At the new input levels, AFW can produce 79.952 million gallons of the new soft drink, which represents an increase of 10.450 million-gallons with no increase in the cost of production.

It should be clear from these results that AFW was not operating efficiently at the original input levels. While AFW is operating more efficiently with the new input mix, it remains an open question whether the company is maximizing output with an operating budget of $2 million and prevailing input prices. In other words, we still do not know whether the new input mix is optimal.

d. If AFW sells its output at a fixed price, new input levels clearly will cause the company's total profit to rise. The total cost of producing the new soft drink last year and this year was $2,000,000. If AFW can sell the new soft drink to regional bottlers for $0.05 per gallon, last year's total revenues amounted to $3,475,100 ($0.05 × 69,502,000), for a total profit of $1,475,100 ($3,475,100 – $2,000,000). By reallocating the budget and changing the input mix, AFW total revenues increased to $3,997,600 ($0.05 × 79,952,000) for a total profit of $1,997,600, or an increase in profit of $522,500.

THEORY OF THE FIRM

The concept of the "firm" or the "company" is commonly misunderstood. Too often, the corporate entities are confused with the people who own or operate the organizations. In fact, a firm is an activity that combines scarce productive resources to produce goods and services that are demanded by society. Firms are more appropriately viewed as an activity that transforms productive inputs into outputs of goods and services. The manner in which productive resources are combined and organized will depend of the organizational objective of the owner–operator or, as in the case of publicly owned companies, the decisions of the designated agents of the company's shareholders.

Scarce productive resources are many and varied. Consider, for example, the productive resources that go into the production of something as simple as a chair. First, there are various types of labor employed, such as designers, machine tool operators, carpenters, and sales personnel. If the chair is made of wood, decisions must be made regarding the type or types of wood that will be used. Will the chair have upholstery of some kind? If so, then decisions must be made on material, quality, and patterns. Will the chair have any attachments, such as small wheels on the bottom of the legs for easy moving? Will the wheels be made of metal, plastic, or some composite material?

The point is that even something as relatively simple as a chair may require quite a large number of resources in the production process. It should be clear, therefore, that when one is discussing economic and business rela-

tionships in the abstract, making too many allowances for reality has its limitations. To overcome this problem, we will assume that production is functionally related to two broad categories of inputs, labor and capital.

THE OBJECTIVE OF THE FIRM

Economists have traditionally assumed that the goal of the firm is to maximize profit π. This behavioral assumption is central to the neoclassical theory of the firm, which posits the firm as a profit-maximizing "black box" that transforms inputs into outputs for sale in the market. While the precise contents of the "black box" are unknown, it is generally assumed to contain the "secret formula" that gives the firm its competitive advantage. In general, neoclassical theory makes no attempt to explain what actually goes on inside the "black box," although the underlying production function is assumed to exhibit certain desirable mathematical properties, such as a favorable position with respect to the law of diminishing returns, returns to scale, and substitutability between and among productive inputs. The appeal of the neoclassical model is its application to a wide range of profit-maximizing firms and market situations.

Neoclassical theory attempts to explain the behavior of profit-maximizing firms subject to known resource constraints and perfect market information. It is important, however, to distinguish between current period profits and the stream of profits over some period of time. Often, managers are observed making decisions that reduce this year's profit in an effort to boost net income in future. Since both present and future profits are important, one approach is to maximize the present, or discounted, value of the firm's stream of future profits, that is,

$$\begin{aligned} \text{Maximize: } PV(\pi) &= \frac{\pi_1}{(1+i)} + \frac{\pi_2}{(1+i)^2} + \ldots + \frac{\pi_n}{(1+i)^n} \\ &= \sum \frac{\pi_t}{(1+i)^t} \end{aligned} \tag{1.4}$$

where profit is defined in Equation (1.3), t is an index of time, and i the appropriate discount rate.[4] The behavior characterized in Equation (1.4) assumes that the objective of the firm is that of wealth maximization over some arbitrarily determined future time period. Equation (1.4) gives the

[4] The concept of the time value of money is discussed in considerable detail in Chapter 12. The time value of money recognizes that $1 received today does not have the same value as $1 received tomorrow. To see this, suppose that $1 received today were deposited into a savings account paying a certain 5% annual interest rate. The value of that deposit would be worth $1.05 a year later. Thus, receiving $1 today is worth $1.05 a year from now. Stated differently, the future value of $1 received today is $1.05 a year from now. Alternatively, the present value of $1.05 received a year from now is $1 received today. The process of reducing future values to their present values is often referred to as discounting. For this reason, the interest rate used in present value calculations is often referred to as the discount rate.

immediate value of the firm's profit stream, which is expected to grow to a specified value at some time in the future. Discounting is necessary because profits obtained in some future period are less valuable than profits earned today, since profits received today may be reinvested at an interest rate i.

Note that Equation (1.4) may be rewritten as

$$PV(\pi) = \sum \frac{\pi_t}{(1+i)^t} = \sum \frac{(TR_t - TC_t)}{(1+i)^t} \tag{1.5}$$

Equation (1.5) explicitly recognizes the importance of decisions made in separate divisions of a prototypical business organization. The marketing department, for example, might have primary responsibility for company sales, which are reflected in total revenue (TR). The production department has responsibility for monitoring the firm's costs of production (TC), while corporate finance is responsible for acquiring financing to support the firm's capital investment activities and is therefore keenly interested in the interest rate (i) on acquired investment capital (i.e., the discount rate).

This more complete model of firm behavior also has the advantage of incorporating the important elements of time and uncertainty. Here, the primary goal of the firm is assumed to be expected wealth maximization, and is generally considered to be the primary objective of the firm.

