experience a decline in the growth rate and popularity of private brands. However, many industry experts view private labeling as a game changing force for competing against well-established brands in the long run. Their viewpoint is supported by the fact that increasing number of consumers are showing interests in buying private label products that compete on the basis of offering superior economy while maintaining an above average quality.

3. **Rising Cost of Raw Materials:** In recent years, manufacturers across many industries, including the beverages, have been faced with rising costs of raw materials, as well as higher energy and insurance costs. For instance, within the four-year period from 2007 to 2011, PET resin prices increased 39.2%, oil prices increased 21.6%, sugar prices increased 163.2%, and corn prices increased over 67.5%.¹² Rising costs, particularly those related to the costs of ingredients and packaging have translated into increased prices for consumers. If this trend, as forecasts suggest, continues its growth in the next few years, the well-established brands and premium products that often compete on a point of difference other than the promise of superior economy, may find themselves in a better position to strive against the value brands and private label products.
4. **Asset-Light Business Model:** As businesses are forced to adapt to change faster than ever and to control the increasing level of complexity in their operations, they realize that achieving profitable growth is becoming more challenging. Business model modification has been among the employed strategies for overcoming such challenges in an effort to maximize the value for the company and its customers through combining various elements of a business into a well-integrated system.

As more companies began to employ an *asset-light* business model for their operations, an underlying trend in the beverage industry started to take form with respect to business model modification. Under this operational framework, companies, often the start-ups, seek to maximize their value by investing less on manufacturing, and instead, to focus limited resources on branding and other marketing activities, and in some cases, on distribution. Availability of a wide range of private label operators and ingredient houses; for filling and product development, respectively; as well as growing demand for premium and healthy products that offer higher profit margins have allowed these companies to adopt the asset-light business model approach.

Despite the fact that several start-up companies in the beverage industry have proved to be successful in recent years, many industry experts point out the potential risks associated with the use of this business model in the long run. For instance, the strategic dependence of the firm throughout the value chain, lower profit margins, and the exposure of the firm to raw material price fluctuations are among the most discussed downsides. Gaining distribution for products and the high costs associated with building a strong brand have been the major barriers to entry. The case of “Innocent Drinks Co.”¹³ is a success story of a company that went through a fast phase of growth by using the light-asset model. After establishing a promising niche market, a strong and consistent brand image, and a loyal consumer base, it was partly acquired by Coca-Cola in 2010.

ENERGY DRINKS

Energy drinks provide an interesting example of the mechanics of supply and demand.”¹⁴ Energy drinks are not really a new category in the beverage market. In fact, Coca-Cola was marketed as an energy drink when it was first launched over 100 years ago. But the key to the growing popularity of this category today was the development and marketing of Red Bull, currently the most popular energy drink in the world.

¹²R. Colbert, Global Drinks Report 2012, the best times, the worst times, March 2012. *Beverage Innovation*, (93), 42–43.

¹³<http://www.innocentdrinks.co.uk>.

¹⁴For a brief but thorough overview of the energy drink market, see http://en.wikipedia.org/wiki/Energy_drink.

Other popular brands include Burn (from Coca Cola), Monster (from Monster Beverage Co.), and 5-Hour Energy (from Living Essentials). The key regional players are Power Horse in the Middle East, Lipovitan (from Taisho Pharmaceutical) in Asia Pacific, and V (from Frucor) in Australasia. Regions such as eastern Europe, western Europe, and Latin America have no clear regional leader.¹⁵

The energy drinks category has been growing steadily over the past decade. Energy drink sales are highly concentrated in developed markets, and their general popularity is mainly because of their popularity in North America, Australasia, and western Europe. Collectively, the United States, the United Kingdom, and Australia account for almost half the world's consumption of this type of drink.¹⁶ On a per capita basis, people in these markets consume more than ten times the amount of energy drinks than the rest of the world.¹⁷

Energy drinks have incited controversy because of their alleged negative health effects.¹⁸ Nevertheless, this controversy has apparently not slowed the increase in their consumption in certain regions or countries in the world. Most of the negative perceptions are held by older consumers, who are not the primary target market for these products. In fact, the negative perception of energy drinks may have helped increase the popularity of energy drinks among younger consumers.¹⁹

In terms of the dynamics of supply and demand, we see from the above discussion that the energy drinks category provides a good illustration of a monopolistically competitive market. Entry is relatively easy, but each seller tries to compete by differentiating itself in one way or another. Red Bull has a large concentration of caffeine and sugar, but so do other energy drinks. Red Bull's distinguishing feature is its use of lipovitan in its formulation. The founder of Red Bull got the idea from a Thai drink called Krating Daeng, which also uses this ingredient.

In addition to differentiation based on ingredients, companies have tried to differentiate their products on the basis of packaging as well as the new-usage-occasions. For example, in recent years, energy drinks have been sold in smaller one- and two-ounce servings called energy shots. The amount of the ingredients providing energy (i.e., caffeine and sugar) is about the same as a normal size serving. Thus, consumers get their energy in a highly concentrated form. The idea behind this product development strategy was to focus more on the functionality of products and less on the aspirational attributes and image of the brands. This strategy aimed to create new-usage-occasions by increasing the appeal to the more conservative consumers who might consider drinking these products during the day on the basis of their functionality. Unfortunately for sellers, the idea of offering energy shots to expand the consumer base and usage of energy drinks has not been very successful. These new product introductions remained unpopular with most of the new target consumers and as a result energy shots remained a niche category in most markets. Energy shots are more expensive than the energy drinks, which themselves are already expensive in comparison to other soft drinks. Here we see another example of the economic laws of supply and demand at work. No matter how creative the marketing effort, it is very difficult to overcome limited demand and high price elasticity.

¹⁵C. Martinez, Global Energy Drinks, February 2012. *Beverage Innovation*, pp. 14–15. (Original source: Zenith International Ltd, www.zenithinternational.com)

¹⁶Euromonitor International, Energy Drinks Entering a New Phase of Growth, 2012.

¹⁷Martinez, op. cit.

¹⁸<http://www.helium.com/items/316590-energy-drinks-negative-health-effects> 19.

¹⁹Martinez, op. cit.

Overall, the energy drinks category has a strong potential for growth in the future. In developed markets, companies will try to spur growth through strategies such as: creating new-usage-occasions, i.e., positioning energy drinks as a coffee substitute; and also by expanding the category's consumer base through developing products with a more healthful image. Developing markets such as Asia-Pacific, Middle East, Africa, and Latin America, on the other hand, are likely to duplicate the successful strategies of energy drinks in the developed markets by competing on the basis of image and functionality, with a heavier focus placed on image as a main source of differentiation.

This page intentionally left blank

Statistical and Financial Tables

Table A.1a Future Value of \$1 at the End of n Periods

Period	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
1	1.0100	1.0200	1.0300	1.0400	1.0500	1.0600	1.0700	1.0800	1.0900	1.1000
2	1.0201	1.0404	1.0609	1.0816	1.1025	1.1236	1.1449	1.1664	1.1881	1.2100
3	1.0303	1.0612	1.0927	1.1249	1.1576	1.1910	1.2250	1.2597	1.2950	1.3310
4	1.0406	1.0824	1.1255	1.1699	1.2155	1.2625	1.3108	1.3605	1.4116	1.4641
5	1.0510	1.1041	1.1593	1.2167	1.2763	1.3382	1.4026	1.4693	1.5386	1.6105
6	1.0615	1.1262	1.1941	1.2653	1.3401	1.4185	1.5007	1.5869	1.6771	1.7716
7	1.0721	1.1487	1.2299	1.3159	1.4071	1.5036	1.6058	1.7138	1.8280	1.9487
8	1.0829	1.1717	1.2668	1.3686	1.4775	1.5938	1.7182	1.8509	1.9926	2.1436
9	1.0937	1.1951	1.3048	1.4233	1.5513	1.6895	1.8385	1.9990	2.1719	2.3579
10	1.1046	1.2190	1.3439	1.4802	1.6289	1.7908	1.9672	2.1589	2.3674	2.5937
11	1.1157	1.2434	1.3842	1.5395	1.7103	1.8983	2.1049	2.3316	2.5804	2.8531
12	1.1268	1.2682	1.4258	1.6010	1.7959	2.0122	2.2522	2.5182	2.8127	3.1384
13	1.1381	1.2936	1.4685	1.6651	1.8856	2.1329	2.4098	2.7196	3.0658	3.4523
14	1.1495	1.3195	1.5126	1.7317	1.9799	2.2609	2.5785	2.9372	3.3417	3.7975
15	1.1610	1.3459	1.5580	1.8009	2.0789	2.3966	2.7590	3.1722	3.6425	4.1772
16	1.1726	1.3728	1.6047	1.8730	2.1829	2.5404	2.9522	3.4259	3.9703	4.5950
17	1.1843	1.4002	1.6528	1.9479	2.2920	2.6928	3.1588	3.7000	4.3276	5.0545
18	1.1961	1.4282	1.7024	2.0258	2.4066	2.8543	3.3799	3.9960	4.7171	5.5599
19	1.2081	1.4568	1.7535	2.1068	2.5270	3.0256	3.6165	4.3157	5.1417	6.1159
20	1.2202	1.4859	1.8061	2.1911	2.6533	3.2071	3.8697	4.6610	5.6044	6.7275
21	1.2324	1.5157	1.8603	2.2788	2.7860	3.3996	4.1406	5.0338	6.1088	7.4002
22	1.2477	1.5460	1.9161	2.3699	2.9253	3.6035	4.4304	5.4365	6.6586	8.1403
23	1.2572	1.5769	1.9736	2.4647	3.0715	3.8197	4.7405	5.8715	7.2579	8.9543
24	1.2697	1.6084	2.0328	2.5633	3.2251	4.0489	5.0724	6.3412	7.9111	9.8497
25	1.2824	1.6406	2.0938	2.6658	3.3864	4.2919	5.4274	6.8485	8.6231	10.835
26	1.2953	1.6734	2.1566	2.7725	3.5557	4.5494	5.8074	7.3964	9.3992	11.918
27	1.3082	1.7069	2.2213	2.8834	3.7335	4.8223	6.2139	7.9881	10.245	13.110
28	1.3213	1.7410	2.2879	2.9987	3.9201	5.1117	6.6488	8.6271	11.167	14.421
29	1.3345	1.7758	2.3566	3.1187	4.1161	5.4184	7.1143	9.3173	12.172	15.863
30	1.3478	1.8114	2.4273	3.2434	4.3219	5.7435	7.6123	10.063	13.268	17.449
40	1.4889	2.2080	3.2620	4.8010	7.0400	10.286	14.974	21.725	31.409	45.259
50	1.6446	2.6916	4.3839	7.1067	11.467	18.420	29.457	46.902	74.358	117.39
60	1.8167	3.2810	5.8916	10.520	18.679	32.988	57.946	101.26	176.03	304.48

Table A.1a (continued)

Period	12%	14%	15%	16%	18%	20%	24%	28%	32%	36%
1	1.1200	1.1400	1.1500	1.1600	1.1800	1.2000	1.2400	1.2800	1.3200	1.3600
2	1.2544	1.2996	1.3225	1.3456	1.3924	1.4400	1.5376	1.6384	1.7424	1.8496
3	1.4049	1.4815	1.5209	1.5609	1.6430	1.7280	1.9066	2.0972	2.3000	2.5155
4	1.5735	1.6890	1.7490	1.8106	1.9388	2.0736	2.3642	2.6844	3.0360	3.4210
5	1.7623	1.9254	2.0114	2.1003	2.2878	2.4883	2.9316	3.4360	4.0075	4.6526
6	1.9738	2.1950	2.3131	2.4364	2.6996	2.9860	3.6352	4.3980	5.2899	6.3275
7	2.2107	2.5023	2.6600	2.8262	3.1855	3.5832	4.5077	5.6295	6.9826	8.6054
8	2.4760	2.8526	3.0590	3.2784	3.7589	4.2998	5.5895	7.2058	9.2170	11.703
9	2.7731	3.2519	3.5179	3.8030	4.4355	5.1598	6.9310	9.2234	12.166	15.917
10	3.1058	3.7072	4.0456	4.4114	5.2338	6.1917	8.5944	11.806	16.060	21.647
11	3.4785	4.2262	4.6524	5.1173	6.1759	7.4301	10.657	15.112	21.199	29.439
12	3.8960	4.8179	5.3502	5.9360	7.2876	8.9161	13.215	19.343	27.982	40.037
13	4.3635	5.4924	6.1528	6.8858	8.5994	10.699	16.386	24.759	36.937	54.451
14	4.8871	6.2613	7.0757	7.9875	10.147	12.839	20.319	31.691	48.756	74.053
15	5.4736	7.1379	8.1371	9.2655	11.974	15.407	25.196	40.565	64.359	100.71
16	6.1304	8.1372	9.3576	10.748	14.129	18.488	31.243	51.923	84.954	136.97
17	6.8660	9.2765	10.761	12.468	16.672	22.186	38.741	66.461	112.14	186.28
18	7.6900	10.575	12.375	14.463	19.673	26.623	48.039	85.071	148.02	253.34
19	8.6128	12.056	14.232	16.777	23.214	31.948	59.568	108.89	195.39	344.54
20	9.6463	13.743	16.367	19.461	27.393	38.338	73.864	139.38	257.92	468.58
21	10.804	15.668	18.822	22.574	32.324	46.005	91.592	178.41	340.45	637.26
22	12.100	17.861	21.645	26.186	38.142	55.206	113.57	228.36	449.39	866.67
23	13.552	20.362	24.891	30.376	45.008	66.247	140.83	292.30	593.20	1178.7
24	15.179	23.212	28.625	35.236	53.109	79.497	174.63	374.14	783.02	1603.0
25	17.000	26.462	32.919	40.874	62.669	95.396	216.54	478.90	1033.6	2180.1
26	19.040	30.167	37.857	47.414	73.949	114.48	268.51	613.00	1364.3	2964.9
27	21.325	34.390	43.535	55.000	87.260	137.37	332.95	784.64	1800.9	4032.3
28	23.884	39.204	50.066	63.800	102.97	164.84	412.86	1004.3	2377.2	5483.9
29	16.750	44.693	57.575	74.009	121.50	197.81	511.95	1285.6	3137.9	7458.1
30	29.960	50.950	66.212	85.850	143.37	237.38	634.82	1645.5	4142.1	10143
40	93.051	188.88	267.86	378.72	750.38	1469.8	5455.9	19427	66521	*
50	289.00	700.23	1083.7	1670.7	3927.4	9100.4	46890	*	*	*
60	897.60	2595.9	4384.0	7370.2	20555	56348	*	*	*	*

