Chapter 8

Comparative-Static Analysis of General-Function Models

The study of partial derivatives has enabled us, in Chap. 7, to handle the simpler type of comparative-static problems, in which the equilibrium solution of the model can be explicitly stated in the reduced form. In that case, partial differentiation of the solution will directly yield the desired comparative-static information. You will recall that the definition of the partial derivative requires the absence of any functional relationship among the independent variables (say, x_i), so that x_1 can vary without affecting the values of $x_2, x_3, \ldots, x_n$. As applied to comparative-static analysis, this means that the parameters and/or exogenous variables which appear in the reduced-form solution must be mutually independent. Since these are indeed defined as predetermined data for purposes of the model, the possibility of their mutually affecting one another is inherently ruled out. The procedure of partial differentiation adopted in Chap. 7 is therefore fully justifiable.

However, no such expediency should be expected when, owing to the inclusion of general functions in a model, no explicit reduced-form solution can be obtained. In such cases, we will have to find the comparative-static derivatives directly from the originally given equations in the model. Take, for instance, a simple national-income model with two endogenous variables Y and C:

$$\begin{aligned} Y &= C + I_0 + G_0 \\ C &= C(Y, T_0) \qquad [T_0\text{: exogenous taxes}] \end{aligned}$$

which is reducible to a single equation (an equilibrium condition)

$$Y = C(Y, T_0) + I_0 + G_0$$

to be solved for Y^*. Because of the general form of the C function, however, no explicit solution is available. We must, therefore, find the comparative-static derivatives directly from this equation. How might we approach the problem? What special difficulty might we encounter?

Let us suppose that an equilibrium solution Y^* does exist. Then, under certain rather general conditions (to be discussed in Section 8.5), we may take Y^* to be a differentiable function of the exogenous variables I_0, G_0, and T_0. Hence we may write the equation

$$Y^* = Y^*(I_0, G_0, T_0)$$

even though we are unable to determine explicitly the form which this function takes. Furthermore, in some neighborhood of the equilibrium value Y^*, the following identical equality will hold:

$$Y^* \equiv C(Y^*, T_0) + I_0 + G_0$$

This type of identity will be referred to as an *equilibrium identity* because it is nothing but the equilibrium condition with the Y variable replaced by its equilibrium value Y^*. Now that Y^* has entered into the picture, it may seem at first blush that simple partial differentiation of this identity will yield any desired comparative-static derivative, say, $\partial Y^*/\partial T_0$. This, unfortunately, is not the case. Since Y^* is a function of T_0, the two arguments of the C function are *not* independent. Specifically, T_0 can in this case affect C not only *directly*, but also *indirectly* via Y^*. Consequently, partial differentiation is no longer appropriate for our purposes. How, then, do we tackle this situation?

The answer is that we must resort to *total differentiation* (as against partial differentiation). Based on the notion of *total differentials*, the process of total differentiation can lead us to the related concept of *total derivative*, which measures the rate of change of a function such as $C(Y^*, T_0)$ with respect to the argument T_0, when T_0 also affects the other argument, Y^*. Thus, once we become familiar with these concepts, we shall be able to deal with functions whose arguments are not all independent, and that would remove the major stumbling block we have so far encountered in our study of the comparative statics of a general-function model. As a prelude to the discussion of these concepts, however, we should first introduce the notion of *differentials*.

8.1 Differentials

The symbol dy/dx, for the derivative of the function $y = f(x)$, has hitherto been regarded as a single entity. We shall now reinterpret it as a ratio of two quantities, dy and dx.

Differentials and Derivatives

By definition, the derivative $dy/dx = f'(x)$ is the limit of a difference quotient:

$$\frac{dy}{dx} = f'(x) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} \qquad \textbf{(8.1)}$$

Thus, by itself, $\Delta y/\Delta x$ (without requiring $\Delta x \to 0$) is not equal to dy/dx. If we denote the discrepancy between the two quotients by δ, we can write

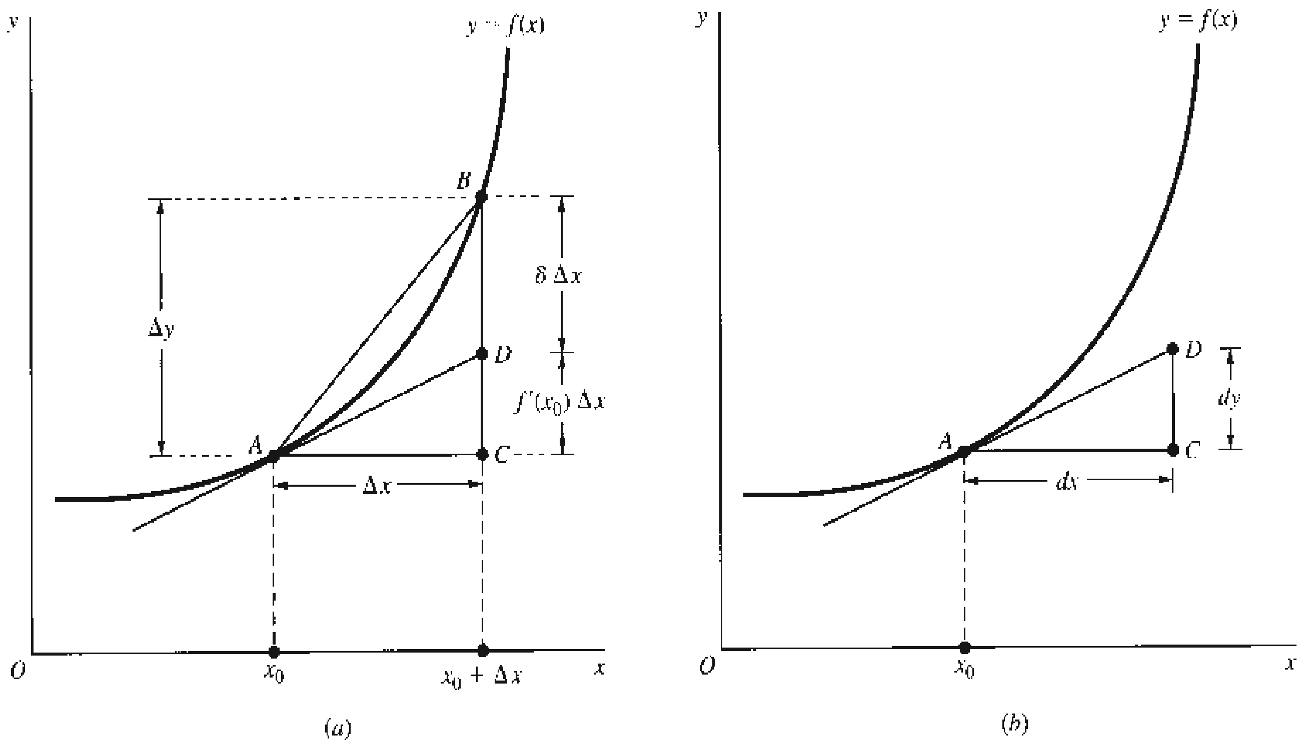
$$\frac{\Delta y}{\Delta x} - \frac{dy}{dx} = \delta \qquad \text{where} \qquad \delta \to 0 \quad \text{as} \quad \Delta x \to 0 \qquad [\text{by (8.1)}] \qquad \textbf{(8.2)}$$