Problem 1.2. The managers of the XYZ Company are in a position to organize production Q in a way that will generate the following two net income streams, where $\pi_{i,j}$ designates the ith production process in the jth production period.

$$\pi_{1,1}(Q) = \$100; \pi_{1,2}(Q) = \$330$$

$$\pi_{2,1}(Q) = \$300; \pi_{2,2}(Q) = \$121$$

For example, $\pi_{1,2}(Q)$ = \$330 indicates that net income from production process 1 in period 2 is \$330. If the anticipated discount rate for both production periods is 10%, which of these two net income streams will generate greater net profit for the company?

Solution. Both profit streams are assumed to be functions of output levels and to represent the results of alternative production schedules. Note that although the first profit stream appears to be preferable to the second, since it yields \$9 more in profit over the two periods, computation of present values (PV) reveals that, in fact, the second π stream is preferable to the first.

$$PV(\pi_1) = \sum \frac{\pi^t}{(1+i)^t} = \frac{\$100}{1.1} + \frac{\$330}{(1.1)^2} = \$363.64$$

$$PV(\pi_2) = \sum \frac{\pi^t}{(1+i)^t} = \frac{\$300}{1.1} + \frac{\$121}{(1.1)^2} = \$372.73$$

HOW REALISTIC IS THE ASSUMPTION OF PROFIT MAXIMIZATION?

The assumption of profit maximization has come under repeated criticism. Many economists have argued that this behavioral assertion is too simplistic to describe the complex nature and managerial thought processes of the modern large corporation. Two distinguishing characteristics of the modern corporation weaken the neoclassical assumption of profit maximization. To begin with, the modern large corporation is generally not owner operated. Responsibility for the day-to-day operations of the firm is delegated to managers who serve as agents for shareholders.

One alternative to neoclassical theory based on the assumption of profit maximization is *transaction cost theory*, which asserts that the goal of the firm is to minimize the sum of external and internal transaction costs subject to a given level of output, which is a first-order condition for profit maximization.[5] According to Ronald Coase (1937), who is regarded as the founder of the transaction cost theory, firms exist because they are excellent resource allocators. Thus, consumers satisfy their demand for goods and services more efficiently by ceding production to firms, rather than producing everything for their own use.

Still another theory of firm behavior, which is attributed to Herbert Simon (1959), asserts that corporate executives exhibit *satisficing behavior*. According to this theory, managers will attempt to maximize some objective, such as executive salaries and perquisites, subject to some minimally acceptable requirement by shareholders, such as an "adequate" rate of return on investment or a minimum rate of return on sales, profit, market share, asset growth, and so on. The assumption of satisficing behavior is predicated on the belief that it is not possible for management to know with certainty when profits are maximized because of the complexity and uncertainties associated with running a large corporation. There are also noneconomic organizational objectives, such as good citizenship, product quality, and employee goodwill.

The closely related theory of manager-utility maximization was put forth by Oliver Williamson (1964). Williamson argued that managers seek to maximize their own utility, which is a function of salaries, perquisites, stock options, and so on. It has been argued, however, that managers who place their own self-interests before the interests of shareholders by failing to exploit profit opportunities may quickly find themselves looking for work. This will come about either because shareholders will rid themselves of

[5] Transactions costs refer to costs not directly associated with the actual transaction that enable the transaction to take place. The costs associated with acquiring information about a good or service (e.g., price, availability, durability, servicing, safety) are transaction costs. Other examples of transaction costs include the cost of negotiating, preparing, executing, and enforcing a contract.

managers who fail to maximize earnings and share prices or because the company finds itself the victim of a corporate takeover. William Baumol (1967), on the other hand, has argued that sales or market share maximization after shareholders' earnings expectations have been satisfied more accurately reflects the organizational objectives of the typical large modern corporation.

Marris and Wood (1971) have argued that the objective of management is to maximize the firm's *valuation ratio*, which is related to the growth rate of the firm. The firm's valuation ratio is defined as the ratio of the stock market value of the firm to its highest possible value. The highest possible value of this ratio is 1. According to this view, since managers are primarily motivated by job security, they will attempt to achieve a corporate growth rate that maximizes profits, dividends, and shareholder value. The importance of the valuation ratio is that it may be used as a proxy for a shareholder satisfaction with the performance of management. The higher the firm's valuation ratio, the less likely that managers will be ousted.

Still another important contribution to an understanding of firm behavior is *principal–agent theory* (see, for example, Alchain and Demsetz, 1972; Demsetz and Lehn, 1985; Diamond and Verrecchia, 1982; Fama and Jensen, 1983a, 1983b; Grossman and Hart, 1983; Harris and Raviv, 1978; Holstrom, 1979, 1982; Jensen and Meckling, 1976; MacDonald, 1984; and Shavell, 1979). According to this theory, the firm may be seen as a nexus of contracts between principals and "stakeholders" (agents). The principal–agent relationship may be that between owner and manager or between manager and worker. The principal–agent problem may be summarized as follows: What are the least-cost incentives that principals can offer to induce agents to act in the best interest of the firm? Principal–agent theory views the principal as a kind of "incentive engineer" who relies on "smart" contracts to minimize the opportunistic behavior of agents. Owner–manager and manager–worker principal–agent problems will be examined in greater detail in the next two sections.

Definition: This principal–agent problem arises when there are inadequate incentives for agents (managers or workers) to put forth their best efforts for principals (owners or managers). This incentive problem arises because principals, who have a vested interest in the operations of the firm, benefit from the hard work of their agents, while agents who do not have a vested interest, prefer leisure.