* > 99,999

Table A.1b Sum of an Annuity of \$1 per Period for n Periods

Number of Periods	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	2.0100	2.0200	2.0300	2.0400	2.0500	2.0600	2.0700	2.0800	2.0900	2.1000
3	3.0301	3.0604	3.0909	3.1216	3.1525	3.1836	3.2149	3.2464	3.2781	3.3100
4	4.0604	4.1216	4.1836	4.2465	4.3101	4.3746	4.4399	4.5061	4.5731	4.6410
5	5.1010	5.2040	5.3091	5.4163	5.5256	5.6371	5.7507	5.8666	5.9847	6.1051
6	6.1520	6.3081	6.4684	6.6330	6.8019	6.9753	7.1533	7.3359	7.5233	7.7156
7	7.2135	7.4343	7.6625	7.8983	8.1420	8.3938	8.6540	8.9228	9.2004	9.4872
8	8.2857	8.5830	8.8923	9.2142	9.5491	9.8975	10.260	10.637	11.028	11.436
9	9.3685	9.7546	10.159	10.583	11.027	11.491	11.978	12.488	13.021	13.579
10	10.462	10.950	11.464	12.006	12.578	13.181	13.816	14.487	15.193	15.937
11	11.567	12.169	12.808	13.486	14.207	14.972	15.784	16.645	17.560	18.531
12	12.683	13.412	14.192	15.026	15.917	16.870	17.888	18.977	20.141	21.384
13	13.809	14.680	15.618	16.627	17.713	18.882	20.141	21.495	22.953	24.523
14	14.947	15.974	17.086	18.292	19.599	21.015	22.550	24.215	26.019	27.975
15	16.097	17.293	18.599	20.024	21.579	23.276	25.129	27.152	29.361	31.772
16	17.258	18.639	20.157	21.825	23.657	25.673	27.888	30.324	33.003	35.950
17	18.430	20.012	21.762	23.698	25.840	28.213	30.840	33.750	36.974	40.545
18	19.615	21.412	23.414	25.645	28.132	30.906	33.999	37.450	41.301	45.599
19	20.811	22.841	25.117	27.671	30.539	33.760	37.379	41.446	46.018	51.159
20	22.019	24.297	26.870	29.778	33.066	36.786	40.995	45.762	51.160	57.275
21	23.239	25.783	28.676	31.969	35.719	39.993	44.865	50.423	56.765	64.002
22	24.472	27.299	30.537	34.248	38.505	43.392	49.006	55.457	62.873	71.403
23	25.716	28.845	32.453	36.618	41.430	46.996	53.436	60.893	69.532	79.543
24	26.973	30.422	34.426	39.083	44.502	50.816	58.177	66.765	76.790	88.497
25	28.243	32.030	36.459	41.646	47.727	54.865	63.249	73.106	84.701	98.347
26	29.526	33.671	38.553	44.312	51.113	59.156	68.676	79.954	93.324	109.18
27	30.821	35.344	40.710	47.084	54.669	63.706	74.484	87.351	102.72	121.10
28	32.129	37.051	42.931	49.968	58.403	68.528	80.698	95.339	112.97	134.21
29	33.450	38.792	45.219	52.966	62.323	73.640	87.347	103.97	124.14	148.63
30	34.785	40.568	47.575	56.085	66.439	79.058	94.461	113.28	136.31	164.49
40	48.886	60.402	75.401	95.026	120.80	154.76	199.64	259.06	337.88	442.59
50	64.463	84.579	112.80	152.67	209.35	290.34	406.53	573.77	815.08	1163.9
60	81.670	114.05	163.05	237.99	353.58	533.13	813.52	1253.2	1944.8	3034.8

Table A.1b (continued)

Number of Periods	12%	14%	15%	16%	18%	20%	24%	28%	32%	36%
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	2.1200	2.1400	2.1500	2.1600	2.1800	2.2000	2.2400	2.2800	2.3200	2.3600
3	3.3744	3.4396	3.4725	3.5056	3.5724	3.6400	3.7776	3.9184	4.0624	4.2096
4	4.7793	4.9211	4.9934	5.0665	5.2154	5.3680	5.6842	6.0156	6.3624	6.7251
5	6.3528	6.6101	6.7424	6.8771	7.1542	7.4416	8.0484	8.6999	9.3983	10.146
6	8.1152	8.5355	8.7537	8.9775	9.4420	9.9299	10.980	12.136	13.406	14.799
7	10.089	10.730	11.067	11.414	12.142	12.916	14.615	16.534	18.696	21.126
8	12.300	13.233	13.727	14.240	15.327	16.499	19.123	22.163	25.678	29.732
9	14.776	16.085	16.786	17.519	19.086	20.799	24.712	29.369	34.895	41.435
10	17.549	19.337	20.304	21.321	23.521	25.959	31.643	38.593	47.062	57.352
11	20.655	23.045	24.349	25.733	28.755	32.150	40.238	50.398	63.122	78.998
12	24.133	27.271	29.002	30.850	34.931	39.581	50.895	65.510	84.320	108.44
13	28.029	32.089	34.352	36.786	42.429	48.497	64.110	84.853	112.30	148.47
14	32.393	37.581	40.505	43.672	50.818	59.196	80.496	109.61	149.24	202.93
15	37.280	43.842	47.580	51.660	60.965	72.035	100.82	141.30	198.00	276.98
16	42.753	50.980	55.717	60.925	72.939	87.442	126.01	181.87	262.36	377.69
17	48.884	59.118	65.075	71.673	87.068	105.93	157.25	233.79	347.31	514.66
18	55.750	68.394	75.836	84.141	103.74	128.11	195.99	300.25	459.45	700.94
19	63.440	78.969	88.212	98.603	123.41	154.74	244.03	385.32	607.47	954.28
20	72.052	91.025	102.44	115.38	146.63	186.69	303.60	494.21	802.86	1298.8
21	81.699	104.77	118.81	134.84	174.02	225.03	377.46	633.59	1060.8	1767.4
22	92.503	120.44	137.63	157.41	206.34	271.03	469.06	812.00	1401.2	2404.7
23	104.60	138.30	159.28	183.60	244.49	326.24	582.63	1040.4	1850.6	3271.3
24	118.16	158.66	184.17	213.98	289.49	392.48	723.46	1332.7	2443.8	4450.0
25	133.33	181.87	212.79	249.21	342.60	471.98	898.09	1706.8	3226.8	6053.0
26	150.33	208.33	245.71	290.09	405.27	567.38	1114.6	2185.7	4260.4	8233.1
27	169.37	238.50	283.57	337.50	479.22	681.85	1383.1	2798.7	5624.8	11198.0
28	190.70	272.89	327.10	392.50	566.48	819.22	1716.1	3583.3	7425.7	15230.3
29	214.58	312.09	377.17	456.30	669.45	984.07	2129.0	4587.7	9802.9	20714.2
30	241.33	356.79	434.75	530.31	790.95	1181.9	2640.9	5873.2	12941	28172.3
40	767.09	1342.0	1779.1	2360.8	4163.2	7343.9	22729	69377	*	*
50	2400.0	4994.5	7217.7	10436	21813	45497	*	*	*	*
60	7471.6	18535	29220	46058	*	*	*	*	*	*

* > 99,999

Table A.1c Present Value of \$1 Received at the End of n Periods

Period	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
1	0.9901	0.9804	0.9709	0.9615	0.9524	0.9434	0.9346	0.9259	0.9174	0.9091
2	0.9803	0.9612	0.9426	0.9246	0.9070	0.8900	0.8734	0.8573	0.8417	0.8264
3	0.9706	0.9423	0.9151	0.8890	0.8638	0.8396	0.8163	0.7938	0.7722	0.7513
4	0.9610	0.9238	0.8885	0.8548	0.8227	0.7921	0.7629	0.7350	0.7084	0.6830
5	0.9515	0.9057	0.8626	0.8219	0.7835	0.7473	0.7130	0.6806	0.6499	0.6209
6	0.9420	0.8880	0.8375	0.7903	0.7462	0.7050	0.6663	0.6302	0.5963	0.5645
7	0.9327	0.8706	0.8131	0.7599	0.7107	0.6651	0.6227	0.5835	0.5470	0.5132
8	0.9235	0.8535	0.7894	0.7307	0.6768	0.6274	0.5820	0.5403	0.5019	0.4665
9	0.9143	0.8368	0.7664	0.7026	0.6446	0.5919	0.5439	0.5002	0.4604	0.4241
10	0.9053	0.8203	0.7441	0.6756	0.6139	0.5584	0.5083	0.4632	0.4224	0.3855
11	0.8963	0.8043	0.7224	0.6496	0.5847	0.5268	0.4751	0.4289	0.3875	0.3505
12	0.8874	0.7885	0.7014	0.6246	0.5568	0.4970	0.4440	0.3971	0.3555	0.3186
13	0.8787	0.7730	0.6810	0.6006	0.5303	0.4688	0.4150	0.3677	0.3262	0.2897
14	0.8700	0.7579	0.6611	0.5775	0.5051	0.4423	0.3878	0.3405	0.2992	0.2633
15	0.8613	0.7430	0.6419	0.5553	0.4810	0.4173	0.3624	0.3152	0.2745	0.2394
16	0.8528	0.7284	0.6232	0.5339	0.4581	0.3936	0.3387	0.2919	0.2519	0.2176
17	0.8444	0.7142	0.6050	0.5134	0.4363	0.3714	0.3166	0.2703	0.2311	0.1978
18	0.8360	0.7002	0.5874	0.4936	0.4155	0.3503	0.2959	0.2502	0.2120	0.1799
19	0.8277	0.6864	0.5703	0.4746	0.3957	0.3305	0.2765	0.2317	0.1945	0.1635
20	0.8195	0.6730	0.5537	0.4564	0.3769	0.3118	0.2584	0.2145	0.1784	0.1486
25	0.7798	0.6095	0.4776	0.3751	0.2953	0.2330	0.1842	0.1460	0.1160	0.0923
30	0.7419	0.5521	0.4120	0.3083	0.2314	0.1741	0.1314	0.0994	0.0754	0.0573
40	0.6717	0.4529	0.3066	0.2083	0.1420	0.0972	0.0668	0.0460	0.0318	0.0221
50	0.6080	0.3715	0.2281	0.1407	0.0872	0.0543	0.0339	0.0213	0.0134	0.0085
60	0.5504	0.3048	0.1697	0.0951	0.0535	0.0303	0.0173	0.0099	0.0057	0.0033

Table A.1c (continued)

Period	12%	14%	15%	16%	18%	20%	24%	28%	32%	36%
1	0.8929	0.8772	0.8696	0.8621	0.8475	0.8333	0.8065	0.7813	0.7576	0.7353
2	0.7972	0.7695	0.7561	0.7432	0.7182	0.6944	0.6504	0.6104	0.5739	0.5407
3	0.7118	0.6750	0.6575	0.6407	0.6086	0.5787	0.5245	0.4768	0.4348	0.3975
4	0.6355	0.5921	0.5718	0.5523	0.5158	0.4823	0.4230	0.3725	0.3294	0.2923
5	0.5674	0.5194	0.4972	0.4761	0.4371	0.4019	0.3411	0.2910	0.2495	0.2149
6	0.5066	0.4556	0.4323	0.4104	0.3704	0.3349	0.2751	0.2274	0.1890	0.1580
7	0.4523	0.3996	0.3759	0.3538	0.3139	0.2791	0.2218	0.1776	0.1432	0.1162
8	0.4039	0.3506	0.3269	0.3050	0.2660	0.2326	0.1789	0.1388	0.1085	0.0854
9	0.3606	0.3075	0.2843	0.2630	0.2255	0.1938	0.1443	0.1084	0.0822	0.0628
10	0.3220	0.2697	0.2472	0.2267	0.1911	0.1615	0.1164	0.0847	0.0623	0.0462
11	0.2875	0.2366	0.2149	0.1954	0.1619	0.1346	0.0938	0.0662	0.0472	0.0340
12	0.2567	0.2076	0.1869	0.1685	0.1372	0.1122	0.0757	0.0517	0.0357	0.0250
13	0.2292	0.1821	0.1625	0.1452	0.1163	0.0935	0.0610	0.0404	0.0271	0.0184
14	0.2046	0.1597	0.1413	0.1252	0.0985	0.0779	0.0492	0.0316	0.0205	0.0135
15	0.1827	0.1401	0.1229	0.1079	0.0835	0.0649	0.0397	0.0247	0.0155	0.0099
16	0.1631	0.1229	0.1069	0.0930	0.0708	0.0541	0.0320	0.0193	0.0118	0.0073
17	0.1456	0.1078	0.0929	0.0802	0.0600	0.0451	0.0258	0.0150	0.0089	0.0054
18	0.1300	0.0946	0.0808	0.0691	0.0508	0.0376	0.0208	0.0118	0.0068	0.0039
19	0.1161	0.0829	0.0703	0.0596	0.0431	0.0313	0.0168	0.0092	0.0051	0.0029
20	0.1037	0.0728	0.0611	0.0514	0.0365	0.0261	0.0135	0.0072	0.0039	0.0021
25	0.0588	0.0378	0.0304	0.0245	0.0160	0.0105	0.0046	0.0021	0.0010	0.0005
30	0.0334	0.0196	0.0151	0.0116	0.0070	0.0042	0.0016	0.0006	0.0002	0.0001
40	0.0107	0.0053	0.0037	0.0026	0.0013	0.0007	0.0002	0.0001	*	*
50	0.0035	0.0014	0.0009	0.0006	0.0003	0.0001	*	*	*	*
60	0.0011	0.0004	0.0002	0.0001	*	*	*	*	*	*

*The factor is zero to four decimal places.

Table A.1d Present Value of an Annuity of \$1 per Period for n Periods

Number of Payments	1%	2%	3%	4%	5%	6%	7%	8%	9%
1	0.9901	0.9804	0.9709	0.9615	0.9524	0.9434	0.9346	0.9259	0.9174
2	1.9704	1.9416	1.9135	1.8861	1.8594	1.8334	1.8080	1.7833	1.7591
3	2.9410	2.8839	2.8286	2.7751	2.7232	2.6730	2.6243	2.5771	2.5313
4	3.9020	3.8077	3.7171	3.6299	3.5460	3.4651	3.3872	3.3121	3.2397
5	4.8534	4.7135	4.5797	4.4518	4.3295	4.2124	4.1002	3.9927	3.8897
6	5.7955	5.6014	5.4172	5.2421	5.0757	4.9173	4.7665	4.6229	4.4859
7	6.7282	6.4720	6.2303	6.0021	5.7864	5.5824	5.3893	5.2064	5.0330
8	7.6517	7.3255	7.0197	6.7327	6.4632	6.2098	5.9713	5.7466	5.5348
9	8.5660	8.1622	7.7861	7.4353	7.1078	6.8017	6.5152	6.2469	5.9952
10	9.4713	8.9826	8.5302	8.1109	7.7217	7.3601	7.0236	6.7101	6.4177
11	10.3676	9.7868	9.2526	8.7605	8.3064	7.8869	7.4987	7.1390	6.8052
12	11.2551	10.5753	9.9540	9.3851	8.8633	8.3838	7.9427	7.5361	7.1607
13	12.1337	11.3484	10.6350	9.9856	9.3936	8.8527	8.3577	7.9038	7.4869
14	13.0037	12.1062	11.2961	10.5631	9.8986	9.2950	8.7455	8.2442	7.7862
15	13.8651	12.8493	11.9379	11.1184	10.3797	9.7122	9.1079	8.5595	8.0607
16	14.7179	13.5777	12.5611	11.6523	10.8378	10.1059	9.4466	8.8514	8.3126
17	15.5623	14.2919	13.1661	12.1657	11.2741	10.4773	9.7632	9.1216	8.5436
18	16.3983	14.9920	13.7535	12.6593	11.6896	10.8276	10.0591	9.3719	8.7556
19	17.2260	15.6785	14.3238	13.1339	12.0853	11.1581	10.3356	9.6036	8.9501
20	18.0456	16.3514	14.8775	13.5903	12.4622	11.4699	10.5940	9.8181	9.1285
25	22.0232	19.5235	17.4131	15.6221	14.0939	12.7834	11.6536	10.6748	9.8226
30	25.8077	22.3965	19.6004	17.2920	15.3725	13.7648	12.4090	11.2578	10.2737
40	32.8347	27.3555	23.1148	19.7928	17.1591	15.0463	13.3317	11.9246	10.7574
50	39.1961	31.4236	25.7298	21.4822	18.2559	15.7619	13.8007	12.2335	10.9617
60	44.9550	34.7609	27.6756	22.6235	18.9293	16.1614	14.0392	12.3766	11.0480

Table A.1d (continued)