Multiplying (8.2) through by Δx, and rearranging, we have

$$\Delta y = \frac{dy}{dx}\Delta x + \delta\,\Delta x \qquad \text{or} \qquad \Delta y = f'(x)\Delta x + \delta\,\Delta x \qquad \textbf{(8.3)}$$

This equation describes the change in y (Δy) that results from a specific—not necessarily small—change in x (Δx) from any starting value of x in the domain of the function

FIGURE 8.1



$y = f(x)$. But it also suggests that we can, by ignoring the discrepancy term $\delta\,\Delta x$, use the $f'(x)\,\Delta x$ term as an approximation to the true Δy value, where the approximation gets progressively better as Δx gets progressively smaller.

In Fig. 8.1*a*, when x changes from x_0 to $x_0 + \Delta x$, a movement from point A to point B occurs on the graph of $y = f(x)$. The true Δy is measured by the distance CB, and the ratio of the two distances $CB/AC = \Delta y/\Delta x$ can be read from the slope of line segment AB. But if we draw a tangent line AD through point A, and use AD in place of AB to approximate the value of Δy, we obtain distance CD, which leaves distance DB as the discrepancy or error of approximation. Since the slope of AD is $f'(x_0)$, distance CD is equal to $f'(x_0)\,\Delta x$ and, by (8.3), distance DB is equal to $\delta\,\Delta x$. Obviously, as Δx decreases, point B would slide along the curve toward point A, thereby reducing the discrepancy and making $f'(x)$ or dy/dx a better approximation to $\Delta y/\Delta x$.

Focusing on the tangent line AD, and taking the distance CD as an approximation to CB, let us relabel the distances AC and CD by dx and dy, respectively, as in Fig. 8.1*b*. Then

$$\frac{dy}{dx} = \text{slope of tangent } AD = f'(x)$$

and, after multiplying through by dx, we get

$$dy = f'(x)\,dx \qquad \textbf{(8.4)}$$

The derivative $f'(x)$ can then be reinterpreted as the factor of proportionality between the two finite changes dy and dx. Accordingly, given a specific value of dx, we can multiply it by $f'(x)$ to get dy as an approximation to Δy, with the understanding that the smaller the Δx, the better the approximation. The quantities dx and dy are called the *differentials* of x and y, respectively.

A few remarks are in order regarding differentials as mathematical entities. First, while dx is an independent variable, dy is a dependent variable. Specifically, dy is a function of x as well as of dx: It depends on x because a different position for x_0 in Fig. 8.1 would mean a different location for point A and for its tangent line; it depends on dx because a different magnitude of dx would mean a different position for point C as well as a different distance CD. Second, if $dx = 0$, then $dy = 0$, because point B would in that case coincide with point A. But if $dx \neq 0$, then it is possible to divide dy by dx to get $f'(x)$, just as we can multiply dx by $f'(x)$ to get dy. Third, the differential dy can be expressed only in terms of some other differential(s)—here, dx. This is because our context calls for the coupling of a dependent change dy with an independent change dx. While it makes sense to write $dy = f'(x)\,dx$, it is not meaningful to chop away the dx term on the right and write $dy = f'(x)$. The coupling of the two changes is effected through the derivative $f'(x)$, which may be viewed as a "converter" that serves to translate a given change dx into a counterpart change dy.

The process of finding the differential dy from a given function $y = f(x)$ is called *differentiation*. Recall that we have been using this term as a synonym for derivation, without having given an adequate explanation. In light of our interpretation of a derivative as a quotient of two differentials, however, the rationale of the term becomes self-evident. It is still somewhat ambiguous, though, to use the single term "differentiation" to refer to the process of finding the differential dy as well as to that of finding the derivative dy/dx. To avoid confusion, the usual practice is to qualify the word *differentiation* with the phrase "with respect to x" when we take the derivative dy/dx.