Although these alternative theories of firm behavior stress some relevant aspects of the operation of a modern corporation, they do not provide a satisfactory alternative to the broader assumption of profit maximization. Competitive forces in product and resource markets make it imperative for managers to keep a close watch on profits. Otherwise, the firm may lose market share, or worse yet, go out of business entirely. Moreover, alternative organizational objectives of managers of the modern corporation

cannot stray very far from the dividend-maximizing self-interests of the company's shareholders. If they do, such managers will be looking for a new venue within which to ply their trade.

Regardless of the specific firm objective, however, managerial economics is less interested in how decision makers actually behave than in understanding the economic environment within which managers operate and in formulating theories from which hypotheses about cause and effect may be inferred. In general, economists are concerned with developing a framework for predicting managerial responses to changes in the firm's operating environment. Even if the assumption of profit maximization is not literally true, it provides insights into more complex behavior. Departures from these assumptions may thus be analyzed and recommendations made. In fact, many practicing economists earn a living by advising business firms and government agencies on how best to achieve "efficiency" by bringing the "real world" closer to the ideal hypothesized in economic theory. Indeed, the assumption of profit maximization is so useful precisely because this objective is rarely achieved in reality.

OWNER–MANAGER/PRINCIPAL–AGENT PROBLEM

A distinguishing characteristic of the large corporation is that it is not owner operated. The responsibility for day-to-day operations is delegated to managers who serve as agents for shareholders. Since the owners cannot closely monitor the manager's performance, how then shall the manager be compelled to put forth his or her "best" effort on behalf of the owners?

If a manager is paid a fixed salary, a fundamental incentive problem emerges. If the firm performs poorly, there will be uncertainty over whether this was due to circumstances outside the manager's control was the result of poor management. Suppose that company profits are directly related to the manager's efforts. Even if the fault lay with a goldbricking manager, this person can always claim that things would have been worse had it not been for his or her herculean efforts on behalf of the shareholders. With absentee ownership, there is no way to verify this claim. It is simply not possible to know for certain why the company performed poorly. When owners are disconnected from the day-to-day operations of the firm, the result is the *owner–manager/principal–agent problem*.

To understand the essence of the owner–manager/principal–agent problem, suppose that a manager's contract calls for a fixed salary of $200,000 annually. While the manager values income, he or she also values leisure. The more time devoted to working means less time available for leisure activities. A fundamental conflict arises because owners want managers to work, while managers prefer leisure. The problem, of course, is that

the manager will receive the same $200,000 income regardless of whether he or she puts in a full day's work or spends the entire day enjoying leisure activity. A fixed salary provides no incentive to work hard, which will adversely affect the firm's profits. Without the appropriate incentive, such as continual monitoring, the manager has an incentive to "goof off."

Definition: The owner–manager/principal–agent problem arises when managers do not share in the success of the day-to-day operations of the firm. When managers do not have a stake in company's performance, some managers will have an incentive to substitute leisure for a diligent work effort.

INCENTIVE CONTRACTS

Will the offer of a higher salary compel the manager to work harder? The answer is no for the same reason that the manager did not work hard in the first place. Since the owners are not present to monitor the manager's performance, there will be no incentive to substitute work for leisure. A fixed-salary contract provides no penalty for goofing off. One solution to the principal–agent problem would be to make the manager a stakeholder by offering the manager an *incentive contract*. An incentive contract links manager compensation to performance. Incentive contracts may include such features as profit sharing, stock options, and performance bonuses, which provide the manager with incentives to perform in the best interest of the owners.

Definition: An incentive contract between owner and manager is one in which the manager is provided with incentives to perform in the best interest of the owner.

Suppose, for example, that in addition to a salary of $200,000 the manager is offered 10% of the firm's profits. The sum of the manager's salary and a percentage of profits is the manager's gross compensation. This profit-sharing contract transforms the manager into a stakeholder. The manager's compensation is directly related to the company's performance. It is in the manager's best interest to work in the best interest of the owners. Exactly how the manager responds to the offer of a share of the firm's profits depends critically on the manager's preferences for income and leisure. But one thing is certain. Unless the marginal utility of an additional dollar of income is zero, it will be in the manager's best interest not to goof off during the entire work day. Making the manager a stakeholder in the company's performance will increase the profits of the owner.

OTHER MANAGEMENT INCENTIVES

The principal–agent problem helps to explain why a manager might not put forth his or her effort on behalf of the owner. There are, however, other reasons why a manager would work in the best interest of the owner that

are quite apart from the direct incentives associated with being a stakeholder in the success of the firm. These indirect incentives relate directly to the self-interest of the manager. One of these incentives is the manager's own reputation.

Managers are well aware that their current position may not be their last. The ability of managers to move to other more responsible and lucrative positions depends crucially on demonstrated managerial skills in previous employments. An effective manager invests considerable time, effort, and energy in the supervision of workers and organization of production. The value of this investment will be captured in the manager's reputation, which may ultimately be sold in the market at a premium. Thus, even if the manager is not made a stakeholder in the firm's success through profit sharing, stock options, or performance bonuses, the manager may nonetheless choose to do a good job as a way of laying the groundwork for future rewarding opportunities.

Another incentive, which was discussed earlier, relates to the manager's job security. Shareholders who believe that the firm is not performing up to its potential, or is not earning profits comparable to those of similar firms in the same industry, may then move to oust the incumbent management. Closely related to a shareholder revolt is the threat of a takeover. Sensing that a firm's poor performance may be the result of underachieving or incompetent managers another company might move to wrest control of the business from present shareholders. Once in control, the new owners will install a more effective management team to increase net earnings and raise shareholder value.