Number of Payments	10%	12%	14%	15%	16%	18%	20%	24%	28%	32%	36%
1	0.9091	0.8929	0.8772	0.8696	0.8621	0.8475	0.8333	0.8065	0.7813	0.7576	0.7353
2	1.7355	1.6901	1.6467	1.6257	1.6052	1.5656	1.5278	1.4568	1.3916	1.3315	1.2760
3	2.4869	2.4018	2.3216	2.2832	2.2459	2.1743	2.1065	1.9813	1.8684	1.7663	1.6735
4	3.1699	3.0373	2.9137	2.8550	2.7982	2.6901	2.5887	2.4043	2.2410	2.0957	1.9658
5	3.7908	3.6048	3.4331	3.3522	3.2743	3.1272	2.9906	2.7454	2.5320	2.3452	2.1807
6	4.3553	4.1114	3.8887	3.7845	3.6847	3.4976	3.3255	3.0205	2.7594	2.5342	2.3388
7	4.8684	4.5638	4.2883	4.1604	4.0386	3.8115	3.6046	3.2423	2.9370	2.6775	2.4550
8	5.3349	4.9676	4.6389	4.4873	4.3436	4.0776	3.8372	3.4212	3.0758	2.7860	2.5404
9	5.7590	5.3282	4.9464	4.7716	4.6065	4.3030	4.0310	3.5655	3.1842	2.8681	2.6033
10	6.1446	5.6502	5.2161	5.0188	4.8332	4.4941	4.1925	3.6819	3.2689	2.9304	2.6495
11	6.4951	5.9377	5.4527	5.2337	5.0286	4.6560	4.3271	3.7757	3.3351	2.9776	2.6834
12	6.8137	6.1944	5.6603	5.4206	5.1971	4.7932	4.4392	3.8514	3.3868	3.0133	2.7084
13	7.1034	6.4235	5.8424	5.5831	5.3423	4.9095	4.5327	3.9124	3.4272	3.0404	2.7268
14	7.3667	6.6282	6.0021	5.7245	5.4675	5.0081	4.6106	3.9616	3.4587	3.0609	2.7403
15	7.6061	6.8109	6.1422	5.8474	5.5755	5.0916	4.6755	4.0013	3.4834	3.0764	2.7502
16	7.8237	6.9740	6.2651	5.9542	5.6685	5.1624	4.7296	4.0333	3.5026	3.0882	2.7575
17	8.0216	7.1196	6.3729	6.0472	5.7487	5.2223	4.7746	4.0591	3.5177	3.0971	2.7629
18	8.2014	7.2497	6.4674	6.1280	5.8178	5.2732	4.8122	4.0799	3.5294	3.1039	2.7668
19	8.3649	7.3658	6.5504	6.1982	5.8775	5.3162	4.8435	4.0967	3.5386	3.1090	2.7697
20	8.5136	7.4694	6.6231	6.2593	5.9288	5.3527	4.8696	4.1103	3.5458	3.1129	2.7718
25	9.0770	7.8431	6.8729	6.4641	6.0971	5.4669	4.9476	4.1474	3.5640	3.1220	2.7765
30	9.4269	8.0552	7.0027	6.5660	6.1772	5.5168	4.9789	4.1601	3.5693	3.1242	2.7775
40	9.7791	8.2438	7.1050	6.6418	6.2335	5.5482	4.9966	4.1659	3.5712	3.1250	2.7778
50	9.9148	8.3045	7.1327	6.6605	6.2463	5.5541	4.9995	4.1666	3.5714	3.1250	2.7778
60	9.9672	8.3240	7.1401	6.6651	6.2492	5.5553	4.9999	4.1667	3.5714	3.1250	2.7778

Table A.2 Areas Under the Normal Curve

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

Table A.3a Critical Values for the *F*-Distribution ($\alpha = .05$)

Degrees of Freedom for Denominator	Degrees of Freedom for Numerator								
	1	2	3	4	5	6	8	10	15
1	161.4	199.5	215.7	224.6	230.2	234.0	238.9	241.9	245.9
2	18.51	19.00	19.16	19.25	19.30	19.33	19.37	19.40	19.43
3	10.13	9.55	9.28	9.12	9.01	8.94	8.85	8.79	8.70
4	7.71	6.94	6.59	6.39	6.26	6.16	6.04	5.96	5.86
5	6.61	5.79	5.41	5.19	5.05	4.95	4.82	4.74	4.62
6	5.99	5.14	4.76	4.53	4.39	4.28	4.15	4.06	3.94
7	5.59	4.74	4.35	4.12	3.97	3.87	3.73	3.64	3.51
8	5.32	4.46	4.07	3.84	3.69	3.58	3.44	3.35	3.22
9	5.12	4.26	3.86	3.63	3.48	3.37	3.23	3.14	3.01
10	4.96	4.10	3.71	3.48	3.33	3.22	3.07	2.98	2.85
11	4.84	3.98	3.59	3.36	3.20	3.09	2.95	2.85	2.72
12	4.75	3.89	3.49	3.26	3.11	3.00	2.85	2.75	2.62
13	4.67	3.81	3.41	3.18	3.03	2.92	2.77	2.67	2.53
14	4.60	3.74	3.34	3.11	2.96	2.85	2.70	2.60	2.46
15	4.54	3.68	3.29	3.06	2.90	2.79	2.64	2.54	2.40
16	4.49	3.63	3.24	3.01	2.85	2.74	2.59	2.49	2.35
17	4.45	3.59	3.20	2.96	2.81	2.70	2.55	2.45	2.31
18	4.41	3.55	3.16	2.93	2.77	2.66	2.51	2.41	2.27
19	4.38	3.52	3.13	2.90	2.74	2.63	2.48	2.38	2.23
20	4.35	3.49	3.10	2.87	2.71	2.60	2.45	2.35	2.20
21	4.32	3.47	3.07	2.84	2.68	2.57	2.42	2.32	2.18
22	4.30	3.44	3.05	2.82	2.66	2.55	2.40	2.30	2.15
23	4.28	3.42	3.03	2.80	2.64	2.53	2.37	2.27	2.13
24	4.26	3.40	3.01	2.78	2.62	2.51	2.36	2.25	2.11
25	4.24	3.39	2.99	2.76	2.60	2.49	2.34	2.24	2.09
26	4.23	3.37	2.98	2.74	2.59	2.47	2.32	2.22	2.07
27	4.21	3.35	2.96	2.73	2.57	2.46	2.31	2.20	2.06
28	4.20	3.34	2.95	2.71	2.56	2.45	2.29	2.19	2.04
29	4.18	3.33	2.93	2.70	2.55	2.43	2.28	2.18	2.03
30	4.17	3.32	2.92	2.69	2.53	2.42	2.27	2.16	2.01
40	4.08	3.23	2.84	2.61	2.45	2.34	2.18	2.08	1.92
50	4.03	3.18	2.79	2.56	2.40	2.29	2.13	2.03	1.87
60	4.00	3.15	2.76	2.53	2.37	2.25	2.10	1.99	1.84
70	3.98	3.13	2.74	2.50	2.35	2.23	2.07	1.97	1.81
80	3.96	3.11	2.72	2.49	2.33	2.21	2.06	1.95	1.79
90	3.95	3.10	2.71	2.47	2.32	2.20	2.04	1.94	1.78
100	3.94	3.09	2.70	2.46	2.31	2.19	2.03	1.93	1.77
125	2.93	3.07	2.68	2.44	2.29	2.17	2.01	1.91	1.75
150	3.90	3.06	2.66	2.43	2.27	2.16	2.00	1.89	1.73
200	3.89	3.04	2.65	2.42	2.26	2.14	1.98	1.88	1.72
∞	3.84	3.00	2.60	2.37	2.21	2.10	1.94	1.83	1.67

Table A.3b Critical Values for the F -Distribution ($\alpha = .01$)

Degrees of Freedom for Denominator	Degrees of Freedom for Numerator									
	1	2	3	4	5	6	8	10	15	
1	4052	4999	5403	5625	5764	5859	5981	6056	6157	
2	98.50	99.00	99.17	99.25	99.30	99.33	99.37	99.40	99.43	
3	34.12	30.82	29.46	28.71	28.24	27.91	27.49	27.23	26.87	
4	21.20	18.00	16.69	15.98	15.52	15.21	14.80	14.55	14.20	
5	16.26	13.27	12.06	11.39	10.97	10.67	10.29	10.05	9.72	
6	13.75	10.92	9.78	9.15	8.75	8.47	8.10	7.87	7.56	
7	12.25	9.55	8.45	7.85	7.46	7.19	6.84	6.62	6.31	
8	11.26	8.65	7.59	7.01	6.63	6.37	6.03	5.81	5.52	
9	10.56	8.02	6.99	6.42	6.06	5.80	5.47	5.26	4.96	
10	10.04	7.56	6.55	5.99	5.64	5.39	5.06	4.85	4.56	
11	9.65	7.21	6.22	5.67	5.32	5.07	4.74	4.54	4.25	
12	9.33	6.93	5.95	5.41	5.06	4.82	4.50	4.30	4.01	
13	9.07	6.70	5.74	5.21	4.86	4.62	4.30	4.10	3.82	
14	8.86	6.51	5.56	5.04	4.69	4.46	4.14	3.94	3.66	
15	8.68	6.36	5.42	4.89	4.56	4.32	4.00	3.80	3.52	
16	8.53	6.23	5.29	4.77	4.44	4.20	3.89	3.69	3.41	
17	8.40	6.11	5.19	4.67	4.34	4.10	3.79	3.59	3.31	
18	8.29	6.01	5.09	4.58	4.25	4.01	3.71	3.51	3.23	
19	8.18	5.93	5.01	4.50	4.17	3.94	3.63	3.43	3.15	
20	8.10	5.85	4.94	4.43	4.10	3.87	3.56	3.37	3.09	
21	8.02	5.78	4.87	4.37	4.04	3.81	3.51	3.31	3.03	
22	7.95	5.72	4.82	4.31	3.99	3.76	3.45	3.26	2.98	
23	7.88	5.66	4.76	4.26	3.94	3.71	3.41	3.21	2.93	
24	7.82	5.61	4.72	4.22	3.90	3.67	3.36	3.17	2.89	
25	7.77	5.57	4.68	4.18	3.85	3.63	3.32	3.13	2.85	
26	7.72	5.53	4.64	4.14	3.82	3.59	3.29	3.09	2.81	
27	7.68	5.49	4.60	4.11	3.78	3.56	3.26	3.06	2.78	
28	7.64	5.45	4.57	4.07	3.75	3.53	3.23	3.03	2.75	
29	7.60	5.42	4.54	4.04	3.73	3.50	3.20	3.00	2.73	
30	7.56	5.39	4.51	4.02	3.70	3.47	3.17	2.98	2.70	
40	7.31	5.18	4.31	3.83	3.51	3.29	2.99	2.80	2.52	
50	7.17	5.06	4.20	3.72	3.41	3.19	2.89	2.70	2.42	
60	7.08	4.98	4.13	3.65	3.34	3.12	2.82	2.63	2.35	
70	7.01	4.92	4.07	3.60	3.29	3.07	2.78	2.59	2.31	
80	6.96	4.88	4.04	3.56	3.26	3.04	2.74	2.55	2.27	
90	6.93	4.85	4.01	3.53	3.23	3.01	2.72	2.52	2.24	
100	6.90	4.82	3.98	3.51	3.21	2.99	2.69	2.50	2.22	
125	6.84	4.78	3.94	3.47	3.17	2.95	2.66	2.47	2.19	
150	6.81	4.75	3.91	3.45	3.14	2.92	2.63	2.44	2.16	
200	6.76	4.71	3.88	3.41	3.11	2.89	2.60	2.41	2.13	
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.51	2.32	2.04	

Table A.4 Critical Values for the t -Distribution

One-Tail $\alpha =$ Two-Tail $\alpha =$	0.10 0.20	0.05 0.10	0.025 0.05	0.01 0.02	0.005 0.01
$df = 1$	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
50	1.299	1.676	2.009	2.402	2.678
60	1.296	1.672	2.000	2.390	2.660
70	1.294	1.667	1.994	2.381	2.648
80	1.292	1.664	1.990	2.374	2.639
90	1.291	1.662	1.987	2.368	2.632
100	1.290	1.660	1.984	2.364	2.626
125	1.288	1.657	1.979	2.357	2.616
150	1.287	1.655	1.976	2.351	2.609
200	1.286	1.653	1.972	2.345	2.601
∞	1.282	1.645	1.960	2.326	2.576

Table A.5a Durbin-Watson Statistic: Significance Points for d_l and d_u (One-Tail Test, $\alpha = .05$)

n	$k = 1$		$k = 2$		$k = 3$		$k = 4$		$k = 5$	
	d_l	d_u	d_l	d_u	d_l	d_u	d_l	d_u	d_l	d_u
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

Note: n = number of observations, k = number of regressors

Table A.5b Durbin-Watson Statistic: Significance Points for d_l and d_u (Two-Tail Test, $\alpha = .05$)

n	$k = 1$		$k = 2$		$k = 3$		$k = 4$		$k = 5$	
	d_l	d_u	d_l	d_u	d_l	d_u	d_l	d_u	d_l	d_u
15	0.95	1.23	0.83	1.40	0.71	1.61	0.59	1.84	0.48	2.09
16	0.98	1.24	0.86	1.40	0.75	1.59	0.64	1.80	0.53	2.03
17	1.01	1.25	0.90	1.40	0.79	1.58	0.68	1.77	0.57	1.98
18	1.03	1.26	0.93	1.40	0.82	1.56	0.72	1.74	0.62	1.93
19	1.06	1.28	0.96	1.41	0.86	1.55	0.76	1.72	0.66	1.90
20	1.08	1.28	0.99	1.41	0.89	1.55	0.79	1.70	0.70	1.87
21	1.10	1.30	1.01	1.41	0.92	1.54	0.83	1.69	0.73	1.84
22	1.12	1.31	1.04	1.42	0.95	1.54	0.86	1.68	0.77	1.82
23	1.14	1.32	1.06	1.42	0.97	1.54	0.89	1.67	0.80	1.80
24	1.16	1.33	1.08	1.43	1.00	1.54	0.91	1.66	0.83	1.79
25	1.18	1.34	1.10	1.43	1.02	1.54	0.94	1.65	0.86	1.77
26	1.19	1.35	1.12	1.44	1.04	1.54	0.96	1.65	0.88	1.76
27	1.21	1.36	1.13	1.44	1.06	1.54	0.99	1.64	0.91	1.75
28	1.22	1.37	1.15	1.45	1.08	1.54	1.01	1.64	0.93	1.74
29	1.24	1.38	1.17	1.45	1.10	1.54	1.03	1.63	0.96	1.73
30	1.25	1.38	1.18	1.46	1.12	1.54	1.05	1.63	0.98	1.73
31	1.26	1.39	1.20	1.47	1.13	1.55	1.07	1.63	1.00	1.72
32	1.27	1.40	1.21	1.47	1.15	1.55	1.08	1.63	1.02	1.71
33	1.28	1.41	1.22	1.48	1.16	1.55	1.10	1.63	1.04	1.71
34	1.29	1.41	1.24	1.48	1.17	1.55	1.12	1.63	1.06	1.70
35	1.30	1.42	1.25	1.48	1.19	1.55	1.13	1.63	1.07	1.70
36	1.31	1.43	1.26	1.49	1.20	1.56	1.15	1.63	1.09	1.70
37	1.32	1.43	1.27	1.49	1.21	1.56	1.16	1.62	1.10	1.70
38	1.33	1.44	1.28	1.50	1.23	1.56	1.17	1.62	1.12	1.70
39	1.34	1.44	1.29	1.50	1.24	1.56	1.19	1.63	1.13	1.69
40	1.35	1.45	1.30	1.51	1.25	1.57	1.20	1.63	1.15	1.69
45	1.39	1.48	1.34	1.53	1.30	1.58	1.25	1.63	1.21	1.69
50	1.42	1.50	1.38	1.54	1.34	1.59	1.30	1.64	1.26	1.69
55	1.45	1.52	1.41	1.56	1.37	1.60	1.33	1.64	1.30	1.69
60	1.47	1.54	1.44	1.57	1.40	1.61	1.37	1.65	1.33	1.69
65	1.49	1.55	1.46	1.59	1.43	1.62	1.40	1.66	1.36	1.69
70	1.51	1.57	1.48	1.60	1.45	1.63	1.42	1.66	1.39	1.70
75	1.53	1.58	1.50	1.61	1.47	1.64	1.45	1.67	1.42	1.70
80	1.54	1.59	1.52	1.62	1.49	1.65	1.47	1.67	1.44	1.70
85	1.56	1.60	1.53	1.63	1.51	1.65	1.49	1.68	1.46	1.71
90	1.57	1.61	1.55	1.64	1.53	1.66	1.50	1.69	1.48	1.71
95	1.58	1.62	1.56	1.65	1.54	1.67	1.52	1.69	1.50	1.71
100	1.59	1.63	1.57	1.65	1.55	1.67	1.53	1.70	1.51	1.72

Note: n = number of observations, k = number of regressors

This page intentionally left blank

Note: Page numbers followed by f indicate figures; numbers followed by t indicate tables; numbers followed by b indicate boxes; numbers followed by n indicate footnotes.