Differentials and Point Elasticity

To illustrate the economic application of differentials, let us consider the notion of the elasticity of a function. Given a demand function $Q = f(P)$, for instance, its elasticity is defined as $(\Delta Q/Q)/(\Delta P/P)$. Using the idea of approximation explained in Fig. 8.1, we can replace the independent change ΔP and the dependent change ΔQ with the differentials dP and dQ, respectively, to get an approximation elasticity measure known as the *point elasticity* of demand and denoted by ε_d (the Greek letter epsilon, for "elasticity"):[†]

$$\varepsilon_d \equiv \frac{dQ/Q}{dP/P} = \frac{dQ/dP}{Q/P} \tag{8.5}$$

Observe that on the extreme right of the expression we have rearranged the differentials dQ and dP into a ratio dQ/dP, which can be construed as the derivative, or the *marginal* function, of the demand function $Q = f(P)$. Since we can interpret similarly the ratio Q/P in the denominator as the *average* function of the demand function, the point elasticity of demand ε_d in (8.5) is seen to be the ratio of the marginal function to the average function of the demand function.

[†] The point-elasticity measure can alternatively be interpreted as the limit of $\dfrac{\Delta Q/Q}{\Delta P/P} = \dfrac{\Delta Q/\Delta P}{Q/P}$ as $\Delta P \to 0$, which gives the same result as (8.5).

Indeed, this last-described relationship is valid not only for the demand function but also for any other function, because for any given *total* function $y = f(x)$ we can write the formula for the point elasticity of y with respect to x as

$$\varepsilon_{yx} = \frac{dy/dx}{y/x} = \frac{\text{marginal function}}{\text{average function}} \tag{8.6}$$

As a matter of convention, the *absolute* value of the elasticity measure is used in deciding whether the function is elastic at a particular point. In the case of a demand function, for instance, we stipulate:

$$\text{The demand is} \begin{Bmatrix} \text{elastic} \\ \text{of unit elasticity} \\ \text{inelastic} \end{Bmatrix} \text{ at a point when } |\varepsilon_d| \gtreqless 1.$$

Example 1

Find ε_d if the demand function is $Q = 100 - 2P$. The marginal function and the average function of the given demand are

$$\frac{dQ}{dP} = -2 \qquad \text{and} \qquad \frac{Q}{P} = \frac{100 - 2P}{P}$$

so their ratio will give us

$$\varepsilon_d = \frac{-P}{50 - P}$$

As written, the elasticity is shown as a function of P. As soon as a specific price is chosen, however, the point elasticity will be determinate in magnitude. When $P = 25$, for instance, we have $\varepsilon_d = -1$, or $|\varepsilon_d| = 1$, so that the demand elasticity is unitary at that point. When $P = 30$, in contrast, we have $|\varepsilon_d| = 1.5$; hence, demand is elastic at that price. More generally, it may be verified that we have $|\varepsilon_d| > 1$ for $25 < P < 50$ and $|\varepsilon_d| < 1$ for $0 < P < 25$ in the present example. (Can a price $P > 50$ be considered meaningful here?)

Example 2

Find the point elasticity of supply ε_s from the supply function $Q = P^2 + 7P$, and determine whether the supply is elastic at $P = 2$. Since the marginal and average functions are, respectively,

$$\frac{dQ}{dP} = 2P + 7 \qquad \text{and} \qquad \frac{Q}{P} = P + 7$$

their ratio gives us the elasticity of supply

$$\varepsilon_s = \frac{2P + 7}{P + 7}$$

When $P = 2$, this elasticity has the value $11/9 > 1$; thus the supply is elastic at $P = 2$.

At the risk of digressing a trifle, it may also be added here that the interpretation of the ratio of two differentials as a derivative—and the consequent transformation of the elasticity formula of a function into a ratio of its marginal to its average—makes possible a quick way of determining the point elasticity graphically. The two diagrams in Fig. 8.2 illustrate the cases, respectively, of a negatively sloped curve and a positively sloped curve. In each case, the value of the marginal function at point A on the curve, or at $x = x_0$ in the domain, is measured by the slope of the tangent line AB. The value of the average function, on the

FIGURE 8.2

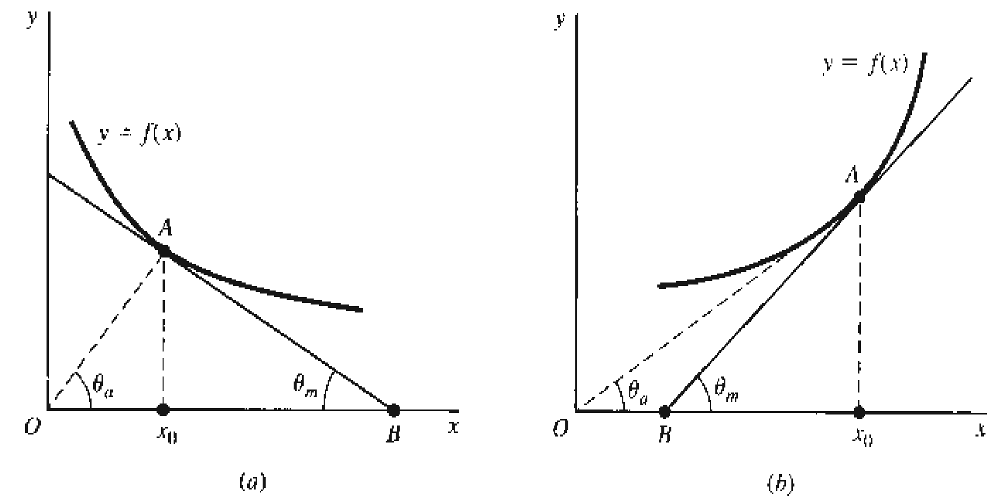
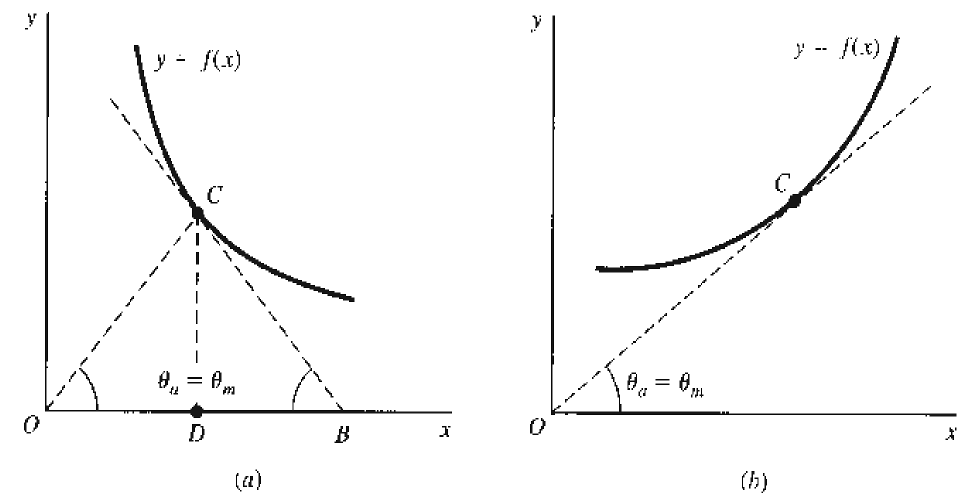