MANAGER–WORKER/PRINCIPAL–AGENT PROBLEM

The principal–agent problem also arises in the relationship between management and labor. Suppose that the manager is a stakeholder in the firm's operations. While manager's well-being is now synonymous with that of the owners, there is potentially a principal–agent problem between manager and worker. Without a stake in the company's performance, there will be an incentive for some workers to substitute leisure for hard work. Since it may not be possible to closely and constantly monitor worker performance, the manager is confronted with the principal–agent problem of providing incentives for diligent work. As before, the solution is to transform the worker into a stakeholder.

Definition: The manager–worker/principal–agent problem arises when workers do not have a vested interest in a firm's success. Without a stake in the company's performance, there will be an incentive for some workers not to put forth their best efforts.

PROFIT AND REVENUE SHARING

As in the case of the owner–manager/principale–agent problem, workers can be encouraged to put forth their best efforts by linking their compensation to the firm's profitability. Another way to enhance worker performance is to tie compensation to the firm's revenues. This method of compensation is particularly important when worker performance directly impact revenues rather than operating costs. The most common form of revenue sharing is the sales commission. When we think of sales commissions, we tend to think of insurance agents, real estate brokers, automobile salespersons, and so on. But sales commissions take a variety of forms. The familiar system in which bartenders and waiters earn tips also constitutes a revenue-based incentive scheme.

There are, however, problems associated with revenue-based incentive schemes. One problem is that such compensation mechanisms may lead to unethical behavior toward customers. This is especially true when customer contact is on a one-time or impersonal basis. The negative stereotypes associated with some professions, such as telephone marketers or used-car salespeople, attest to the potential dangers of linking compensation to revenues. Another problem with linking compensation to revenues is that there is generally no incentive for workers to minimize cost. Corporate executives who inflate expense accounts in attempts to curry favor with potential clients and bartenders who give free drinks attest to some of the problems associated with revenue-based incentive schemes.

OTHER WORKER INCENTIVES

Other methods of encouraging workers to put forth their best efforts are piecework, time clocks, and spot checks. Piecework involves payment based on the number of units produced. Sweatshop operations, once common in the textile industry, are examples of this type of revenue-based incentive scheme. Of course, when worker compensation is based on piecework high quantity often comes at the expense of low product quality. Low-quality products may lead to customer dissatisfaction, which in turn results in lower sales, revenues, and profits.

Time clocks indicate whether workers show up for work on time and stay til the ends of their shifts. However, time clocks do not monitor worker performance while at the workplace. Thus, the use of time clocks is an inferior solution to the manager–worker/principal–agent problem. A more effective solution, which verifies that not only the worker is on the job but that the worker is performing up to expectations, is the spot check. To be effective, spot checks must be unpredictable. Otherwise, workers will know when to work hard and when goofing off will not be noticed.

There are two distinct problems with spot checks. To be effective, random spot checks must be frequent enough to raise the expected penalty to the

worker who is caught goldbricking. Frequent spot checks, however, are costly and reduce the firm's profitability. In addition, frequent spot checks can have a negative effect on worker morale. Low worker morale will negatively affect productivity and profitability. In general, incentive-based schemes based on threats are inferior to compensation-based solutions, such as revenue or profit sharing, to the principal–agent problem.

CONSTRAINTS ON THE OPERATIONS OF THE FIRM

Suppose that the objective of the firm is to maximize short-run profits (or wealth, or value). In attempting to achieve this objective, the firm faces a number of constraints. These constraints might include a scarcity of essential productive resources, such as a certain type of skilled labor, specific raw materials, as might occur because of labor discontent in the country of a foreign supplier, limitations on factory or warehouse space, and unavailability of credit. Constraints might also take the form of legal restrictions on the operations of the firm, such as minimum wage laws, pollution emission standards, and legal restrictions on certain types of business activity. Such constraints are often imposed by government to achieve perceived social (welfare) goals.

For many business and economic applications, it is necessary to think in terms of the optimizing managerial objectives subject to one or more side constraints. This process is referred to as constrained optimization. For example, it might be the goal of a firm to maximize profits subject to limitations on operating budgets or the level of output. The existence of these constraints usually means that the range of possibilities available to the firm is limited. Thus, profit maximization in the strict sense may not be possible. Put differently, the maximum attainable profits in the presence of such constraints are likely to be less than they would have been in the absence of the restrictions. Although most of this text deals with developing principles of firm behavior based on theories of unconstrained profit maximization, we will also introduce the powerful mathematical techniques of Lagrange multipliers and linear programming for dealing with constrained optimization problems.

ACCOUNTING PROFIT VERSUS ECONOMIC PROFIT

To say that products that can be produced profitably will be, and those that cannot be produced profitably will not begs the question of what we mean by "profit." What is commonly thought of as profit by the accountant may not match the meaning assigned to the term by an economist. An econ-

omist's notion of profit goes back to the basic fact that resources are scarce and have alternative uses. To use a certain set of resources to produce a good or service means that certain alternative production possibilities were forgone. Costs in economics have to do with forgoing the opportunity to produce alternative goods and services. The economic, or opportunity, cost of any resource in producing some good or service is its value or worth in its next best alternative use.