A

A priori probabilities, 478
 Abandonment option, 491–493
 Abdul Rahim, Khalid B., 203n
 Abnormal profit. *see* Economic profit
 Above-normal profit. *see* Economic profit
 Abrams, Stan, 22n
 Accounting costs, 33–34, 295
 Accounting methods, changes in, 296
 Accounting rate of return, 464, 470, 498
 ACNielsen, 91, 116–117
 Acquisitions, 548–550. *see also* Mergers; Takeovers
 Adelman, Bob, 22n
 Adjusted R², 123
 Adrangi, Bahram, 85n, 89n
 Adverse selection, 448–450, 452, 455
 Advertising, 370
 Advertising elasticity, 90, 94
 Aeppel, Timothy, 413n
 Afrooz, Ahmad, 203n
 Agency costs, 27n
 Agency issues, 450
 Agency problems. *see* Principal-agent problem
 Agency theory, 411n
 Aggregate production function, 203–205
 Ahlstrand, Bruce, 376n
 AiDi, Inc., 397
 AIG, 545
 Airline industry
 mergers in, 547
 price leadership in, 396–397
 pricing in, 425
 service productivity in, 223
 Airline industry, price discrimination in, 405n, 424–425
 Akerlof, George, 443
 Akin, John S., 85n
 Alcatel, 371
 Allocating function of price. *see* Guiding function of price
 Allocation, of resources, 6–10, 53, 534
 Alternative costs, 33–34
 Amazon.com, 397, 526
 AmBev, 379
 America West, 547
 American Airlines, 396, 547
 American Express, 278
 American Express Financial Advisors, 109
 Ames, Glen C. W., 89n
 Anderson Eric T., 175n
 Andreas, Michael, 394
 Andrews, Grenville S., 485n
 Andreyeva, Tatiana, 85n
 Angulo, A. M., 131n
 Anheuser Busch, 378, 379
 Annual equivalent cost (AEC), 486
 Anthony, Robert, 9
 Antitrust laws
 as function of government, 535–538
 tying arrangements and, 409–410
 violation of, 333
 AOL, Inc., 526
 AOL Time Warner, 333
 Appel, David, 304n
 Apple, Inc., 8, 207, 307–308, 395, 552
 Arc elasticity
 computation of, 70–72
 cross-price elasticity and, 86–87
 definition of, 94
 income elasticity and, 88
 revenue and, 80, 82
 Arcelor Mittal, 424
 Archer Daniels Midland Company (ADM), 394
 Ariely, Dan, 176n
 Asia, elasticities in, 91–92
 Asset specificity, 19
 AstraZeneca, 526
 Asymmetric information, 440, 443–453, 455
 Asymptotic curves, 481
 AT&T, 250, 329, 547
 AT&T Wireless, 255
 AU Optronics Corporation, 395
 Auld, M. Christopher, 85n
 Autocorrelation, 128–129
 Automobile industry
 export restraints on Japan, 105–106
 imperfect competition in, 372

Automobiles
 demand and, 154
 in lemons model problem, 443–446
 used, and market demand, 110–111
 Average cost, 257–258, 265–269.
 see also Long-run average cost; Short-run average cost; Weighted average cost of capital
 Average fixed cost (AFC), 255, 256, 282
 Average product (AP), 183–185, 212, 243, 245, 246–247
 Average total cost (AC or ATC)
 definition of, 256, 282
 long-run, 261, 261t, 262f, 263–264, 268f, 272f
 minimum point of, 257
 short-run, 265–266
 Average variable cost (AVC), 255, 256–257, 259f, 260, 282

B

B. F. Goodrich, 552
 Baker, Thomas L., 350n
 Baltic Beverages, 379
 Bank for International Settlements, 546
 Bank of America, 250, 547
 Bank One, 250, 547
 Banking industry
 mergers in, 547
 subprime loan financial crisis, 2007–2009, 542–546
 United Kingdom, case study, 300–301
 United States, case study, 301, 304–305
 Bargaining, 439–443
 contracts and, 439
 negotiation and, 440
 in oligopoly, 441–442
 tacit, 439, 440, 441, 456
 Barnes, Julian E., 109n
 Barometric price leadership, 396–397, 427
 Barrett, Jessica, 130n

- Barro, Robert J., 225n
 Baseline forecasts, 135
 BASF AG, 395, 423
 Basso, Alberto, 85n, 87n
 Baumol, William, 2–3, 221–222, 224, 376, 398, 398n
 Baumol model, of revenue
 maximization, 398–399, 427
 Baumol's Disease, 221–222, 224f
 Bears Stearns, 544
 Becker, Cary, 553n
 Beef, demand for, 106–108
 Beer industry, marketing in,
 378–379
 Bekö, Jani, 89n
 Benefit externalities, 538, 554
 Benzie, Robert, 406n
 Bernanke, Ben, 543
 Besank, David, 373n, 378n
 Besanko, David, 3n
 Bhattacharya, Prasenjit, 113n
 Bhavasi, Sam, 280
 Big data, 117
 Binary variables, 119
 Bing, 334
 BlackBerry, 207
 Blair, Roger D., 304n
 Blankmeyer, Eric C., 85n
 Block, Stanley, 470n, 491n
 Boeing, 277
 Boston, Massachusetts, demand
 for soft drinks in, case study,
 129–130
 Boston Beer Company, 378
 Bosworth, Barry P., 221n, 224n, 225f
 Bottling plants, cost estimation in,
 case study, 302
 Bowe, Christopher, 109n
 Box-Jenkins forecasting model, 151
 Boyle, Catherine, 424n
 Bradford, Julie, 378n
 Bradley, Michael, 550n
 Brazil, coffee glut in, 113
 Break-even analysis, 346–347,
 357, 486
 Break-even point, 347–348, 357
 Break-even revenue, 348–349, 357
 Brewster, Gary W., 90n
 BRIC countries, 14, 53–54, 513
 Brickley, James A., 3n
 Brigham, Eugene F., 35n,
 474n, 507n
 Briloff, Abraham, 33n
 British Telecom (BT), 547
 Brookings Institution, 221n
 Brooks, Diana D., 394, 395
 Brown, C., 102n
 Brownell, Kelly D., 85n
 Bruner, Robert F., 549–550
 Budgeting. *see* Capital
 budgeting
 Bueso, Alberto T., 158n, 493n
 Burns, G., 394n
 Burstein, Meyer L., 409n
 Bush, George W., 543
 Business analytics, 117, 125
 Business considerations, in foreign
 markets, 34–35
Business Cycle Indicator Handbook
 (Conference Board), 135,
 137, 138n
 Business risk, 30, 36, 478
 Bustillo, Miguel, 398n
 Butler, Eamonn F., 100n, 101n
 Butler, Richard J., 304n
 Buyers, number of, 42, 371
- C**
- Calculations
 cash flow discounting, 30
 of decision trees, 488–490
 expected value, of possible
 outcomes, 479–480
 gross up, 522
 of internal rate of return (IRR),
 466–467
 of simulations, 487–488
 standard deviation, 480–481
 third-degree price
 discrimination, 402–404
 trend line, 146
 Calculus formulas
 for Cobb-Douglas production
 function, 242–243, 245, 246
 for inputs, 243–244
 for marginal product (MP),
 242–247
 for marginal revenue, 83–84
 for monopoly, 344
 for output and pricing decisions,
 343–344
 for point elasticity, 74, 86, 88n
 for production functions,
 242–247
 for returns to scale, 240–241,
 243, 246
 for short-run cost function,
 291–294
 for total product (TP), 242–243
 California, price discrimination
 in, 406
 California Public Employees'
 Retirement System
 (Calpers), 28
 Call centers, production function
 of, case study, 209–210
 Callahan, Patricia, 109n
 Capacity
 average cost and, 265–269
 planning, 205–207
 Cap-and-trade, 540
 Capital
 cost of. *see* Cost of capital
 in production function,
 200–201
 productivity and, 225, 226f
 as a resource, 4, 6f, 9
 working, 472
 Capital asset pricing model (CAPM),
 474–475, 498
 Capital budgeting, 462–477
 under conditions of risk,
 483–484
 decisions, types of, 463
 definition of, 463, 498
 evaluation methods of, 464
 MR = MC rule and, 475–476
 in multinational corporations,
 521–522, 531
 in practice, 470
 real options in, 490–493
 real-life cases of
 international, 494
 Capital equipment, 258, 264
 Capital rationing, 477, 498
 Card, D., 102n
 Cargill, 394
 Carlsberg, 379
 Carlton, Dennis W., 374n
 Carnegie, Andrew, 137
 Carney, Bryan M., 553n
 Carr, Aoife, 113n
 Carroll, C., 525n
 Cartels
 definition of, 427
 in industrial organization, 375
 overview, 391–392
 price fixing by, 393–395
 real-life cases of, 423–424
 Carton-board industry, European
 cartel in, 423
 Caser, Raff, 424n
 Cash flows
 in capital budgeting decisions, 463n
 conventional, 469
 discounting of, 29–30
 free, 499, 507–508
 internal rate of return and,
 466–467
 in multinational capital
 budgeting, 522
 net present value and, 464–466
 overview, 470–471
 time value of money and, 464
 types of, 471–472
 Cassel, Andrew, 112n
 Catron, H. Alan, 155n
 Causal (explanatory)
 forecasting, 160
 Causal forecasting, 134, 153
 Cell phone industry, 207–209
 Celler-Kefauver Act (1950), 536
 Census Bureau, 366
 Census of Manufactures (Census
 Bureau), 203, 204, 366

- Certainty equivalents, 484, 485, 498
 Chaffin, Joshua, 395n
 Chang, M. K., 204n
 Changes in demand, 40, 56
 Changes in supply, 43, 56
 Changes in the quantity demanded, 40, 56
 Changes in the quantity supplied, 43, 56
 Charlier, Marj, 107n
 Chase Bank, 547
 Chase Manhattan, 547
 Chemical Bank, 547
 Chin, Lee, 203n
 China, production in, 210
 Chow, Gregory C., 154n, 204n
 Christensen, L. R., 301n
 Christie's, 394–395
 Chrysler, 548
 Chung, Chanjin, 90n
 Cialdini, Robert, 176n
 Cigarettes, demand for, and price, 113
 Cigliano, J. M., 85n, 89n
 Cingular Wireless, 250
 Citicorp, 547
 Citigroup, 547
 City centers, traffic control in, 109–110
 Clark, Jeffrey A., 301n
 Clayton Act (1914), 535–536
 Clean Air Act (1990), 539–540
 CM (Contract manufacturer), 308
 Coase, Ronald, 19, 22, 540–541
 Coase theorem, 540, 554–555
 Cobb, C. W., 198n, 199, 203
 Cobb-Douglas production function, 198–203
 calculus formulas for, 242–243, 245, 246
 Cost and, 289–290
 definition of, 212
 history of, 204n
 numerical example of, 201
 studies utilizing, 202–203
 Coca Cola Company, 14, 394, 552
 Coefficient of determination (R^2), 123, 159
 Coefficient of variation, 482–483, 498
 Coefficients
 cross-elasticity, 407n
 price elasticity, 70–71
 regression, 120–122, 125
 standard error of, 122, 127
 Coffee, demand for, 154–155
 Coffee industry, supply and demand in, 113
 Cohen, Laurie P., 395n
 Coincident indicators, 137
 Coinsurance, 449, 455
 Colbert, R., 559n
 Collateralized debt obligations (CDOs), 544
 Collusive agreements. *see* Cartels
 Comitimi, Salvatore, 202n
 Command process, 7–8, 16
 Compaq, 306–307, 308, 309
 Comparative statics analysis, 46–53
 Compensation packages, 451
 Competition. *see also* Imperfect competition; Monopolistic competition; Nonprice competition; Oligopoly; Perfect competition
 degree of, 313
 deregulation and, 546–547
 government maintenance of, 534
 managerial decision-making and, 53
 market types and, 315–316, 362, 362t
 measures of, 313–314
 Porter's model of, 376–378
 risk and, 478
 Complementary goods, 44, 86, 94, 371, 416–417
 Compound growth rate projection, 139–140, 160
 Conduct, in industrial organization, 374–375
 The Conference Board, 135, 136, 137, 138–139
 Conflict, strategy of, 433–443
 Consolidations, as cost-cutting measure, 278
 Constant returns to scale (CRTS), 194–195, 243, 263, 282
 Consumer behavior
 measurement of, 116–117
 risk of change in, 478
 Consumer Confidence Index, 136
Consumer Confidence Survey (Conference Board), 136
 Consumer demand, 108
 Consumer panel surveys, 117
 Consumer surveys, 116–117, 160
 Consumers
 preferences of, 478
 protection of, 536
 spending intentions of, 136
 Contestable markets, theory of, 376
 Continental Airlines, 250, 547
 Continuous distribution, 481
 Contract manufacturer (CM), 308
 Contracts and bargaining, 439
 Contribution margin, 324–325, 326f, 338
 Cooper, John C. B., 85n
 Cooperative games, 439, 455
 Coors, 379
 Coot, Ronald S., 298n
 Copyrights, 551–552
 Corn, demand for, 113
 Corporate raiders, 451
 Corporations. *see* Firms;
 Multinational corporations
 Correia, C., 495n
 Cost curves, 259–260, 265–267, 291–300, 302
 Per-unit costs in Excel, 292–294
 Cost efficiency, in short-run cost function, 257–258
 Cost externalities, 538–539, 555
 Cost function. *see* Long-run cost function; Short-run cost function
 Cost leadership, 279, 377
 Cost of capital
 definition of, 465, 498
 in multinational capital budgeting, 522
 overview, 472
 weighted, 475, 499
 Costco, 207n
 Cost-plus pricing, 411–413, 427
 Costs. *see also* Specific type of costs, e.g., Fixed cost, Opportunity costs
 definition of, 251
 diminishing returns and, 253–254
 economic, 317
 economic versus accounting, 33–34, 295–296
 estimation of, 282, 295–305
 importance of, 249–251
 incremental, 283
 internal versus external operating, 20–22
 minimization of, 239–240
 as nonprice determinant of supply, 43–44, 47t
 opportunity, 3, 6, 16, 283
 production and, 253–254
 productivity vs., 221
 as reason for mergers, 548
 reduction of, as a competitive edge, 276–279
 relevant, 283
 timing of, 296
 transaction, 19–22, 37, 298
 types of, 251–252
 Countrywide Mortgage, 545
 Cournot, Antoine-Augustin, 433n
 Coutinho, Maria Cristina, 494n
 Cox, James E., Jr., 151n
 Cradock, Angie I., 130n
 Crane, Alan, 209n
 Credit card industry, imperfect competition in, 372
 Credit default swaps (CDSs), 544
 Critical F-value, 123
 Critical t-value, 122, 160
 Croatia, capital budgeting in, 494
 Croft, Jane, 407n