FIGURE 8.3



other hand, is in each case measured by the slope of line OA (the line joining the point of origin with the given point A on the curve, like a radius vector), because at point A we have $y = x_0A$ and $x = Ox_0$, so that the average is $y/x = x_0A/Ox_0 =$ slope of OA. The elasticity at point A can thus be readily ascertained by comparing the *numerical* values of the two slopes involved: If AB is steeper than OA, the function is elastic at point A; in the opposite case, it is inelastic at A. Accordingly, the function pictured in Fig. 8.2*a* is inelastic at A (or at $x = x_0$), whereas the one in Fig. 8.2*b* is elastic at A.

Moreover, the two slopes under comparison are directly dependent on the respective sizes of the two angles θ_m and θ_a (Greek letter theta; the subscripts m and a indicate marginal and average, respectively). Thus we may, alternatively, compare these two angles instead of the two corresponding slopes. Referring to Fig. 8.2 again, you can see that $\theta_m < \theta_a$ at point A in diagram a, indicating that the marginal falls short of the average in numerical value; thus the function is inelastic at point A. The exact opposite is true in Fig. 8.2*b*.

Sometimes, we are interested in locating a point of unitary elasticity on a given curve. This can now be done easily. If the curve is negatively sloped, as in Fig. 8.3*a*, we should find a point C such that the line OC and the tangent BC will make the same-sized angle with the x axis, though in the opposite direction. In the case of a positively sloped curve, as in Fig. 8.3*b*, one has only to find a point C such that the tangent line at C, when properly extended, passes through the point of origin.

We must warn you that the graphical method just described is based on the assumption that the function $y = f(x)$ is plotted with the dependent variable y on the vertical axis. In particular, in applying the method to a demand curve, we should make sure that Q is on the vertical axis. (Now suppose that Q is actually plotted on the horizontal axis. How should our method of reading the point elasticity be modified?)

EXERCISE 8.1

1. Find the differential dy, given:
   (a) $y = -x(x^2+3)$ (b) $y = (x-8)(7x+5)$ (c) $y = \frac{x}{x^2+1}$
2. Given the import function $M = f(Y)$, where M is imports and Y is national income, express the income elasticity of imports ε_{MY} in terms of the propensities to import.
3. Given the consumption function $C = a - bY$ (with $a > 0$; $0 < b < 1$):
   (a) Find its marginal function and its average function.
   (b) Find the income elasticity of consumption ε_{CY}, and determine its sign, assuming $Y > 0$.
   (c) Show that this consumption function is inelastic at all positive income levels.
4. Find the point elasticity of demand, given $Q = k/P^n$, where k and n are positive constants.
   (a) Does the elasticity depend on the price in this case?
   (b) In the special case where $n = 1$, what is the shape of the demand curve? What is the point elasticity of demand?
5. (a) Find a positively sloped curve with a constant point elasticity everywhere on the curve.
   (b) Write the equation of the curve, and verify by (8.6) that the elasticity is indeed a constant.
6. Given $Q = 100 - 2P + 0.02Y$, where Q is quantity demanded, P is price, and Y is income, and given $P = 20$ and $Y = 5{,}000$, find the
   (a) Price elasticity of demand.
   (b) Income elasticity of demand.

8.2 Total Differentials

The concept of differentials can easily be extended to a function of two or more independent variables. Consider a saving function

$$S = S(Y, i) \tag{8.7}$$

where S is savings, Y is national income, and i is the interest rate. This function is assumed—as all the functions we shall use here will be assumed—to be continuous and to possess continuous (partial) derivatives, or, symbolically, $f \in C'$. The partial derivative $\partial S/\partial Y$ measures the marginal propensity to save. Thus, for any change in Y, dY, the resulting change in S can be approximated by the quantity $(\partial S/\partial Y)\,dY$, which is comparable to the right-hand expression in (8.4). Similarly, given a change in i, di, we may take $(\partial S/\partial i)\,di$

as the approximation to the resulting change in S. The total change in S is then approximated by the differential

$$dS = \frac{\partial S}{\partial Y}\,dY + \frac{\partial S}{\partial i}\,di \tag{8.8}$$

or, in an alternative notation,

$$dS = S_Y\,dY + S_i\,di$$

Note that the two partial derivatives S_Y and S_i again play the role of "converters" that serve to convert the changes dY and di, respectively, into a corresponding change dS. The expression dS, being the *sum* of the approximate changes from both sources, is called the *total differential* of the saving function. And the process of finding such a total differential is called *total differentiation*. In contrast, the two additive components to the right of the equals sign in (8.8) are referred to as the *partial differentials* of the saving function.