Given the notion of opportunity costs, economic costs are the payments a firm must make, or incomes it must provide, to resource suppliers to attract these resources away from alternative lines of production. Economic costs (TC) include all relevant opportunity costs. These payments or incomes may be either explicit, "out-of-pocket" or cash expenditures, or implicit. Implicit costs represent the value of resources used in the production process for which no direct payment is made. This value is generally taken to be the money earnings of resources in their next best alternative employment. When a computer software programmer quits his or her job to open a consulting firm, the forgone salary is an example of an implicit cost. When the owner of an office building decides to open a hobby shop, the forgone rental income from that store is an example of an implicit cost. When a housewife decides to redeem a certificate of deposit to establish a day-care center for children, the forgone interest earnings represent an implicit cost. In short, any sacrifice incurred when the decision is made to produce a good or service must be taken into account if the full impact of that decision is to be correctly assessed. These relationships may be summarized as follows:

$$\text{Accounting profit: } \pi_A = TR - TC_{\text{explicit}} \tag{1.6}$$

$$\text{Economic profit: } \pi = TR - TC = TR - TC_{\text{explicit}} - TC_{\text{implicit}} \tag{1.7}$$

Problem 1.3. Andrew operates a small shop specializing in party favors. He owns the building and supplies all his own labor and money capital. Thus, Andrew incurs no explicit rental or wage costs. Before starting his own business Andrew earned $1,000 per month by renting out the store and earned $2,500 per month as a store manager for a large department store chain. Because Andrew uses his own money capital, he also sacrificed $1,000 per month in interest earned on U.S. Treasury bonds. Andrew's monthly revenues from operating his shop are $10,000 and his total monthly expenses for labor and supplies amounted to $6,000. Calculate Andrew's monthly accounting and economic profits.

Solution. Total accounting profit is calculated as follows:

Total revenue	$10,000
Total explicit costs	6,000
Accounting profit	$4,000

Andrew's accounting profit appears to be a healthy $4,000 per month. However, if we take into account Andrew's implicit costs, the story is quite different. Total economic profit is calculated as follows:

Total revenue	$10,000
Total explicit costs	6,000
Forgone rent	1,000
Forgone salary	2,500
Forgone interest income	1,000
Total implicit costs	4,500
Total economic costs	10,500
Economic profit (loss)	$ (500)

Economic profits are equal to total revenue less total economic costs, which is the sum of explicit and implicit costs. Accounting profits, on the other hand, are equal to total revenue less total explicit costs.

It is, of course, a simple matter to make accounting profit equivalent to economic profit by making explicit all relevant implicit costs. Suppose, for example, that an individual quits a $40,000 per year job as the manager a family restaurant to open a new restaurant. Since this is a sacrifice incurred by the budding restauranteur, the forgone salary is an implicit cost. On the other hand, this implicit cost can easily be made explicit by putting the restaurant owner "on the books" for a salary of $40,000. The somewhat arbitrary distinction between explicit and implicit costs is illustrated in the following problem.

Problem 1.4. Adam is the owner of a small grocery store in a busy section of Boulder, Colorado. Adam's annual revenue is $200,000 and his total explicit cost (Adam pays himself an annual salary of $30,000) is $180,000 per year. A supermarket chain wants to hire Adam as its general manager for $60,000 per year.

a. What is the opportunity cost to Adam of owning and managing the grocery store?
b. What is Adam's accounting profit?
c. What is Adam's economic profit?

Solution

a. Opportunity cost is the $60,000 in forgone salary that Adam might have earned had he decided to work as general manager for the supermarket chain.
b. $\pi_A = TR - TC_{explicit} = \$200{,}000 - \$180{,}000 = \$20{,}000$
c. $\pi = TR - TC_{explicit} - TC_{implicit} = \$200{,}000 - \$180{,}000 - \$30{,}000$
   $= -\$10{,}000$
   Another way of looking at this problem is to consider Adam's forgone income following his decision to continue to operate the grocery store. Adam's forgone income may be summarized as follows:

$$\pi_A = \text{grocery store salary} - \text{supermarket salary}$$
$$= \$20,000 + \$30,000 - \$60,000 = -\$10,000$$

This is the same as the result in part b, since the grocery store salary less the supermarket salary is just the opportunity cost as defined.

NORMAL PROFIT

Another important concept in economics is that of *normal profit*. Normal profit, sometimes referred to as *normal rate of return*, is the level of profit required to keep a firm engaged in a particular activity. Alternatively, normal profit represents the rate of return on the next best alternative investment of equivalent risk. It is the level of profit necessary to keep people from pulling their investments in search of higher rates of return. If a firm is well established and has a good earnings record, and if there is little to no possibility of financial loss during a specified period of time, then the normal rate of return is approximately equal to the interest rate on a risk-free government bond. If the firm's earnings are erratic and its future prospects questionable, then the *risk-free rate of return* must be augmented by a *risk premium*. Either way, normal profit is a form of opportunity cost. Viewed in this way, it is easy to see that normal profit represents a component of total economic cost.

Definition: Normal profit refers to the level of profits required to keep a firm engaged in a particular activity. Alternatively, normal profit represents the rate of return on the next best alternative investment of equivalent risk. Normal profits are a kind of opportunity cost.

Normal profit is an implicit cost of doing business. To see the relationship between economic profit and normal profit, let us assume that we have explicitly accounted for all economic costs except for normal profits. Define the firm's total operating costs TC_O as total economic costs TC minus normal profit π_N. This relationship is summarized in Equation (1.8).

$$TC_o = TC - \pi_N \tag{1.8}$$

Definition: Total operating cost is the difference between total economic cost and normal profit.