- Cross, Robert G., 407n
 Cross-elasticity. *see* Cross-price elasticity
 Cross-price elasticity, 86–87, 94
 Cross-sectional analysis
 long-run cost function and, 295, 299–300
 time-series analysis vs., 204
 when to use, 201
 Cross-sectional data, 118, 160
 Cubic production function, 196, 197f, 245–247
 Cultural factors, 517, 530
 Currency options, 519–520
 Currency swaps, 520
 Cut-off rate, 465
 Cyclical fluctuations, in time series analysis, 143, 147
- D**
- Daimler Benz, 548
 Dairy farms, case study, 302
 D’Amico, Michael F., 350n
 Daniels, J. D., 557n
 Danielson, Morris G., 470n
 Data. *see* Cross-sectional data; Time series data
 Data mining, 116–117
 Daves, Phillip R., 35n
 David, Edmond, 301n
 Davis, Donna F., 151n
 Dean, Joel, 2, 295, 296, 297, 298, 300
 Debt, 473
 Decision making. *see* Managerial decision making
 Decision trees, 488–490, 498
 Decisions
 capital budgeting, 463
 investment, 463
 optimal, 22–23, 36
 production function and, 205–207
 regression analysis and, 125
 Decreasing returns to scale (DRTS), 194–195, 243, 263, 282
 Dedi, Lidija, 494n
 Deductibles, 449, 455
 Deere, D., 102n
 Deflators, 223
 Degree of operating leverage (DOL), 350–353, 357
 Degrees of freedom (df), 122, 160
 Dell, Michael, 280, 306
 Dell Inc., 279–280, 306–307, 395
 Deloitte, 54f
 Delphi method, 161
 Delta Airlines, 250, 396, 547
 Demand
 asymmetric information and, 443–446
 break-even analysis and, 350
 complementary goods and, 416–417
 cross-price elasticity of, 86–87
 definition of, 39, 56
 elasticity and, 90
 guiding function of price and, 48–50, 51f
 joint products and, 417–420
 law of, 39–40
 market, 39–42
 market equilibrium and, 44–45, 47f, 48f
 market power and, 53
 price elasticity of. *see* Price elasticity of demand
 rationing function of price and, 47–48, 51f, 52t
 real-life cases in, 105–113
 scarcity and, 5–9
 substitute goods and, 417
 Demand curves
 cost-plus pricing, 412–413
 demand estimation and, 126–127
 elasticity and, 72–75, 77
 kinked, 367–369, 380, 433n
 in monopolistic competition, 364–365
 perfectly elastic, 319
 short-run versus long-run, 79–80
 Demand elasticity. *see* Price elasticity of demand
 Demand estimation
 identification problem and, 126–127
 point elasticity and, 74
 white Zinfandel in Los Angeles, case study, 167–174
 Democratic Party, 545
 Demographics, 374n
 Denesuk, Matthew, 220n, 229n
 Denso Corporation, 395
 Dependent variables, 119
 Depreciation, 200–201, 471
 Deregulation, 546–547
 Derived demand, 79, 94, 190–191
 Differential pricing. *see* Price discrimination
 Diminishing returns
 law of, 185–188, 212
 marginal cost (MC) and, 253–254
 marginal product (MP) and, 196–198
 in production function, 196–199
 in total cost function, 258–260
 Diocletian, edict of, 100n
 Direct costs. *see* Out-of-pocket costs
 Discount rate
 definition of, 465, 499
 in multinational capital budgeting, 522
 risk-adjusted, 484–485, 499
 Discounted payback method, 464n
 Discrete distribution, 481
 Discrimination, as risk for
 multinational corporations, 493.
 see also Price discrimination
 Diseconomies of scale, 263–265, 266f, 282
 Disney Corporation, 552
 Distributions, frequency, 481–482
 Dividend growth model, 30–31, 473–474, 499
 Dividend payments, 523
 Dixit, A. K., 490n
 Dodd–Frank Wall Street Reform and Consumer Protection Act (2010), 545–546
 Doh, J. P., 557n
 Dolan, Matthew, 395n
 Dollar General Corporation, 397
 Dominant price leadership, 397–398, 427
 Dominant strategy (DS), 435, 455
 Dominant strategy equilibrium (DSE), 435–437, 455
 Douglas, P. H., 198n, 199, 203, 204
 Dow, William H., 85n
 Downsizing, as cost-cutting measure, 278
 Dranova, David, 373n
 Dressings industry, global market expansion of, 511–513
 Drucker, Peter, 2n, 9
 Duchesne, Erich, 129n
 Dummy variables, 119
 Durand, David, 204
 Durbin-Watson test, 129
 Dutta, Devangshu, 208n
 DWK/Saarstahl, 424
 Dynamic random access memory chips (DRAMs) industry, price-fixing in, 395, 424
- E**
- Eagles, Munroe, 129n
 Earnings, reinvestment of, 523
 Echikson, William, 111n
 Econometric forecasting models, 153–156, 160
 Economic conditions, as business risk, 478
 Economic cost, 33–34, 36, 295–296, 317, 338
 Economic decisions for the firm, 8, 16
 Economic goals, 22–25
 Economic indicators, 136–139, 160
 Economic loss, 318, 324, 325f, 338
 Economic profit, 33–34, 317–318, 324–327
 competition and, 313–314
 definition of, 34, 36, 338

- Economic Value Added (EVA), 32–33, 36
- Economics, 2, 7–9, 16. *see also* Managerial economics
- Economies of scale, 263–265, 273, 282, 301–302, 304
- Economies of scope, 272–273, 282
- Economy
 global, and U. S. financial crisis, 2007–2009, 546
 knowledge-based, 551
 stabilization of, 541–542
- Efficient scale of production, 263
- Ehrhardt, Michael C., 507n
- Eichenwald, Kurt, 394n
- Eiteman, David K., 493n, 494n, 510n
- Elasticity. *see also* Advertising elasticity; Income elasticity; Point elasticity; Supply elasticity
 categories of, 74
 computation of, 70–72
 definition of, 69, 94
 determinants of, 75–77
 interest rates and, 90
 mathematics of, 83–84
 population size and, 90
- Electric utilities industry, case study, 301, 302
- Elevator industry, European cartel in, 424
- Employment, 102
- Endogenous variables, 155–156
- Energy Merchant Corporation, 406
- Engineering cost estimating, 282, 303
- Enterprise resource planning (ERP), 278
- Entrepreneurship, 9
- Environmental protection
 equipment, 463
- Equilibrium. *see* Dominant strategy equilibrium (DSE); Market equilibrium; Nash equilibrium; Separating equilibrium
- Equilibrium price, 45, 46–51, 56
- Equilibrium quantity, 45, 46–51, 56
- Equity, cost of, 473–474
- Erfle, Stephen, 129
- ERP (enterprise resource planning), 278
- Estimation. *see also* Demand estimation
 of costs, 282, 295–305
 in long-run cost functions, 282, 299–305
 of production functions, 196–205
 of short-run cost functions, 282, 295–299
 statistical, of production functions, 200–205
- Etter, Lauren, 112n, 113n
- European Central Bank, 546
- European Commission, 423–424, 537–538, 553
- European Court of Justice, 407
- European Union, 334
- EVA Dimensions LLC, 32, 33
- EVA Momentum, 33
- Evening Standard (London), 109
- Exchange rate, 156–158, 518, 521, 531
- Exchange rate hedging, 518–519, 531
- Exchange rate risk, 517, 531
- Excise taxes and product prices, 102–105
- Executive compensation, 28–29, 399
- Executive opinion, 135
- Exogenous variables, 156
- Expectations. *see* Future expectations
- Expectations Index, 136
- Expected value, 479–480, 499
- Experience curve, 272, 282
- Expert opinion, 135
- Explanatory forecasting, 134, 153
- Explanatory variables, 119
- Explicit costs. *see* Out-of-pocket costs
- Exponential smoothing, 134, 147, 149–151, 152, 160
- Exports
 elasticity and, in Asia, 91–92
 restraints on Japan, 105–106
- Expropriation, 493
- Externalities. *see* Benefit externalities; Cost externalities; Markets, externalities of
- ## F
- Facebook, 334
- Facilities, expansion of, 463
- Fackler, Martin, 335n
- Fannie Mae, 545
- Favole, Jared, 407n
- Federal Funds, 541
- Federal Reserve. *see* U. S. Federal Reserve
- Federal Trade Commission Act (1914), 536
- Federal Trade Commission (FTC), 106n, 406
- Ferdows, K., 208n
- Fernández-Kranz, Daniel, 89n
- Fickers, Karin, 130n
- Filippini, Massimo, 302n
- Financial derivative products, 544
- Financial risk. *see* Business risk
- Financing
 cost of capital and, 472
 terms and conditions, 371
- Firer, Colin, 485n
- Firms
 definition of, 19, 36
 determining value of, 507–508
 economic decisions for the, 8
 economic goals of, 22–25
 multinational, 34–35
 noneconomic objectives of, 25
 valuation of, 31–33
- Firm, David, 525n
- First-degree price discrimination, 400
- Fiscal policy, 541–542
- Fisher, Franklin M., 537
- Five Forces model, 376–378, 380
- 5 S's, of lean manufacturing, 208
- Fixed asset reductions, as cost-cutting measure, 278–279
- Fixed cost, 273, 399. *see also* Average fixed cost (AFC); Total fixed cost (TFC)
- Fleet Bank, 547
- Flexibility, 490–491
- Flextronics, 309
- Flotation costs, 474
- Flow variables, 200
- Flows. *see* Cash flows
- Fluctuations, as business risk, 478
- Focal points, 441, 455
- Focus groups, 116–117
- Focus List (Calpers), 34
- Forecasting techniques
 econometric model, 153–156, 160
 least squares projection, 127, 142, 146, 155
 multiple-equation model, 155
 projections, 139–151
 single-equation model for, 155
 smoothing, 147–151
 types of, 133–135
- Forecasts and forecasting
 categories of, 132–133
 demand and, 133
 economic indicators and, 136–139
 error trends in, 152
 exchange rates, 156–158
 opinion as basis for, 135–136
 overview, 132
 in pharmaceutical industry, 152
 planning production and, 267–269
 prerequisites for, 133
 regression analysis and, 155
 sales, accuracy study of, 151–152
 surveys and, 137
- Foreign direct investment (FDI), 520–521, 531
- Foreign tax credits, 522
- Forster, Julie, 394n
- Forward markets, 519

France, Mike, 553n
 France, wine industry in, 111
 Freddie Mac, 545
 Free cash flow, 499, 507–508
 Freeman, Sholem, 111n
 Frey, B. S., 102n
 Friedman, Milton, 79n, 101n, 416n
 Friess, Steve, 407n
 F-test, 123, 160
 Furniture Brands International, Inc., 112
 Furniture industry, in United States, 111
 Furukawa Electric Company, 395
 Future expectations
 nonprice competition and, 371
 as nonprice determinant of demand, 42, 47t
 as nonprice determinant of supply, 44, 47t
 Future products, 491
 Futures markets, 519
 F-value, 123–124

G

Galbraith, Craig S., 203n
 Game theory, 433–443
 Games, 434, 455
 Gapenski, Louis C., 507n
 Gapper, John, 55n
 Garcia, A., 89n
 Gasoline industry, zone pricing in, 406
 Gattorna, John, 208n
 GE Capital Aviation Services (GECAS), 553
 General Electric Company (GE), 391, 393, 536, 553
 Genericide, 552
 Ghose, Anindya, 87n
 Gil, J. M., 89n, 131n
 Gilbert, G., 102n
 Gilroy, C., 102n
 Giordano, James N., 304n
 Giovino, Gary, 85n
 Gitman, Lawrence J., 471n
 Glaister, Stephen, 85n, 89n
 Glaxo Wellcome, 250
 GlaxoSmithKline (GSK), 250, 525
 Global applications
 beer market, 378–379
 decline of cartels in Europe, 423–425
 entering foreign markets, 526–527
 forecasting exchange rates, 156–158
 political risks, 493
 price elasticity of demand, 91–92
 pricing and output decisions in, 335–336

production functions, 210
 profit maximization hypothesis, 34–35
 shifting trends in global outsourcing, 210
 supply chain management, 280–281
 Global Transfer Pricing Survey, 526
 GM, 208–209
 Goals. *see* Economic goals
 Goel, Rajeev K., 85n
 Golden parachute, 451, 455
 Golicic, Susan L., 151n
 Gompertz curve, 152, 153f
 Google, 334, 552–553
 Gordon, Myron J., 473
 Gordon, T. J., 135n
 Gordon model, 473–474
 Gortmaker, Steven, 130n
 Government
 antitrust laws and, 535–538
 deregulation and, 546–547
 economy stabilization and, 541–542
 intellectual property and, 551
 intervention by, 538–541
 in market economy, functions of, 534–535
 policy and time lags, 542
 Government sponsored enterprises (GSEs), 545
 Gracia, A., 131n
 Graham, Daniel J., 85n, 89n
 Graham, John R., 470n
 Gramm-Leach-Bliley Act (1999), 542–546
 Grant, Linda, 130n
 Greenberg, Carol, 354n
 Greene, W. H., 301n
 Greenspan, Alan, 545
 Gronroos, C., 223
 Groothoff, Frederike, 130n
 Gross up calculation, 522
 Growth accounting, 224–225
 Growth options, 490
 Growth-maximization model, 399
 Gruber, Jonathan, 85n
 GSK (GlaxoSmithKline), 250
 Guerrero, Francesco, 423n
 Guiding function of price, 48–50, 51f, 52t, 56

H

Hall, R. L., 411n
 Halteren, Hans van, 130n
 Halvorsen, R., 87n
 Hamadeh, Mohamad, 85n, 89n
 Hanke, John E., 133n, 135n, 151n
 Hanks, Richard D., 407n
 Harden, Blaine, 335
 Harding, Riyon, 551n

Harrison, Crayton, 280n
 Hart-Scott-Rodino Act (1976), 536
 Harvey, Campbell R., 470n
 Hawes, Jon M., 350n
 Heckman, M. A., 18n
 Hedging, 518–519, 531
 Heien, D., 85n
 Heilbrun, James, 221n
 Heineken, 379
 Helmer, O., 135n
 Henkel, 424
 Herfindahl-Hirschman (HH) index, 367
 Hertz, David B., 488n
 Hewlett-Packard Company, 395, 526
 Hilsenrath, Jon E., 112n
 Hilzenrath, David S., 525n
 Historical costs, 34, 251, 282–283
 Hitch, C. J., 411n
 Hodgetts, R. M., 557n
 Hon, Mark T., 89n
 Honeywell, 536, 553
 Hoof, Lian van, 130n
 Hormel Foods, Inc., 108
 Hosiery mill industry, case study, 297–298
 Hotel industry, price discrimination in, 407–409
 Hotten, Russell, 424n
 Houston, Joel F., 474n
 Houthakker, H. S., 85n, 89n
 Hsieh, Chee-Ruey, 87n, 132n
 Hsieh, Edward T., 85n
 Hsieh, Rajeev K., 89n
 Hu, Teh Wei, 87n, 132n
 Huang, Chung I., 89n
 Huang, Cliff J., 154n
 Huang, David S., 202n
 Huettner, David A., 301n
 Huffington Post, 335n
 Hurdle rate, 465
 Hurwich, Michael, 368n
 Husten, Corinne, 85n
 Hynix, 424

I

IBM
 antitrust lawsuit against, 333–334
 price leadership by, 369, 396
 supply chain management at, 307
 tying arrangements and, 409
 IBP, Inc., 108
 Identification problem, in regression analysis, 126–127
 Imperfect competition
 definition of, 361–362
 managerial decision-making and, 53
 nonprice competition, 369–371