It is possible, of course, that Y may change while i remains constant. In that case, $di = 0$, and the total differential will reduce to $dS = (\partial S/\partial Y)\,dY$. Dividing both sides by dY, we get

$$\frac{\partial S}{\partial Y} = \left(\frac{dS}{dY}\right)_{i\text{ constant}}$$

Thus it is clear that the partial derivative $\partial S/\partial Y$ can also be interpreted, in the spirit of Fig. 8.1b, as the ratio of two differentials dS and dY, with the proviso that i, the other independent variable in the function, is held constant. Analogously, we can interpret the partial derivative $\partial S/\partial i$ as the ratio of the differential dS (with Y held constant) to the differential di. Note that although dS and di can now each stand alone as a differential, the expression $\partial S/\partial i$ remains as a single entity.

The more general case of a function of n independent variables can be exemplified by, say, a utility function in the general form

$$U = U(x_1, x_2, \ldots, x_n) \tag{8.9}$$

The total differential of this function can be written as

$$dU = \frac{\partial U}{\partial x_1}\,dx_1 + \frac{\partial U}{\partial x_2}\,dx_2 + \cdots + \frac{\partial U}{\partial x_n}\,dx_n \tag{8.10}$$

or $$dU = U_1\,dx_1 + U_2\,dx_2 + \cdots + U_n\,dx_n = \sum_{i=1}^{n} U_i\,dx_i$$

in which each term on the right side indicates the approximate change in U resulting from a change in one of the independent variables. Economically, the first term, $U_1\,dx_1$, means the marginal utility of the first commodity times the increment in consumption of that commodity, and similarly for the other terms. The sum of these, dU, thus represents the total approximate change in utility originating from all possible sources of change. As the reasoning in (8.3) shows, dU, as an approximation, tends toward the true change ΔU as all the dx_i terms tend to zero.

Like any other function, the saving function (8.7) and the utility function (8.9) can both be expected to give rise to point-elasticity measures similar to that defined in (8.6). But each elasticity measure must in these instances be defined in terms of the change in *one* of the independent variables only; there will thus be *two* such elasticity measures to the saving function, and n of them to the utility function. These are accordingly called *partial elasticities*. For the saving function, the partial elasticities may be written as

$$\varepsilon_{SY} = \frac{\partial S/\partial Y}{S/Y} = \frac{\partial S}{\partial Y}\frac{Y}{S} \qquad \text{and} \qquad \varepsilon_{Si} = \frac{\partial S/\partial i}{S/i} = \frac{\partial S}{\partial i}\frac{i}{S}$$

For the utility function, the n partial elasticities can be concisely denoted as follows:

$$\varepsilon_{Ux_i} = \frac{\partial U}{\partial x_i}\frac{x_i}{U} \qquad (i = 1, 2, \ldots, n)$$

Example 1

Find the total differential for the following utility functions, where $a, b > 0$:

(*a*) $U(x_1, x_2) = ax_1 + bx_2$
(*b*) $U(x_1, x_2) = x_1^2 + x_2^3 + x_1x_2$
(*c*) $U(x_1, x_2) = x_1^a x_2^b$

The total differentials are as follows:

(*a*)
$$\frac{\partial U}{\partial x_1} = U_1 = a \qquad \frac{\partial U}{\partial x_2} = U_2 = b$$

and

$$dU = U_1\,dx_1 + U_2\,dx_2 = a\,dx_1 + b\,dx_2$$

(*b*)
$$\frac{\partial U}{\partial x_1} = U_1 = 2x_1 + x_2 \qquad \frac{\partial U}{\partial x_2} = U_2 = 3x_2^2 + x_1$$

and

$$dU = U_1\,dx_1 + U_2\,dx_2 = (2x_1 + x_2)\,dx_1 + \left(3x_2^2 + x_1\right)dx_2$$

(*c*)
$$\frac{\partial U}{\partial x_1} = U_1 = ax_1^{a-1}x_2^b = \frac{ax_1^a x_2^b}{x_1} \qquad \frac{\partial U}{\partial x_2} = U_2 = bx_1^a x_2^{b-1} = \frac{bx_1^a x_2^b}{x_2}$$

and

$$dU = \left(\frac{ax_1^a x_2^b}{x_1}\right)dx_1 + \left(\frac{bx_1^a x_2^b}{x_2}\right)dx_2$$

EXERCISE 8.2

1. Express the total differential dU by using the gradient vector ∇U.
2. Find the total differential, given
   (*a*) $z = 3x^2 + xy - 2y^3$
   (*b*) $U = 2x_1 + 9x_1x_2 + x_2^2$
3. Find the total differential, given
   (*a*) $y = \dfrac{x_1}{x_1 + x_2}$ (*b*) $y = \dfrac{2x_1x_2}{x_1 + x_2}$
4. The supply function of a certain commodity is
   $Q = a + bP^2 + R^{1/2}$ $(a < 0,\ b > 0)$ [*R*: rainfall]
   Find the price elasticity of supply ε_{QP}, and the rainfall elasticity of supply ε_{QR}.