From Equation (1.8) we may define the firm's total economic profit as the sum of *total operating profit* and normal profit. This relationship is summarized by the relation.

$$\pi = TR - TC = TR - TC_o - \pi_N = \pi_o - \pi_N \tag{1.9}$$

where $\pi_O = (TR - TC_O)$ is referred to as the firm's *total operating profit*.

Definition: Operating profit is the sum of economic profit and normal profit.

quantity, quality, and price of the good or service being exchanged. Asymmetric information exists when some market participants have more and better information about the goods and services being exchanged. Fraud can arise in the presence of asymmetric information. In extreme cases, the knowledge that some market participants have access to privileged information may result in a complete breakdown of the market, such as might occur if it became widely believed that stock market transactions were dominated by insider trading.

Goods and services are said to be "private" when all the production costs and consumption benefits are borne exclusively by the market participants. That is, there are no indirect, third-party effects. Such third-party effects, called *externalities*, may affect either consumers or producers. The most common example of a *negative externality* in production is pollution. Finally, "market power" refers to the ability to influence the market price of a good or service by shifting the demand or supply curve.

A violation of any of the three assumptions just given could lead to failure of the market to provide socially optimal levels of particular goods or services. When this occurs, direct or indirect government intervention in the market may be deemed to be in the public's best interest. Market failure and government intervention will be discussed at some length in Chapter 15.

For many readers, most of what is presented in this chapter will constitute a review of material learned in a course in the fundamentals of economics. Students who are familiar with the application of elementary algebraic methods to the concepts of demand, supply, and the market process may proceed to Chapter 4 without any loss of continuity.

THE LAW OF DEMAND

The assumption of profit-maximizing behavior assumes that owners and managers know the demand for the firm's good or service. The demand function asserts that there is a measurable relationship between the price that a company charges for its product and the number of units that buyers are willing and able to purchase during a specified time period. Economists refer to this behavioral relationship as the *law of demand*, which is sometimes called the *first fundamental law of economics*.

Definition: The law of demand states that the quantity demanded of a good or service is inversely related to the selling price, *ceteris paribus* (all other determinants remaining unchanged).

The term "law of demand" is actually a misnomer. As discussed in Chapter 1, laws are facts. Laws are assertions of fact. Laws predict events with certainty. By contrast, theories are probabilistic statements of cause

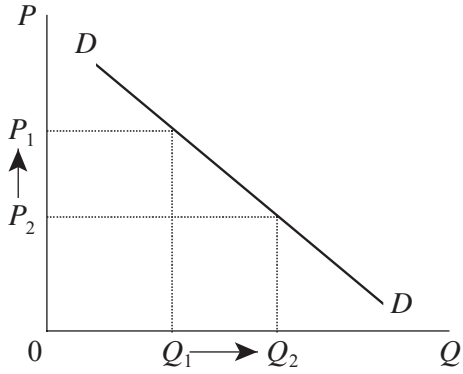

FIGURE 3.1 The demand curve.

and effect. The law of demand is a theory, as is invariably the case when human nature is involved.

Symbolically, the law of demand may be summarized as

$$Q_D = f(P) \tag{3.1a}$$

and

$$\frac{dQ_D}{dP} < 0 \tag{3.1b}$$

Equation (3.1a) states that Q_D, the quantity demanded of a good or service, is functionally related to the selling price P. Inequality (3.1b) asserts that quantity demanded and price are inversely related. This relationship is illustrated in Figure 3.1. The downward-sloping *demand curve* illustrates the inverse relationship between the quantity demanded of a good or service and its selling price.

The validity of the law of demand may be argued on the basis of common sense and simple observation. At a more sophisticated level, the validity of the law of demand may be argued on the basis of diminishing marginal utility and income and substitution effects.[1]

INCOME EFFECT

For most goods, the *income effect* asserts that as a product's price declines (increases), an individual's real income (purchasing power) increases (decreases). The increase in real purchasing power resulting from a fall in prices enables the individual to consume greater quantities of a commodity, while the opposite is true for an increase in prices. In other words, an

[1] A formal derivation of the demand curve is presented in Appendix 3A.

increase in real purchasing power generally (although not always) leads to increase in quantity demanded. The goods of the types for which this phenomenon holds are referred to as *normal goods*. Unfortunately, the income effect does not always have the expected positive effect on the quantity demanded of a good. In some cases, as an individual's purchasing power increases, the quantity demand for that good falls. Goods of these types are called *inferior goods*. Examples of such goods may be potatoes, bus tickets, soup bones, and bologna. We will return to this issue shortly when considering separately the effect of changes in money income on the demand for goods and services.

SUBSTITUTION EFFECT

The more powerful *substitution effect* entails no such ambiguitys. The substitution effect reflects changes in consumers' opportunity costs. The substitution effect states that as a product's price declines, consumers will substitute the now less expensive product for similar goods that are more expensive.[2]

In the majority of cases, the income effect and the substitution effect complement and reinforce each other. That is, a decline in the price of a good will not only have a positive substitution effect, but will have a positive income effect as well. As a result, the ordinary demand curve will be downward sloping. Even in the case of inferior goods, where the income effect is negative, the ordinary demand curve will exhibit a downward slope because the substitution effect, which is always positive with a drop in price, outweighs the negative income effect.[3]

THE MARKET DEMAND CURVE

The law of demand is a theoretical explanation of the expected behavior of individual economic units when confronted with a change in the price of a commodity. Yet our concern, at the present, is less with the behavior

[2] The interaction of the income and substitution effects are summarized in the Slutsky equation, the proof of which may be found in Silberberg (1990), Chapter 10.