- reality of, 371–372
strategy in, 372–378
- Imperfect substitutes, 233
- Implicit costs, 34, 251. *see also*
Opportunity costs
- Imports
elasticity and. in Asia, 91–92
quotas on, 106
- InBev, 379
- Incantalupo, Tom, 406n
- Incentives, 20, 25, 451–452
- Income
demand and, 41
market segmentation and,
370–371
as nonprice determinant of
demand, 47t
redistribution of, 534
- Income elasticity, 87–89, 94
- Increasing returns to scale (IRTS),
193–195, 243, 263, 282
- Incremental costs, 252, 283
- Incremental pricing, 414–415
- Independent projects, 467, 499
- Independent variables, 119, 127
- Index of present value, 464n
- Indexed stock options, 452
- India
call centers in, 210
Kellogg Company in,
526–527
key industries in, 513–515
McDonald's in, 513, 516–517
poverty in, 515
- Indicators, economic, 136–139
- Indirect command, 7
- Indirect costs. *see* Opportunity costs
- Industrial organization, 374–376
- Inferior goods, 88, 94
- Infineon, 424
- Inflation rates, 521
- Inflows. *see* Cash flows
- Information, asymmetric, 443–453
- Information Resources, Inc.
(IRI), 116, 167
- Information technology (IT), cost
reductions and use of, 278
- InfoScan Reviews, 167
- Initial cash flows, 471
- Initial public offerings (IPOs),
507–508
- Inkpen, A., 557n
- Inputs. *see also* Resources
alternative terms for, 183
calculus formulas for, 243–244
call center, case study utilizing,
209–210
Cobb-Douglas production
function and, 199
definition of, 5, 16, 181, 212
in long-run analysis, 183, 196n
multiple, 191–193, 231–241
- optimal level of, 190–191,
236–241
in production function,
181–185, 193
substitution of, 232–236
- Institute of Business Forecasting
and Planning, 134, 152
- Insurance industry
case studies, 304
moral hazard problem in,
450, 452
screening in, 448–450
- Intel, 307
- Intellectual property, 551–553
- Interbrew, 379
- Interdependency, in game
theory, 442
- Interest rates
elasticity and, 90
riskless, 474
- Internal operating costs, 20
- Internal rate of return (IRR)
calculation of, 466–467
in capital budgeting practice,
470, 476
definition of, 464, 499
modified, 469n
net present value versus, 467–469
- Internal Revenue Service (IRS),
525–526
- International Air Transport
Association (IATA), 391
- International Coffee Growers
Association, 314
- International Monetary Fund
(IMF), 546
- International Monetary Market, 519
- Internet, impact of, on transaction
costs, 22
- Inverse elasticity rule, 377
- Investments, 463–464, 472
- Irregular patterns, in time series
forecasting, 143
- Isocost curves, 212, 237–239
- Isoquant curves, 212, 231, 232f
- Issenberg, Sasha, 335
- Iwata, Edward, 395n
- J**
- Jabarin, Amer S., 87n
- Jackson, A., 33n
- Jaffe, Jeffrey, 469n
- Jain, Chaman L., 135n, 152n
- James, R. Warren, 300n
- Japan
automobile export restraints
on, 105
cigarette sales in, 113
use of learning curve in, 271
- Jarman, Max, 406n
- Jennen, Birgit, 423n
- Jensen, Elizabeth J., 535n
- Jensen, Michael C., 27n, 411n
- JetBlue, 396
- Jobs and Growth Tax Relief
Reconciliation Act (2003), 542
- Johnson, Jo, 111n
- Johnston, J., 298, 300–301
- Joint products, 417–420
- Joste, Ruan, 526n
- JP Morgan, 547
- JP Morgan Chase Bank, 250, 547
- Jury of executive opinion, 135, 160
- K**
- Kahn, Kenneth B., 151n
- Kaiser, Harry A., 90n
- Kaizan, 208
- Kapoor, A., 513n
- Katz, I. F., 102n
- Kaufman, Leslie, 109n
- Keat, Paul G., 303n
- Keating, Barry, 151n
- Kellogg Company, 526–527
- Kendall, Brent, 395n
- Kershaw, Sarah, 365
- Khoueiri, Roy, 85n, 89n
- Kilman, Scott, 108n, 394n
- Kinked demand curves, 367–369,
380, 433n
- Kishore, Dash, 517n
- Kleenex, 552
- Knowledge-based economy, 551
- Knox, Kris Joseph, 85n
- Koenigsaecker, George, 208n
- Kohen, A., 102n
- Kollar-Kotelly, Colleen, 537
- Koshal, Manjulika, 87n
- Koshal, Rajindar K., 87n
- Kranhold, Kathy, 395n
- Krueger, A. B., 102n
- Kulatilaka, N., 490n
- L**
- Labor
cost of, 190–193, 200
productivity and, 227–230
as a resource, 4, 9
- Labor costs, 269–272
- Labor productivity (LP), 194,
220–222, 224, 226f,
227–230
- Lagging indicators, 137
- Lampel, Joseph, 376n
- Lance, Peter M., 85n
- Land, 4, 9
- Land, William, 440
- Landon, John H., 301n
- Lanfranco, Bruno A., 89n
- Latta, Michael, 152n

- Laux, Judith A., 302n
 Lavalife, 406
 Law of demand, 39–40, 56
 Law of diminishing returns, 185–188, 212. *see also* Diminishing returns
 Law of marginal rate of technical substitution, 234
 Lawler, Edmund O., 276n, 363n
 Layoffs, as cost-cutting measure, 278
 Lazich, Robert S., 366n
 Le Quotidien (French newspaper), 109
 Leading indicators, 137–139
 Lean approach, to manufacturing, 208–209
 Learning curve, 269–272, 283
 Learning curve effect, 269, 271–272
 Leasing decisions, 463
 Least squares projection technique, 127, 142, 146, 155
 Lego, 552
 Lehman Brothers, 544
 Lehrer, Jonah, 176n
 Lemons model, 443–446
 Lenz, John, 90n
 Leonard, Devin, 395n
 Lerner, Abba, 377
 Lerner Index, 377
 Leunig, Tom, 22n
 Leverage. *see* Financial risk
 Levitt, Theodore, 227n
 Levy, Haim, 483n
 Lewis, M., 208n
 Li, Kui-Wai, 204n
 Li & Fung, 280–281
 License fees, 523
 Limit pricing, 422, 427
 Limited war, concept of, 441–442, 455
 Lin, Chien-Fue Jeff, 87n, 132n
 Linear cost function, 259, 260
 Liquidity, 544–545, 546
 Loh, Chung-Ping, 85n
 London, traffic control in, 110
 Long, Michael W., 85n
 Long run
 definition of, 50, 56
 economies of scale in, 273
 elasticity, 79–80
 market changes and price in, 48–52
 Long run (market analysis), 313, 327, 338
 Long-run average cost (LRAC), 261t, 262f, 263–264, 267–269, 272f
 Long-run cost function, 260–269
 estimation in, 282, 299–305
 short-run cost function versus, 263
 studies illustrating, 300–303
 Long-run marginal cost (LRMC), 261–262, 264n
 Long-run monopolistic competition equilibrium (LRMCE), 364, 387–389
 Long-run production function, 182, 193–195, 212
 Los Angeles Times (newspaper), 109
 Loss. *see* Economic loss; Profit maximization hypothesis
 Lovell, C. A. Knox, 302n
 Lovelock, Christopher H., 276n
 Lowy, Alex, 22n
 Loyalty programs, 370
 Lucent, 371
 Lundegaard, Karen, 111n
 Luthans, F., 557n
- M**
- Macluca, J., 208n
 Macroeconomics, 5
 Make or buy decisions, 463
 Makridakis, Spyros, 132n, 133n, 135n
 Management
 definition of, 2
 of personnel, 8–9
 Peter Drucker on, 9
 shareholder conflict with, 26–29
 Management control, 11n
 Management skills, 9
 Managerial decision making
 competition and, 53
 cost reduction and, 278–279
 costs and, 249–251
 monopoly markets and, 333
 optimization analysis and, 433
 perfect competition and, 53, 333
 production function and, 205–207
 resource allocation and, 3–4, 8–9
 Managerial economics
 application to business disciplines, 4
 definition of, 2, 16
 economic terms and concepts in, 5–9
 optimal decisions in, 23
 overview, 2–3
 relationship to microeconomics, 3
 strategy and, 372–378
 Managerial utility function, 450–451
 Managers
 labor market for, 28
 market environment and, 53
 Manufacturing, lean, 208–209
 Marchant, Mary A., 87n
 Marcus, A. J., 490n
 Marginal cost (MC)
 average variable cost and, 260
 definition of, 256, 283
 diminishing returns and, 253–254
 long-run, 261–262, 264n
 marginal revenue rule, 239
 optimal decisions and, 23
 shape of cost curve for, 298–299
 in short-run cost function, 257
 Marginal factor cost (MFC), 191n, 212. *see also* Marginal labor cost (MLC)
 Marginal labor cost (MLC), 190t, 191, 212. *see also* MRP = MLC rule
 Marginal pricing, 413–414
 Marginal product (MP)
 calculus formulas for, 242–247
 in Cobb-Douglas production function, 198–200
 definition of, 212
 diminishing returns and, 196–198
 multiple inputs and, 191–193, 235–237
 short-run analysis of, 183–186
 Marginal rate of technical substitution (MRTS), 212, 233–236
 Marginal resource cost (MRC). *see* Marginal labor cost (MLC)
 Marginal revenue (MR), 23, 81–84, 94
 Marginal revenue product (MRP), 190t, 191, 212, 239, 243–244
 Marginal revenue-marginal cost. *see* MR = MC rule
 Marginal utility of money, 483n
 Market concentration, 365–367, 375–376
 Market demand, 39–42, 56
 Market entry and exit, 313, 316, 327–328, 362
 Market environment and managers, 53
 Market equilibrium
 comparative statics analysis of, 46–53
 demand and, 44–45, 47f, 48f
 rationing function of price and, 48f
 short-run versus long-run adjustments and, 50–51
 supply and, 44–45
 Market power, 53
 definition of, 56, 338
 market types and, 312, 313
 Market process, 7–8, 16
 Market research, 136
 Market segmentation, 9, 370, 407
 Market share and price fixing, 391
 Market structure, 312, 338
 Market supply, 42–44
 Market types

- business examples of, 314–315
 - competition and, 315–316, 362
 - Market Value Added (MVA), 31–32, 36
 - Marketing mix, 370–371
 - Markets
 - analysis of changes in, 46–53
 - externalities of, 538–541
 - informational asymmetry in, 443–453
 - for protecting against risk, 519–520
 - Markon, Jerry, 395n
 - Markup, 411–412
 - Marsh, Peter, 22n, 34n
 - Marshall, Alfred, 65, 78
 - Marston, Stephen C., 306
 - Martin, Andrew, 113n
 - Martin, Gina, 85n
 - Martinez, C., 560n
 - Martins, Carlos V., 494n
 - Mason, Christopher, 395n
 - Mastercard, 372
 - Material costs, reductions in, 277
 - Mathematical solutions
 - cost-plus and marginal pricing, 413–414
 - of elasticity and revenue, 83–84
 - profit maximization, 416n
 - Mathias, F. John, 542
 - Maximization. *see* Profit
 - maximization hypothesis;
 - Revenue, maximization of;
 - Wealth maximization
 - Maximum efficient scale (MaxES), 263
 - MBNA, 250
 - McCarthy, Teresa, 151n
 - McCartney, Scott, 424n
 - McConnell, Campbell, 2n
 - McCormick & Co., 406
 - McDonald's, 513, 516–517
 - McGregor, Jena, 28n
 - McHugh, Anne, 130n
 - McMillan, John, 440n
 - McNulty, Sheila, 109n
 - McRae, Robert, 85n, 89n
 - Meckling, William H., 3, 27n, 411n
 - Mefford, Robert N., 203n
 - Menzer, John T., 151n
 - Merchant, Mary A., 85n
 - Merck & Co., 525–526
 - Mergers. *see also* Acquisitions
 - in airline industry, 547
 - in banking industry, 547
 - as cost-cutting measure, 278
 - reasons for, 548–550
 - in telecommunications industry, 547
 - Microeconomics, 3, 5, 375
 - Microsoft Corporation, 371, 526, 552
 - anti-trust cases against, 333–334, 536–538
 - as example of a monopoly, 315
 - Miller Brewing, 379
 - Milliman, N., 394n
 - Milunovich, S., 33n
 - Minimum efficient scale (MES or MinES), 263
 - Minimum required rate of return, 465
 - Minimum wages, 101–102
 - Minneapolis, Minnesota, traffic control in, 109
 - Mintz, Ilse, 106, 106n
 - Mintzberg, Henry, 376n
 - MNCs. *see* Multinational corporations
 - Mobile device industry, 207–209
 - Modified internal rate of return (MIRR), 469n
 - Moffett, Michael H., 493n, 494n, 510n
 - Mohd Noor, Zaleha Bt, 203n
 - Molsen, 379
 - Molsen Coors Brewing Company, 379
 - Monetary policy, 541
 - Monopolistic competition
 - business examples of, 315
 - definition of, 380
 - demand curves in, 364–365
 - as imperfect competition, 361–362
 - long-run monopolistic competition equilibrium (LRMCE), 387–389
 - market characteristics of, 312f
 - market entry and exit and, 313
 - market power and, 312, 313
 - MR = MC rule and, 363–364
 - overview, 363–365
 - Monopoly
 - calculus formulas for, 344
 - natural, 535, 555
 - Monopoly markets, 329–332
 - business examples of, 315
 - characteristics of, 312f, 313f
 - definition of, 338
 - managerial decision making and, 333
 - MR = MC rule in, 329–332, 338
 - outputs in, 329
 - price discrimination in, 399–401, 410
 - pricing in, 329
 - Monopsony, 555
 - Mont-Ferguson, Helen, 130n
 - Monti, Mario, 334
 - Moore, T. G., 102n
 - Moral hazard problem, 450, 452, 455
 - Morgan Stanley, 397
 - Morgenstern, Oskar, 483n
 - Moroney, John R., 204n
 - Morris, Betsy, 108n
 - Mortgage industry, 543
 - Mosheim, Robert, 302n
 - Motorola, 207
 - Motorola Mobility, 552
 - Moving average method, 145, 147–149, 151, 160
 - MR = MC rule
 - capital budgeting and, 475–476
 - in Five Forces model, 376–377
 - in graphs, 323, 324f
 - monopolistic competition and, 363–364
 - in monopoly markets, 329–332, 338
 - multiple inputs and, 239
 - optimal output levels using, 321–324
 - MRP = MFC rule, 191, 212
 - MRP = MLC rule, 191, 212
 - Muda, 208
 - Multicollinearity, 127–128
 - Multifactor productivity (MFP), 220, 221, 225, 226f, 229, 230f
 - Multinational capital budgeting, 521–522, 531
 - Multinational corporation risk, 517–518, 531
 - Multinational corporations
 - budgeting in, 521–522
 - Kellogg Company, case study, 526–527
 - opportunities and challenges for, 510–517
 - transfer pricing in, 523–526
 - Multinational transfer pricing, 523–526, 531
 - Multiple inputs, working with, 191–193, 231–241
 - Multiproduct pricing, 415–420, 427
 - Murphy, K. M., 102n
 - Music industry, pricing and demand in, 112
 - Mutual dependence recognized, 440
 - Mutual interdependence, 362, 367, 380–381
 - Mutual interdependence and game theory, 440
 - Mutually exclusive projects, 467–468, 499
- ## N
- Naive forecasting, 134, 152, 160
 - Nash, John, 435–436
 - Nash equilibrium, 435–437, 440–441, 455
 - Nassauer, Sarah, 111n
 - National Association of Purchasing Agents, 136

- National Bureau of Economic Research, 137
- National Cattlemen's Association, 107
- National Electric Manufacturers Association, 393
- National Restaurant Association, 354
- Nations Bank, 547
- Natural monopoly, 535, 555
- Negotiation and bargaining, 440
- Nelson, Edward T., 89n
- Nelson, Michael A., 85n
- Net present value (NPV)
 calculation of, 464–466
 definition of, 499
 internal rate of return versus, 467–469
 in practice, 470
 risk and, 483–484
- Netscape Communications Corporation, 537
- Nevada Equal Rights Commission, 407
- New York City, traffic control in, 110
- New York Post (newspaper), 109
- Newman, Maria, 110n
- Newspapers, price elasticity of demand and, 109
- Ng, Y. C., 204n
- Niisu, Izumi, 335–336
- Nkwenti-Zamcho, Emmanuel, 203n
- Nokia, 207
- Noland, R. Paul, 407n
- Noncash investments, 472
- Non-constant-sum games, 439
- Noncooperative games, 439, 440, 455
- Noneconomic objectives, 25, 36
- Nonmarginal pricing, 410–415
- Nonprice competition, 314, 362, 369–371
- Nonprice determinants of demand, 41, 47t, 57, 370
- Nonprice determinants of supply, 43–44, 47t, 57
- Nonprice variables, 370
- Non-zero-sum games, 439
- Normal curves, 481
- Normal profit, 34, 36, 317, 324–327, 338
- Nortel, 552
- North American Industry Classification System (NAICS), 366
- Northwest Airlines, 250, 396, 547
- Nugent, Anne, 378n
- O**
- Obama, Barack, 545, 546
- O'Connor, Dennis J., 158n, 493n
- Ocunade, Albert A., 301n
- ODM (Original design manufacturer), 308
- OEM (Original equipment manufacturer), 308
- Offshoring, 21–22. *see also* Outsourcing
- Ogier, Thierry, 113n
- Oil
 price analysis, 52–53
 supply and demand for, 53–54
- Ojasalo, K., 223
- Okunade, A. A., 85n
- Oligopoly
 bargaining in, 441–442
 business examples of, 315
 characteristics of, 312f
 definition of, 381
 game theory and, 434
 as imperfect competition, 361–362
 market concentration, 365–367
 market power of, 312, 314
 price leaders in, 369
 price leadership in, 396
 pricing in, 367–369
 prisoners' dilemmas in, 437–439, 454
 revenue maximization, 398
- Oliver, R. K., 306n
- Olsen, Elizabeth, 38n
- O'Neill, Jim, 14, 513
- One-tail test, 122, 160
- OPEC, 314
- Operating cash flows, 471–472
- Operating costs, 20
- Operating leverage, degree of, 350–353, 357
- Operational control, 11n
- Opinion polls, 135–136, 161
- Opportunistic behavior, 19, 21, 36
- Opportunity costs, 3, 6, 16, 33–34, 251–252, 283
- Optimal decision, 22–23, 36
- Optimal levels
 of inputs, 190–191, 236–241
 of output, 320–324
 of production, 239
- Options, currency, 519–520
- Orden, David, 112
- Organization of Petroleum Exporting Countries (OPEC), 51, 391, 392
- Original design manufacturer (ODM), 308
- Original equipment manufacturer (OEM), 308
- Orsag, Silvije, 494n
- Oster, Sharon, 373n, 378n
- Otis Elevator Company, 552
- Outflows. *see* Cash flows
- Out-of-pocket costs, 251–252
- Outputs
 alternative terms for, 183
 calculus formulas for, 343–344
 learning curve and, 269–272
 in monopoly markets, 329
 optimal levels of, 320–324
 in perfect competition market, 316–320
 in production function, 181, 183–185, 191
 tangibility of, 223
- Outsourcing, 21, 210
- Overseas Private Investment Corporation, 493
- P**
- P = MC rule, 322–323, 338–339
- Palmisano, Sam, 307
- Parker, John, 109n
- Parker Hannifin Corporation, 413
- Partial productivity measures (PPMs), 220, 221
- PARTS framework, 442
- Patents, 551–552
- Patki, Mukul, 175n
- Payback, 499
- Payback method, 464, 470
- Payoff matrix, 434, 455
- Payoffs, 434, 455
- Peck, Merton J., 87n
- Peers, Alexandra, 395n
- Peltzman, S., 375n
- Penang, Malaysia, 307
- Penetration pricing, 422, 427
- PepsiCo, 273, 394, 554n
- Perdue, Frank, 363
- Perfect competition
 calculus formulas for, 343–344
 characteristics of, 312f, 313f
 definition of, 312, 339
 examples of, 314
 managerial decision making and, 53, 333
 market entry and exit and, 313, 316
 optimal output level and, 320–323
 outputs in, 316–320
 pricing in, 316–320, 415
- Perfect complements, 233
- Perfect elasticity, 74–75
- Perfect inelasticity, 75
- Perfect substitutes, 195, 232
- Performance, 374–375
- Perloff, Jeffrey M., 374n
- Personnel, management of, 8–9
- Per-unit cost curves, 292–294
- Pfizer, 250