5. How do the two partial elasticities in Prob. 4 vary with P and R? In a strictly monotonic fashion (assuming positive P and R)?
6. The foreign demand for our exports X depends on the foreign income Y_f and our price level P: $X = Y_f^{1/2} + P^{-2}$. Find the partial elasticity of foreign demand for our exports with respect to our price level.
7. Find the total differential for each of the following functions:
   (a) $U = -5x^3 - 12xy - 6y^5$
   (b) $U = 7x^2y^3$
   (c) $U = 3x^3(8x - 7y)$
   (d) $U = (5x^2 + 7y)(2x - 4y^3)$
   (e) $U = \dfrac{9y^3}{x - y}$
   (f) $U = (x - 3y)^3$

8.3 Rules of Differentials

A straightforward way of finding the total differential dy, given a function

$$y = f(x_1, x_2)$$

is to find the partial derivatives f_1 and f_2 and substitute these into the equation

$$dy = f_1\, dx_1 + f_2\, dx_2$$

But sometimes it may be more convenient to apply certain rules of differentials which, in view of their striking resemblance to the derivative formulas studied before, are very easy to remember.

Let k be a constant and u and v be two functions of the variables x_1 and x_2. Then the following rules are valid:†

Rule I	$dk = 0$	(cf. constant-function rule)
Rule II	$d(cu^n) = cnu^{n-1}\, du$	(cf. power-function rule)
Rule III	$d(u \pm v) = du \pm dv$	(cf. sum-difference rule)
Rule IV	$d(uv) = v\, du + u\, dv$	(cf. product rule)
Rule V	$d\left(\dfrac{u}{v}\right) = \dfrac{1}{v^2}(v\, du - u\, dv)$	(cf. quotient rule)

Instead of proving these rules here, we shall merely illustrate their practical application.

† All the rules of differentials discussed in this section are also applicable when u and v are themselves the independent variables (rather than functions of some other variables x_1 and x_2).

Example 1

Find the total differential dy of the function

$$y = 5x_1^2 + 3x_2$$

The straightfoward method calls for the evaluation of the partial derivatives $f_1 = 10x_1$ and $f_2 = 3$, which will then enable us to write

$$dy = f_1\,dx_1 + f_2\,dx_2 = 10x_1\,dx_1 + 3\,dx_2$$

We may, however, let $u = 5x_1^2$ and $v = 3x_2$ and apply the previously given rules to get the identical answer as follows:

$$\begin{aligned} dy &= d\left(5x_1^2\right) + d(3x_2) && \text{[by Rule III]} \\ &= 10x_1\,dx_1 + 3\,dx_2 && \text{[by Rule II]} \end{aligned}$$

Example 2

Find the total differential of the function

$$y = 3x_1^2 + x_1x_2^2$$

Since $f_1 = 6x_1 + x_2^2$ and $f_2 = 2x_1x_2$, the desired differential is

$$dy = \left(6x_1 + x_2^2\right) dx_1 + 2x_1x_2\,dx_2$$

By applying the given rules, the same result can be arrived at thus:

$$\begin{aligned} dy &= d\left(3x_1^2\right) + d\left(x_1x_2^2\right) && \text{[by Rule III]} \\ &= 6x_1\,dx_1 + x_2^2\,dx_1 + x_1\,d\left(x_2^2\right) && \text{[by Rules II and IV]} \\ &= \left(6x_1 + x_2^2\right) dx_1 + 2x_1x_2\,dx_2 && \text{[by Rule II]} \end{aligned}$$

Example 3

Find the total differential of the function

$$y = \frac{x_1 + x_2}{2x_1^2}$$

In view of the fact that the partial derivatives in this case are

$$f_1 = \frac{-(x_1 + 2x_2)}{2x_1^3} \qquad \text{and} \qquad f_2 = \frac{1}{2x_1^2}$$

(check these as an exercise), the desired differential is

$$dy = \frac{-(x_1 + 2x_2)}{2x_1^3}dx_1 + \frac{1}{2x_1^2}dx_2$$

However, the same result may also be obtained by application of the rules as follows:

$$\begin{aligned} dy &= \frac{1}{4x_1^4}\left[2x_1^2 d(x_1 + x_2) - (x_1 + x_2)\,d\left(2x_1^2\right)\right] && \text{[by Rule V]} \\ &= \frac{1}{4x_1^4}\left[2x_1^2(dx_1 + dx_2) - (x_1 + x_2)4x_1\,dx_1\right] && \text{[by Rules III and II]} \\ &= \frac{1}{4x_1^4}\left[-2x_1(x_1 + 2x_2)\,dx_1 + 2x_1^2\,dx_2\right] \\ &= \frac{-(x_1 + 2x_2)}{2x_1^3}dx_1 + \frac{1}{2x_1^2}dx_2 \end{aligned}$$

These rules can naturally be extended to cases where more than two functions of x_1 and x_2 are involved. In particular, we can add the following two rules to the previous collection:

Rule VI $$d(u \pm v \pm w) = du \pm dv \pm dw$$

Rule VII $$d(uvw) = vw\,du + uw\,dv + uv\,dw$$

To derive Rule VII, we can employ the familiar trick of first letting $z = vw$, so that

$$d(uvw) = d(uz) = z\,du + u\,dz \qquad \text{[by Rule IV]}$$

Then, by applying Rule IV again to dz, we get the intermediate result

$$dz = d(vw) = w\,dv + v\,dw$$

which, when substituted into the preceding equation, will yield

$$d(uvw) = vw\,du + u(w\,dv + v\,dw) = vw\,du + uw\,dv + uv\,dw$$

as the desired final result. A similar procedure can be employed to derive Rule VI.

EXERCISE 8.3

1. Use the rules of differentials to find (*a*) dz from $z = 3x^2 + xy - 2y^3$ and (*b*) dU from $U = 2x_1 + 9x_1x_2 + x_2^2$. Check your answers against those obtained for Exercise 8.2-2.
2. Use the rules of differentials to find dy from the following functions:

   (*a*) $y = \dfrac{x_1}{x_1 + x_2}$ (*b*) $y = \dfrac{2x_1x_2}{x_1 + x_2}$

   Check your answers against those obtained for Exercise 8.2-3.
3. Given $y = 3x_1(2x_2 - 1)(x_3 + 5)$
   (*a*) Find dy by Rule VII.
   (*b*) Find the differential of y, if $dx_2 = dx_3 = 0$.
4. Prove Rules II, III, IV, and V, assuming u and v to be the independent variables (rather than functions of some other variables).