[3] If, on the other hand, the negative income effect, which is associated with inferior goods, outweighs the always positive substitution effect, the ordinary demand curve will be upward sloping! This is precisely what happened in 1845 when famine in Ireland greatly increased the price of potatoes, which in turn caused real incomes to fall sharply. Irish families ended up consuming more rather than less of the high-priced potatoes. Why? Being forced to pay so high a price for a basic necessity made it impossible for the average family to purchase any meat at all, and hence most were forced to become even more dependent on potatoes. The explanation of this curious phenomenon was first attributed to the Victorian economist Sir Robert Giffen. As a result, such goods have been dubbed Giffen goods. See Samuelson and Nordhaus (1985), p. 416.

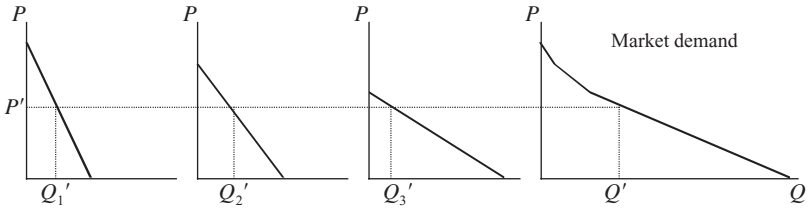

FIGURE 3.2 The market demand curve is the horizontal summation of the individual demand curves.

of individual economic agents than with the market demand for the product of an industry. We must, therefore, extend our analysis to justify what we would hope to be a downward-sloping market demand curve. To derive the hypothetical market demand function for a particular industry's product, let us first consider three hypothetical individual demand functions for the product in question

$$Q_{D,1} = a_1 + b_1 P \tag{3.2a}$$

$$Q_{D,2} = a_2 + b_2 P \tag{3.2b}$$

$$Q_{D,3} = a_3 + b_3 P \tag{3.2c}$$

where the $Q_{D,i}$ terms represent the individual's demand for the commodity, the a_i terms are positive constants, and the b_i terms the unit change in quantity demanded given a change in the selling price.

Definition: The market demand curve is the horizontal summation of the individual demand curves.

For any given price, the market demand curve is the sum of the horizontal distances from the vertical axis to each individual demand curve. Summing together Equations (3.2) we get

$$Q_{D,1} + Q_{D,2} + Q_{D,3} = (a_1 + a_2 + a_3) + (b_1 + b_2 + b_3)P \tag{3.3}$$

or

$$Q_D = a + bP \tag{3.4}$$

where $a = a_1 + a_2 + a_3$ and $b = b_1 + b_2 + b_3$. In general, for the n-consumer case

$$Q_{D,M} = \sum_{i=1\to n} Q_{D,i} = \sum_{i=1\to n} a_t + \sum_{i=1\to n} b_i P \tag{3.5}$$

where $Q_{D,M}$ is market demand. Equation (3.5) is illustrated in Figure 3.2.

Problem 3.1. Suppose that the total market demand for a product comprises the demand of three individuals with identical demand equations.

$$Q_{D,1} = Q_{D,2} = Q_{D,3} = 50 - 25P$$

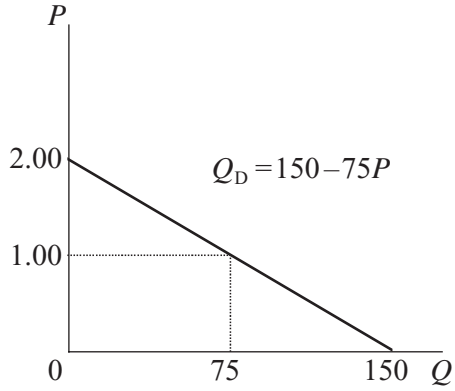

FIGURE 3.3 Diagrammatic solution to problem 3.1.

What is the market demand equation for this product?

Solution. The market demand curve is the horizontal summation of the individual demand curves. The market demand equation is

$$\begin{aligned} Q_D &= Q_{D,1} + Q_{D,2} + Q_{D,3} \\ &= (50 - 25P) + (50 - 25P) + (50 - 25P) = 150 - 75P \end{aligned}$$

This result is illustrated in Figure 3.3.

Problem 3.2. Suppose that the total market demand for a product consists of the demands of individual 1 and individual 2. The demand equations of the two individuals are given by the following equations:

$$Q_{D,1} = 20 - 2P$$

$$Q_{D,2} = 40 - 5P$$

What is the market demand equation for this product?

Solution. Solving the individuals' demand curves for price we obtain

$$P = 10 - 0.5Q_{D,1}$$

$$P = 8 - 0.2Q_{D,2}$$

These demand equations are illustrated in Figure 3.4.

It should be apparent from Figure 3.4 that for $P > \$8$, only individual 1 will purchase units of commodity Q. Thus, the market demand curve is $Q_{D,1} = 20 - 2P$. For prices $P \leq \$8$, both individuals 1 and 2 will purchase units of the commodity Q. Thus, the market demand curve is $Q = Q_{D,1} + Q_{D,2} =$

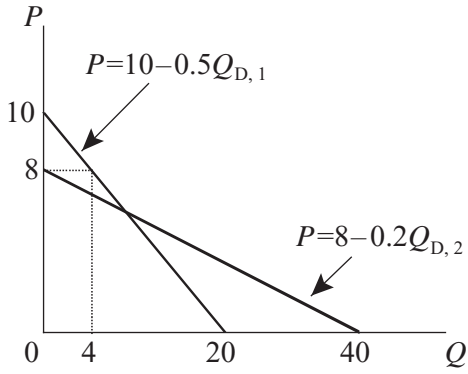

FIGURE 3.4 Demand curves for individuals 1 and 2 in problem 3.2.