- Pharmaceutical industry
 in India, 514
 monopolies in, 334
 Pharmacia, 250
 Philip Morris, 379
 Philippines, 210
 Pierer, Heinrich von, 34n
 Pindyck, R. S., 490n
 Planning
 capacity, 205–207
 enterprise resource, 278
 production, 267–269
 Plant closings, as cost-cutting
 measure, 278
 Planter's Peanuts, 277
 Plastics industry, case study, 298
 Players, 434, 442–443, 456
 Point elasticity
 calculus formulas for, 74, 86, 88n
 computation of, 72, 83, 86n, 88n
 definition of, 94
 estimation of demand and, 74
 marginal revenue and, 82
 regression analysis and, 122–123
 Poison pills, 451, 456
 Polaroid, 329, 334
 Polinsky, A. M., 541n
 Political risk, 493
 Pommerehne, W., 102n
 Pompelli, G., 85n
 Population size and elasticity, 90
 Porter, Michael, 3n, 279,
 376–378
 Portugal, capital budgeting in,
 494–495
 Postponement option, 491
 Poultry industry, 363
 Power functions, 197–198, 212
 Predatory pricing, 422, 427
 Preferences, as nonprice
 determinant of demand, 41, 47t
 Present Situation Index, 136
 Prestige pricing, 422, 427
 Price
 effect of elasticity on, 77–78
 equilibrium, 45, 46–51, 56
 future expectations of, 42
 guiding function of, 48–50, 51f,
 52t, 56
 market demand and, 41–42
 market power and, 53
 as nonprice determinant of
 demand, 47t
 as nonprice determinant of
 supply, 44, 47t
 rationing function of, 46–48, 51f,
 52t, 57
 rigidity. *see* Kinked demand
 curves
 supply and, 42
 Price ceilings, 100
 Price controls, 100
 Price discrimination, 399–410
 in business practice, 404–406
 definition of, 427
 price elasticity of demand and,
 406–407
 protection against, 535–537
 real-life cases of, 406–409,
 424–425
 social welfare implications of, 410
 Price elasticity of demand
 in Baumol's theory, 224
 definition of, 69, 94
 global applications of, 91–92
 measurement of, 69–75
 newspapers and, 109
 price discrimination and,
 406–407
 real-life cases of, 84–85
 revenue and, 80–84
 Verizon and, 112
 Price elasticity of supply, 90–91, 94
 Price fixing
 by cartels, 393–395
 real-life cases of, 423–424
 Price floors, 101–102
 Price leaders, 369, 381
 Price leadership, 396–398, 427
 Price makers, 312, 339, 362
 Price skimming, 422, 427
 Price takers, 312, 339, 362
 Price theory, 375–376
 Price-performance ratios, 264
 Pricing. *see also* Nonprice
 competition
 in airline industry, 425
 asymmetric information and,
 443–446
 calculus formulas for, 343–344
 decisions, 440
 in monopoly markets, 329
 in oligopolies, 367–369
 in perfect competition market,
 316–320
 Pricing for profit, 332, 339
 Pricing for revenue, 332, 339
 Principal-agent problem, 26–27,
 450–451, 456
 Prisoners' dilemma, 437–439, 440,
 454, 456
 Probability, 478, 499
 Probability distribution,
 478–487, 499
 Processing costs, reductions in, 278
 Procter & Gamble, 424
 Product differentiation, 363,
 367–369, 375, 377
 Production
 in China, 210
 costs and, 253–254
 decisions, 8–9
 factors of, 5, 50
 of goods, 220, 224, 225
 issues and challenges to
 competitive, 207–209
 optimal level of, 239
 planning and long-run average
 cost, 267–269
 stages of, 188–189, 196,
 205–207, 213
 Production functions
 aggregate, 203–205
 calculus formulas for, 242–247
 call center, case study for,
 209–210
 Cobb-Douglas. *see* Cobb-Douglas
 production function
 cubic, 196, 197f, 245–247
 definition of, 181, 212
 derived demand and, 190–191
 estimation of, 196–205
 managerial decision-making
 and, 205–207
 power, 197–198, 212
 quadratic, 196, 197f, 247
 short-run versus long-run,
 182–183
 statistical estimation of, 200–205
 studies utilizing, 202–205
 Productivity
 employee, determining,
 446–448
 improving, 221
 labor and, 227–230
 services and, 220–230
 Products
 as capital budgeting
 decisions, 463
 decisions about, 8
 joint, 417–420
 relationships of, 415–420
 Profit. *see also* Economic profit
 defining, 33–34
 definition of, 19
 incentives and, 451–452
 market concentration and,
 375–376
 required, 349–350, 357
 Profit centers, 420
 Profit margin, 411n
 Profit maximization hypothesis,
 23–29, 34–35, 36, 399
 Profit sharing, 452, 456
 Profitability index (PI), 464n, 467,
 499
 Projections. *see* Forecasting
 techniques
 Projects, evaluation of, 467–468, 499
 Promises, 456
 Promotion, 370
 Property. *see* Intellectual property
 Pruitt, Stephen W., 471n
 Psychographics, 374n
 Psychological pricing, 422–423, 427
 Punjabi, Navin, 513n, 526n

- Q**
- Quadratic production function, 196, 197f, 247, 259
- Quaker Oats, 554n
- Qualitative forecasting, 134, 161
- Quantitative forecasting, 134, 161
- Quantity
- change in demand by, 40
 - effect of elasticity on, 77–78
 - equilibrium, 45, 46–51
 - and price, in monopoly markets, 331
 - price elasticity of demand and, 69–70
 - in production function, 182–185, 196, 242
 - in short-run cost function, 255
- Quantity demanded, 39–40, 57
- Quantity supplied, 42, 57
- R**
- R software, 120
- R^2 , 123, 159
- Radebaugh, L. H., 557n
- Raffiee, Kambiz, 85n, 89n
- Ram, Michael A., 89n
- Ram, Rati, 85n
- Ramaswamy, K., 557n
- Rand Corporation, 135
- Randomness, in time series analysis, 147
- Randomness, in time series forecasting, 143
- Rate changes, 296
- Rate of return. *see also* Accounting rate of return; Internal rate of return (IRR)
- capital asset pricing model and, 474–475
 - minimum required, 465
- Rationing, capital, 477, 498
- Rationing function of price, 46–48, 51f, 52t, 57
- Rawsthorn, Alice, 109n
- Real options, 490–493, 499
- Reductions, cost, 277–278
- Refreshment beverage industry. *see* Soft drink industry
- Regression analysis
- analysis of consumer demand using, 129–132
 - calculation of, 124–125
 - capital asset pricing model and, 474
 - coefficients, 120–122, 125
 - definition of, 160
 - demand analysis and, 117–120
 - forecasting and, 155
 - management decisions and, 125
 - problems with, 125–129
 - production function and, 201
 - short-run cost function and, 295
 - variables, types of, in, 119
 - white Zinfandel in Los Angeles (case study), 167–174
- Regulated monopoly, 547
- Regulations, 493, 534. *see also* Antitrust laws
- Reitsch, Arthur G., 133n, 135n, 151n
- Relative elasticity of demand, 74
- Relative inelasticity of demand, 74
- Relevant costs, 251, 283
- Relocation, as cost-cutting measure, 278
- Replacement costs, 34, 251
- Replacement decisions, 463
- Republican Party, 545
- Required profit, 349–350, 357
- Resale values, 472
- Reshoring, 21–22
- Residual term, 128
- Resources
- allocation of, 6–10, 53, 534
 - definition of, 16
 - in microeconomics, 5
 - in production function, 181
 - scarcity of, 5–9
 - traditional categories of, 4, 9
- Retail industry, in India, 514–515
- Return on assets (ROA), 498
- Return on investment (ROI), 498
- Returns to scale, 193–195, 204–205
- calculus formulas for, 240–241, 243, 246
 - definition of, 212–213
 - economies of scale, 263
- Revenue
- mathematics of, 83–84
 - maximization of, 398–399
 - price elasticity of demand and, 80–84
 - profit maximization versus, 23
 - as reason for mergers, 548
- Revolutionary War, price controls during, 101
- RFID technology, 117
- RIM, 553
- Risk. *see also* Business risk
- budgeting under conditions of, 483–484
 - containment of international, 494
 - definition of, 499
 - exchange rate, 517, 531
 - importance of, 5
 - incorporation methods for, 483–485
 - measures of, 478
 - in multinational corporations, 517–518, 531
 - net present value (NPV) and, 483–484
 - political, 493
 - project, 486–487
 - sources of, 478
 - types of, 30
 - uncertainty and, 4–5, 477
 - of wars, 493
- Risk-adjusted discount rate (RADR), 484–485, 499
- Rivalry, in oligopolistic markets, 367–369
- Robbins, Lionel, 7n
- Robinson-Patman Act (1936), 406, 536
- Roche Holding AG, 395, 423
- Rogers, R. P., 304
- Rometty, Virginia, 307n
- Romney, Mitt, 545
- Ross, Stephen A., 469n
- Royal Grolsch, 379
- Royalties, 523
- Ruddick, Graham, 526n
- Rule of 2, 123, 160
- S**
- SABMiller, 379
- Safety equipment, 463
- Sala-i-Martin, Xavier, 225n
- Sales forecasting, 151–152
- Salvage, 472
- Samsung Electronics Co., Ltd., 207, 395
- Samsung of South Korea, 424
- Sande, Eva van de, 130n
- Sanders, Jim, 406n
- Sanders, Nada R., 135n
- SAP, 278
- Saranow, Jennifer, 110n
- Sarbanes-Oxley Act (2002), 28
- Sarnat, Marshall, 483n
- SAS software, 120
- Satisficing, 26–27, 36
- SBC, 250
- Scale economies, 263–265, 273, 282
- Scarcity, 2, 5–9, 16
- Scenario analysis, 486, 487, 499
- Schelling, Thomas C., 434n, 440n, 441
- Schmalensee, Richard L., 537
- Schnabel, Constance, 303n
- Schneider, F., 102n
- Schroeder, Ted C., 90n
- Schuettinger, Robert L., 100n, 101n
- Scope, economies of, 272–273, 282
- Scott, Jonathan A., 470n
- Scottish & Newcastle, 379
- Screening, 448–450, 456
- Seale, James L., Jr., 85n, 87n
- Seasonal indices, 161
- Seasonality fluctuation, in time series analysis, 143, 145–146

- Second-degree price
discrimination, 401
- Securitization, 543–544
- Sellers, number of, 44, 47t
- Selling Areas Marketing, Inc., 108
- Semilogarithmic graphs, 141
- Sen, Anindya, 85n
- Sensitivity analysis, 486, 499
- Separating equilibrium,
446–447, 456
- Serial correlation, 128–129
- Service marks, 551–552
- Services
productivity and, 220–230
shared, 249–250
- Services industry, in India, 514
- Shaffer, Sherrill, 301n
- Shanley, Mark, 373n
- Shared services, 249–250
- Shareholders
impact of mergers on, 549–550
profit maximization and, 26–29
wealth maximization for,
29–35, 469
- Sharpe, William F., 474n
- Sherman Anti-Trust Act (1890),
391, 397, 535
- Shishkin, Philip, 553n
- Shoe store industry, case study, 300
- Short run
break-even analysis, 345–357
definition of, 50, 57
economies of scale in, 273
elasticity in, 79–80
market changes and price in, 52t
- Shortage, 45, 57
- Short-run average cost (SRAC)
learning curve effect on, 271
long-run average cost curve and,
265–267
- Short-run cost function, 255–260
calculus formulas for, 291–294
cost efficiency in, 257–258
estimation of, 282, 295–299
graphs, 292
long-run cost function versus, 263
regression analysis and, 295
studies illustrating, 297–298
- Short-run production function,
182–193
definition of, 182, 213
derived demand and, 190–191
law of diminishing returns and,
185–188
multiple inputs in, 191–193
stages of, 188–189
terminology, 183–185
- Short-run supply curve, 327, 339
- Shugart, William E., II, 374n, 375n
- Shutdown point, 326f, 327, 339
- Siegfried, John J., 154n
- Signals, 446–448, 456
- Silliere, Ed, 406
- Simester, Duncan, 175n
- Simon, Herbert, 411n
- Simon, Julian J., 368n
- Simulation analysis, 487–488, 499
- Singapore, traffic control in,
109–110
- Sinha, Dipendra, 92n
- Slotting fees, 406
- Smith, Adam, 21, 50n, 391
- Smith, Ethan, 112n
- Smith, Michael D., 87n
- SmithKline Beecham, 250
- Snapple, 554n
- Soares, Joao Oliveria, 494n
- Sobotka, Stephen P., 303n
- Social welfare, 375, 410
- Socially optimal price, 539, 555
- Soft drink industry, 9–10, 556–561
- Sonoco, 277
- Sony Publishing Company, 552
- Sotheby's, 394
- South Africa, capital budgeting
in, 495
- South African Breweries (SAB), 379
- Southwest Airlines, 276–277, 396
- Spagat, Elliot, 109n
- Spain, food demand in, 130–131
- Specialization, 186, 189f, 221, 222f
- Specifications, cost function, 296
- Spence, Michael, 443n,
446–447, 454
- Spending plans, surveys of,
136, 161
- SPSS software, 120
- Stabile, Mark, 85n
- Stages of production, 188–189, 196,
205–207, 213
- Standard deviation, 480–481, 499
- Standard error of coefficient (SEC),
122, 127, 160
- Standard error of estimate (SEE), 160
- Standard Industrial Classification
(SIC), 366
- Standard linear costs, 260
- Standard normal distribution, 481
- Standardization, 223
- Statistical probabilities, 478
- Steel industry, import quotas
and, 106
- Steel manufacturing industry,
European cartel in, 424
- Stern Stewart, 32
- Stevenson, R. W., 109n
- Steward, Devin T., 335
- Stewart, G. Bennett, III, 32n
- Stigler, George J., 21, 199, 304, 368n,
375, 392n, 540n
- Stiglitz, Joseph, 443n, 448
- Stimpert, J. L., 302n
- Stock options, 452, 456
- Stockholders. *see* Shareholders
- Stockwell, Tim, 85n
- Stonehill, Arthur I., 493n,
494n, 510n
- Strategic planning, 11n
- Strategy
definition of, 381
Five Forces model of,
376–378
in game theory, 434, 456
in imperfect competition,
372–378
price theory model, 375–376
structure-conduct-performance
(S-C-P) paradigm, 375–376
The Strategy of Conflict (Schelling),
434n
- Stroh's Brewery, 379
- Structure-conduct-performance
(S-C-P) paradigm, 374–375, 381
- Stutzman, J. R., 85n
- Subprime loan financial crisis,
2007–2009, 542–546
- Substitute goods, 86, 94, 371, 417
- Substitution
elasticity and, 75–77
of input factors, 232–236
as nonprice determinant of
supply, 44
- Sugar supply, import quotas
and, 106
- Suits, Daniel B., 87n, 154
- Sullivan, D. P., 557n
- Sun Microsystems, 537
- Sundaram, Anant, 550n
- Sunk costs, 252, 283
- Supermarkets and consumer
demand, 108
- Supply
definition of, 42, 57
guiding function of price and,
48–50, 51f
law of, 42
market, 42–44
market equilibrium and,
44–45
market power and, 53
rationing function of price and,
46–48, 51f, 52t
real-life cases in, 105–113
scarcity and, 5–9
- Supply chain management (SCM),
273–276, 283, 306–309
- Supply curve, 327, 339
- Supply elasticity, 90–91, 94, 105
- Surplus, 45, 57
- Survey of Consumers* (Survey
Research Center, University of
Michigan), 136
- Survey of Manufactures* (Census
Bureau), 365–369
- Survey Research Center, University
of Michigan, 136