8.4 Total Derivatives

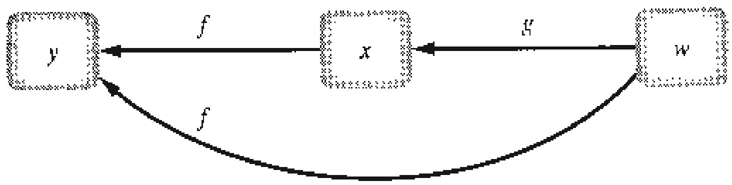
We shall now tackle the question posed at the beginning of the chapter; namely, how can we find the rate of change of the function $C(Y^*, T_0)$ with respect to T_0, when Y^* and T_0 are related? As previously mentioned, the answer lies in the concept of total derivative. Unlike a *partial* derivative, a *total* derivative does not require the argument Y^* to remain constant as T_0 varies, and can thus allow for the postulated relationship between the two arguments.

Finding the Total Derivative

To carry on the discussion in a general framework, let us consider any function

$$y = f(x, w) \qquad \text{where} \qquad x = g(w) \tag{8.11}$$

FIGURE 8.4



The two functions f and g can also be combined into a composite function

$$y = f[g(w), w] \tag{8.11'}$$

The three variables y, x, and w are related to one another as shown in Fig. 8.4. In this figure, which we shall refer to as a *channel map*, it is clearly seen that w—the ultimate source of change—can affect y through two separate channels: (1) *indirectly*, via the function g and then f (the straight arrows), and (2) *directly*, via the function f (the curved arrow). The direct effect can simply be represented by the partial derivative f_w. But the indirect effect can only be expressed by a product of two derivatives, $f_x \dfrac{dx}{dw}$, or $\dfrac{\partial y}{\partial x}\dfrac{dx}{dw}$, by the chain rule for a composite function. Adding up the two effects gives us the desired total derivative of y with respect to w:

$$\begin{aligned}\frac{dy}{dw} &= f_x \frac{dx}{dw} + f_w \\ &= \frac{\partial y}{\partial x}\frac{dx}{dw} + \frac{\partial y}{\partial w}\end{aligned} \tag{8.12}$$

This total derivative can also be obtained by an alternative method: We may first differentiate the function $y = f(x, w)$ totally, to get the total differential

$$dy = f_x\,dx + f_w\,dw$$

and then divide through by dw. The result is identical with (8.12). Either way, the process of finding the total derivative dy/dw is referred to as the *total differentiation of y with respect to w*.

It is extremely important to distinguish between the two look-alike symbols dy/dw and $\partial y/\partial w$ in (8.12). The former is a *total* derivative, and the latter, a *partial* derivative. The latter is in fact merely a component of the former.

Example 1 Find the total derivative dy/dw, given the function

$$y = f(x, w) = 3x - w^2 \qquad \text{where} \qquad x = g(w) = 2w^2 + w + 4$$

By virtue of (8.12), the total derivative should be

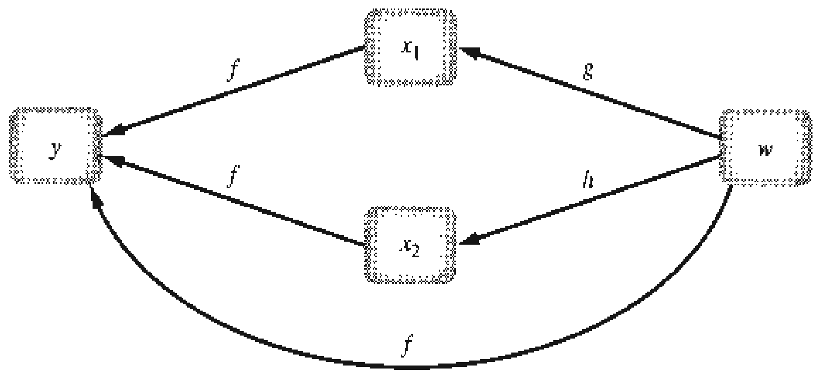
$$\frac{dy}{dw} = 3(4w + 1) + (-2w) = 10w + 3$$

As a check, we may substitute the function g into the function f, to get

$$y = 3(2w^2 + w + 4) - w^2 = 5w^2 + 3w + 12$$

which is now a function of w alone. The derivative dy/dw is then easily found to be $10w + 3$, the identical answer.

FIGURE 8.5



Example 2

If we have a utility function $U = U(c, s)$, where c is the amount of coffee consumed and s is the amount of sugar consumed, and another function $s = g(c)$ indicating the complementarity between these two goods, then we can simply write the composite function

$$U = U[c, g(c)]$$

from which it follows that

$$\frac{dU}{dc} = \frac{\partial U}{\partial c} + \frac{\partial U}{\partial g(c)} g'(c)$$

A Variation on the Theme

The situation is only slightly more complicated when we have

$$y = f(x_1, x_2, w) \qquad \text{where} \begin{cases} x_1 = g(w) \\ x_2 = h(w) \end{cases} \tag{8.13}$$

The channel map will now appear as in Fig. 8.5. This time, the variable w can affect y through three channels: (1) indirectly, via the function g and then f, (2) again indirectly, via the function h and then f, and (3) directly via f. From our previous experience, these three effects are expected to be expressible, respectively as $\frac{\partial y}{\partial x_1}\frac{dx_1}{dw}$, $\frac{\partial y}{\partial x_2}\frac{dx_2}{dw}$, and $\frac{\partial y}{\partial w}$. By adding these together, we get the total derivative

$$\begin{aligned} \frac{dy}{dw} &= \frac{\partial y}{\partial x_1}\frac{dx_1}{dw} + \frac{\partial y}{\partial x_2}\frac{dx_2}{dw} + \frac{\partial y}{\partial w} \\ &= f_1\frac{dx_1}{dw} + f_2\frac{dx_2}{dw} + f_w \end{aligned} \tag{8.14}$$

which is comparable to (8.12). If we take the total differential dy, and then divide through by dw, we can arrive at the same result.