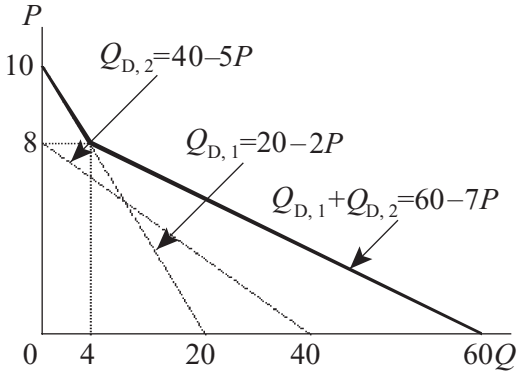

FIGURE 3.5 The market demand curve as the sum of the demand curves for individuals 1 and 2 in problem 3.2.

$(20 - 2P) + (40 - 5P) = 60 - 7P$. The market demand curve for commodity Q is illustrated by the heavy line in Figure 3.5.

The reader will note that the demand curve for commodity Q is discontinuous at $P = \$8$. Figure 3.5 is often referred to as a "kinked" demand curve. Compare this with the smooth and continuous curve in Problem 3.1 (Figure 3.3), in which both individuals enter the market at the same time (i.e., for $P < \$2$).

The market demand curve establishes a relationship between the product's price and the quantity demanded; all other determinants of market demand are held constant. The relationship between changes in price and changes in quantity demanded are illustrated as movements along the demand curve. When economists refer to a change in the quantity demanded (in response to a change in price), they are referring to a movement along the demand curve. As we will see, this is to be distinguished

from a change in demand (illustrated as a shift in the entire demand curve), which results when a determinant of demand, other than its selling price, is changed. This semantic distinction is made necessary because two-dimensional representations of a demand function can accommodate a relationship between two variables only—in this case price and quantity, the independent and dependent variables, respectively.[4] What are some of these other determinants of demand?

OTHER DETERMINANTS OF MARKET DEMAND

We know, of course, that price is not the only factor that influences an economic agent's decision to purchase quantities of a given good or service. Other demand determinants include income, consumer preferences, the prices of related goods, price expectations, and population.

INCOME (I)

Typically, an increase in a consumer's money income will result in increased purchases of goods and services, other things remaining equal (including the selling price). More precisely, a *ceteris paribus* increase in an individual's money income will usually lead to an increase in the demand for a good or service. Conceptually, this is not the same thing as an increase in quantity demanded of a good or service due to an increase in an individual's real income that has resulted from a fall in price. Similarly, a *ceteris paribus* decline in an individual's money income will result in a decrease in demand. As before, such goods are called normal goods. Most goods and services fall into this category. An increase in demand for a normal good resulting from an increase in income may be illustrated in Figure 3.6.

In the case of so-called interior goods, however, the demand for a good or a service actually declines with an increase in money income. The result would be a leftward shift in the demand curve. Inferior goods are largely a matter of individual preferences. As their income rises, some individuals prefer to substitute train or plane travel for slower, and presumably less expensive, long-distance bus rides. On the other hand, other people really like riding buses. For this group, long-distance bus travel is a normal good.

[4] Although it is possible to represent a three-dimensional object on a two-dimensional surface, such as in a photograph, in practice, drawing such diagrams is quite difficult. Moreover, beyond three dimensions, graphically illustrating a relationships that includes, say, four variables is impossible, although depicting its three-dimensional shadow on a two-dimensional surface is not! After all, we live in three-dimensional space, so what does a fourth-dimensional object look like?

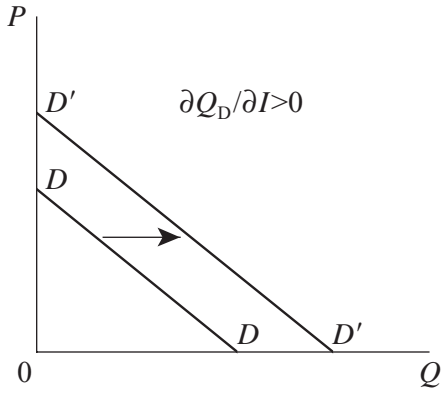

FIGURE 3.6 An increase in demand resulting from an increase in money income.

TASTES OR PREFERENCES (T)

Another determinant of market demand is individuals' tastes, or preferences, for a particular product. After seeing a McDonald's television commercial, for example, one person might be compelled to purchase an increased quantity of hamburgers, even though the price of hamburgers had not fallen or his income had remained the same. This increased demand for hamburgers would be represented as a rightward shift in the demand curve. Similarly, if after reading an article in the *New York Times* about the health dangers associated with diets high in animal fat and salt, the same person might decide to cut down on his intake of hamburgers, which would be shown as a left-shift in his demand curve for hamburgers. The effect of an increase in taste is similar to that depicted for an increase in income in Figure 3.6.

PRICES OF RELATED GOODS: SUBSTITUTES (P_s) AND COMPLEMENTS (P_c)

The prices of related goods can also affect the demand for a particular good or service. Related goods are generally classified as either substitute goods or complementary goods.

Substitutes are goods that consumers consider to be closely related. As the price of good X rises, the quantity demanded of that good will fall according to the law of demand. If good Y is a substitute for good X, the demand for good Y will rise as the consumer substitutes into it. The willingness of the consumer to substitute one good for another varies from good to good and is rarely an either/or proposition. For example, although not perfect substitutes for most consumers, Coca-Cola and Pepsi-Cola might be classified as "close" substitutes. Other examples of goods that may