- Surveys
 consumer, 116–117
 forecasts and forecasting
 and, 137
 of spending plans, 136, 161
- Survivorship techniques, 283,
 304–305
- Sweezy, Paul, 367
- Switzerland, 302
- Symantec Corporation, 526
- T**
- Tacit bargaining, 439, 440, 441, 456
- Taiwan, demand for cigarettes in,
 131–132
- Takeovers, 451
- Tapscott, Don, 22n
- Tariffs, two-part, 401n
- Tarr, D. G., 304
- Tastes
 as nonprice determinant of
 demand, 47t
 and preferences, 41
- Taubman, A. Alfred, 394
- Taut, Nikki, 424n
- Tax credits, 522
- Tax rates
 in multinational capital
 budgeting, 521–522
 multinational transfer pricing
 and, 523–526
- Taxes
 excise, 102–105
 socially optimal pricing and, 539
- Taylor, Colleen, 395n
- Taylor, L. D., 85n, 89n
- Teboul, James, 222n
- Technology
 and data collection, 117
 as nonprice determinant of
 supply, 43–44, 47t
- Tedeschi, Bob, 22n
- Teinowitz, Ira, 406n
- Telang, Rahul, 87n
- Telecommunications industry,
 mergers in, 547
- Tennant, Anthony J., 394–395
- Thaler, Robert, 176n
- The Strategy of Conflict
 (Schelling), 440n
- Theijssen, Daphne, 130n
- Thermos GmbH, 552
- Third-degree price discrimination,
 401–404, 409, 410
- Thompson, Bill, 113n
- Threats, 440, 456
- ThyssenKrupp AG, 424
- Ticoll, David, 22n
- Tie-in sales, 409–410
- Tiems, Jorieke, 130n
- Time lags and government policy, 542
- Time periods. *see* Long run;
 Short run
- Time series analysis, 142–147, 201,
 204, 295
- Time series data, 118, 128–129, 160
- Time series forecasting, 161
- Time value of money, 29, 464, 499
- Times (London), 109
- T-Mobile, 547
- Tobacco production, 112
- Tomkins, Richard, 108n
- Total analysis approach, 320–321
- Total cost function, 258–260, 283
- Total cost (TC), 20, 253, 255,
 256f, 283
- Total factor productivity (TFP). *see*
 Multifactor productivity (MFP)
- Total fixed cost (TFC), 253, 255,
 256f, 283, 357
- Total labor cost (TLC), 190, 191
- Total product (TP). *see also*
 Quantity
 calculus formulas for, 242–243
 curves, 196n
 definition of, 183, 213
 short-run analysis, 183–185
- Total revenue product (TRP), 191
- Total variable cost (TVC), 253, 255,
 256f, 283, 357
- Toyota Corporation, 208–209, 407
- Trade secrets, 551–552
- Trademarks, 551–552
- Traditional process, 7, 8, 16
- Traffic control and demand
 elasticity, 109–110
- Transaction costs, 19–22, 37, 298
- Transactions, offsetting, 519
- Transfer pricing, 420–422, 427,
 521n, 523–526
- Transportation Center,
 Northwestern University, 303
- Transportation costs, 265
- Transportation industry, case
 study, 304
- Traveler's Insurance, 547
- Travelocity, 425
- Trend lines, in time series
 forecasting, 143, 146, 152
- Trend projections
 definition of, 161
 types of, 139–151
- Triplett, Jack E., 221n, 224n, 225f
- t-statistic, 160
- Tsuei, A., 33n
- Tsurumi, Yoshi, 106n
- t-table, 122, 160
- t-test, 122–123, 125, 127, 128, 160
- Tucker, Emma, 423n
- Tully, Brian, 525n
- Tully, S., 33n
- TWA, 547
- Two-part tariffs, 401n
- Two-tail test, 122–123, 160
- Tying arrangements, 409–410
- Tyson Foods, 363
- U**
- U.S. Airways, 547
- U.S. Census of Manufactures, 366n
- U.S. Department of Agriculture
 (USDA), 113
- U.S. Department of Justice,
 333–334, 394, 395, 397,
 536–537, 538, 553
- U.S. Department of Justice Merger
 Guidelines, 366–367
- U.S. Federal Reserve, 203, 541,
 543–544, 546
- U.S. National Recovery Act (NRA),
 392n
- U.S. Patent and Trademark Office
 (USPTO), 551
- U.S. Treasury, 544
- Uncertainty
 definition of, 19, 499
 in game theory, 442
 risk and, 4–5, 477
 transaction costs and, 19
- Unemployment, 102
- Unilever, 424
- Unit costs, 259–260
- Unitary elasticity of demand, 74
- United Airlines, 250, 547
- United Kingdom, banking industry
 in, case study, 300–301
- United Nations Conference on
 Trade and Development
 (UNCTAD), 515
- United States
 2007–2009 subprime loan
 financial crisis in, 542–545
 automobile imports and Japan,
 105–106
 banking industry, case study, 301,
 304–305
 demand for beef, 106–107
 demand for corn in, 113
 furniture industry in, 111
 import quotas and, 106
 price controls in, 100–101
 regulation of commerce in,
 535–537
 resource allocation in, 7, 534
- United Technologies, 553
- Universal Music Group, 112
- V**
- Valuation
 project, in multinational
 corporation, 521–522
 as reason for mergers, 548–550

- Variable sum games, 439, 456. *see*
also Prisoners' dilemma
- Variables
 capital, 196n
 endogenous, 155–156
 exogenous, 156
 labor, 196n
 types of, in regression analysis, 119
- Variation, coefficient of. *see*
 Coefficient of variation
- Velocity of money, 544
- Verhagen, Véronique, 130n
- Veritas Software Corporation, 526
- Verizon, 112, 250, 547
- Vietnam, price discrimination in, 425
- Visa, 372
- Visual time series projections,
 140–141
- Vitamin industry, European cartel
 in, 423
- Vogel, Ronald J., 304n
- Volkswagen, 209
- Volume-cost-profit analysis. *see*
 Break-even analysis
- Von Neumann, John, 483n
- Vortsman, Lisa, 511–513
- W**
- Wachman, Richard, 526n
- Wages, price floors on, 101–102
- Walden, Gene, 276n, 363n
- Waldman, Don E., 535n
- Walker, David A., 298n
- Wal-Mart
 cost reduction and, 276
 data collection and, 117
 price leadership by, 397
- Wanamaker, John, 370n
- Wang, Claire, 130n
- Warranties, as market signals,
 446, 448
- Wars, risk of, 493
- Washkewicz, Donald, 413
- Wealth, redistribution of, 534
- Wealth maximization, 29–33, 37, 469
- Wearden, Graeme, 424n
- Weather conditions, as nonprice
 determinant of supply, 44, 47t
- Webber, M. D., 306n
- Weighted average cost of capital
 (WACC), 475, 499
- Weir, Tom, 406n
- Welch, F., 102n
- Werden, Gregory J., 85n
- Werdigier, Julia, 379n
- Westerfield, Randolph W., 469n
- Westinghouse, 391, 393
- What-if analysis, 46
- Wheeler-Lea Act (1938), 536
- Wheelwright, Steven C., 132n,
 133n, 135n
- Wichern, Dean W., 133n, 135n, 151n
- Wiesenfelder, Margaret, 303n
- Wilke, John R., 395n
- Williamson, Oliver E., 19n
- Wilson, J. Holton, 151n
- Wine industry
 break-even analysis example,
 355–366
 by the glass restaurant
 example, 61–64
 demand analysis of white
 zinfandel in Los Angeles
 (case study), 167–174
 French, case study, 111
- Wines, Michael, 101n
- Wonacott, Peter, 112n
- Woolley, S., 405n
- Woolway, Mark D., 205n
- Working capital, 472
- World War II, price controls during,
 100n, 101–102
- Worrall, John D., 304n
- X**
- Xerox, 552
- Y**
- Yahoo, Inc., 526
- Yamagata, Tadashi, 336
- Yazaki Corporation, 395
- Yeager, Timothy J., 305n
- Z**
- Zande, Patrick van der, 130n
- Zara, 207–208
- Zardoshty, Farangis, 154n
- Zavodny, M., 102n
- Zero sum games, 439, 456
- Zhan, Ping, 85n
- Zhao, Jinhui, 85n
- Zola, Marika, 302n
- Zone pricing, 406
- z-statistic, 481

This page intentionally left blank



Excel Apps Available on the Companion Website

Ch. #	App Title	Worksheet (if multiple worksheets)	What is Discussed on this Worksheet
0	Introduction To Excel	Main	How to structure a problem in Excel and overview of other worksheets
		Referencing	How to fix a row or column so an equation can be dragged
		CreatingPatterns	Create patterns by beginning them and dragging, or by using equations
		AboutEquations	Rules for creating equations in Excel
		AboutShortcuts	How to rapidly maneuvering about large data sets and basic commonly used Excel shortcuts
		CreditCards	The true cost of paying off a credit card using minimum required payments
3	Demand	SingleDemand	How to create and modify a graph with a single linear demand curve
		MultipleDemands	How to create a graph with multiple linear demand curves
4	Elasticity		The relation between elasticity and total revenue given linear demand
	Constant Elasticity Demand (CED)	Demand	Graphing a single constant elasticity demand function
		InverseDemand	Comparing CED curves through a point with different elasticities
5	Geometric Introduction to Univariate Regression	Main	Visualizing the regression line, running a regression in Excel, and understanding Excel's regression idiosyncrasies
		(P,Q)	How to fit a data set with a line. What is the equation for that line?
		UnivariateTake1	Simple (univariate) linear regression, what you should initially examine on a regression output
		Excel'sRegression Areas	The three main parts of Excel's regression output and the meaning of R^2
		VisualizingUnivariate Regression	Predicted values versus residuals
		HowRegressionWorks (Optional)	Interactive sheet showing that regression works by choosing coefficients that minimizes the SS residuals.
	Geometric Introduction to Multivariate Regression	Bivariate Model	Bivariate regression, how to get independent variables contiguous
		Visualizing Multivariate Regression	Graphic depiction of bivariate regression, geometric meaning of estimate
	Multivariate Regression	t,SEC,CI& Estimation	How to create an estimate, t statistics, SEE, SEC, confidence intervals, and the difference between magnitude and significance
		Adjusted R^2	A bit more about the predictive power of a regression. Adjusted R^2 versus R^2
		Sheet 5.1 Table 5.3	Excel's raw regression output (from the text, p. 520) Sample data: The Demand for pizza (from the text, p. 119)
	Dummy Variables	VisualizingDummy Variables	Graphic depiction of how to interpret dummy variables
		MultipleDummies	Significance of dummy variables and general rules for creating dummies
		OmittedVariables	Introduction to residual analysis
	Integrated Introduction to Regression	(Multiple Sheets)	Integrated workbook providing the four above apps in one file with Index and printable Outline
Nonlinearities	Data	Estimating a demand relation when the relation is nonlinear (review of logarithm rules)	
	Linear	Estimates and residuals from a linear model given nonlinear data	
	Quadratic	Estimates and residuals from a quadratic model given nonlinear data	
	CED	Estimating CED function using Log Log form (both Log and Ln versions)	
Multicollinearity	Main	Data sheet and examination of correlation among dependent variables	
	Regressions	24 regression models (3×8) used to show the effect of multicollinearity	
	CoefficientMaps	Graphical depiction of coefficients and 95% CIs in multicollinear models	

Ch. #	App Title	Worksheet (if multiple worksheets)	What is Discussed on this Worksheet
	Serial Correlation and Durbin Watson (DW)	Main	Understanding serial correlation (also called autocorrelation) and interpreting the DW statistic
		SampleRegression	How to calculate DW in Excel using the Residuals option
		Mini-Lectures	Visualizing SC, Interpreting DW using d_t and d_U
	Growth Models	Data	Unit sales of merlot in LA (with optional discussion of deseasonalization)
		LnVersusLog	Understanding why the LnQ model is more heuristic than the LogQ model
	Seasonality and Seasonal Indices	ReadMe	Instructions on how to create seasonal indices using regression
		Data	Data from Table 5.5 p. 144 columns 1 & 2. How to create seasonal dummies
		PrelimRegLn	Find the peak season using a preliminary regression
		RegLn	Obtaining seasonal indices from seasonal dummies (with Ln and Log)
		RegAdditive	Additive indices
Demand with Seasonality	CIwithLogarithms	Estimation in pieces and avoiding mistakes in estimating CI with log models	
		(Multiple Sheets)	The Appendix 5A white zinfandel in LA data and analysis
6	Cobb Douglas Production		Interactive graph shows how varying a , b , and c changes $Q = aL_bK_c$
Cobb Douglas Production (Regression)	Data	Data used for estimating CRTS and general CD production models	
	Regressions	Regressions for both the general CD estimate reported in Table 6.6 and CRTS CD estimate discussed in footnote 15, p. 201	
	Table 6.6	From p. 202 expanded to show CRTS and general CD production functions	
Figure 6.10		Interactive version of Figure 6.10 p. 206 examining the choice among plant sizes	
Cubic Production 3D	Sheet 6C.1	How to set up the calculations needed to find the labor level L where CRTS occurs given capital intensity k in Excel (from the text, p. 247)	
	Figure 6C.1	Returns to scale given cubic production (from the text, p. 246), plus how it was made	
	3DGraphs	3D views showing SR production and isoquant maps and overhead view	
7	Figure 7.3		Dynamic version of Figure 7.3 p. 258 allowing you to vary FC and VC
Figure 7.10		Interactive Figure 7.10 p. 268 examines discrete versus continuous LAC	
Figure 7.12		Interactive Figure 7.12 p. 270 examines learning curve effects	
Cobb Douglas Cost		Dynamic figures show how costs respond to changes in the production function and input cost	
Estimating Variable Cost (Regression)	Main	Data sheet and background notes regarding estimating VC	
	Quadratic&BadCubic	When is a quadratic functional form appropriate and when does a cubic form fail	
	Cubic	Estimating cubic cost function using Excel's Constant is Zero option	
Per-Unit Costs (Appendix 7A)	Main	Color coded version of Sheet 7A.1 p. 293	
	Equations	Color coded version of Sheet 7A.2 p. 293	
	Sheet 7A.1	How to go from a cubic cost equation to graph all per-unit cost curves	
	Sheet 7A.2	Equations for going from a cubic cost equation to graph all per-unit cost curves	
8	Breakeven Analysis	Sheet 8B.1	Estimating breakeven sales using three increasingly sophisticated breakeven calculations (from the text, p. 356)
9	Monopolistic Competition	Sheet 9A.1	Monopolistic competition in SR and LR (from the text, p. 389)
	Kinked Demand		Useful for doing the kinked-demand end of chapter problems
12	Present Value	PresentValue	How to do present value problems in Excel
		IRR	Using Goal Seek to obtain IRR on any project