Example 3

Let the production function be

$$Q = Q(K, L, t)$$

where, aside from the two inputs K and L, there is a third argument t, denoting time. The presence of the t argument indicates that the production function can shift over time in reflection of technological changes. Thus this is a dynamic rather than a static production function. Since capital and labor, too, can change over time, we may write

$$K = K(t) \qquad \text{and} \qquad L = L(t)$$

Then the rate of change of output with respect to time can be expressed, in line with the total-derivative formula (8.14), as

$$\frac{dQ}{dt} = \frac{\partial Q}{\partial K}\frac{dK}{dt} + \frac{\partial Q}{\partial L}\frac{dL}{dt} + \frac{\partial Q}{\partial t}$$

or, in an alternative notation,

$$\frac{dQ}{dt} = Q_K K'(t) + Q_L L'(t) + Q_t$$

Another Variation on the Theme

When the ultimate source of change, w in (8.13), is replaced by two coexisting sources, u and v, the situation becomes the following:

$$y = f(x_1, x_2, u, v) \qquad \text{where } \begin{cases} x_1 = g(u, v) \\ x_2 = h(u, v) \end{cases} \tag{8.15}$$

While the channel map will now contain more arrows, the principle of its construction remains the same; we shall, therefore, leave it to you to draw. To find the total derivative of y with respect to u (while v is held constant), let us take the total differential of y, and then divide through by the differential du, with the result:

$$\begin{aligned} \frac{dy}{du} &= \frac{\partial y}{\partial x_1}\frac{dx_1}{du} + \frac{\partial y}{\partial x_2}\frac{dx_2}{du} + \frac{\partial y}{\partial u}\frac{du}{du} + \frac{\partial y}{\partial v}\frac{dv}{du} \\ &= \frac{\partial y}{\partial x_1}\frac{dx_1}{du} + \frac{\partial y}{\partial x_2}\frac{dx_2}{du} + \frac{\partial y}{\partial u} \qquad \left[\frac{dv}{du} = 0 \text{ since } v \text{ is held constant}\right] \end{aligned}$$

In view of the fact that we are varying u while holding v constant (as a single derivative cannot handle changes in u and v both), however, the result obtained must be modified in two ways: (1) the derivatives dx_1/du and dx_2/du on the right should be rewritten with the partial sign as $\partial x_1/\partial u$ and $\partial x_2/\partial u$, which is in line with the functions g and h in (8.15); and (2) the ratio dy/du on the left should also be interpreted as a *partial* derivative, even though—being derived through the process of total differentiation of y it is actually in the nature of a *total* derivative. For this reason, we shall refer to it by the explicit name of *partial total derivative*, and denote it by $\S y/\S u$ (with $\S$ rather than ∂), in order to distinguish it from the simple partial derivative $\partial y/\partial u$ which, as our result shows, is but one of three component terms that add up to the partial total derivative.†

With these modifications, our result becomes

$$\frac{\S y}{\S u} = \frac{\partial y}{\partial x_1}\frac{\partial x_1}{\partial u} + \frac{\partial y}{\partial x_2}\frac{\partial x_2}{\partial u} + \frac{\partial y}{\partial u} \tag{8.16}$$

which is comparable to (8.14). Note the appearance of the symbol $\partial y/\partial u$ on the right, which necessitates the adoption of the new symbol $\S y/\S u$ on the left to indicate the broader

† An alternative way of denoting this partial total derivative is

$$\left.\frac{dy}{du}\right|_{v \text{ constant}} \qquad \text{or} \qquad \left.\frac{dy}{du}\right|_{dv=0}$$

concept of a partial total derivative. In a perfectly analogous manner, we can derive the other partial total derivative, $\S y/\S v$. Inasmuch as the roles of u and v are symmetrical in (8.15), however, a simpler alternative is available to us. All we have to do to obtain $\S y/\S v$ is to replace the symbol u in (8.16) by the symbol v throughout.

The use of the new symbols $\S y/\S u$ and $\S y/\S v$ for the partial total derivatives, if unconventional, serves the good purpose of avoiding confusion with the simple partial derivatives $\partial y/\partial u$ and $\partial y/\partial v$ that can arise from the function f alone in (8.15). However, in the special case where the f function takes the form of $y = f(x_1, x_2)$ without the arguments u and v, the simple partial derivatives $\partial y/\partial u$ and $\partial y/\partial v$ are not defined. Hence, it may not be inappropriate in such a case to use the latter symbols for the partial total derivatives of y with respect to u and v, since no confusion is likely to arise. Even in that event, though, the use of a special symbol is advisable for the sake of greater clarity.

Some General Remarks

To conclude this section, we offer three general remarks regarding total derivative and total differentiation:

1. In the cases we have discussed, the situation involves without exception a variable that is functionally dependent on a second variable, which is in turn dependent functionally on a third variable. As a consequence, the notion of a *chain* inevitably enters the picture, as evidenced by the appearance of a product (or products) of two derivative expressions as the component(s) of a total derivative. For this reason, the total-derivative formulas in (8.12), (8.14), and (8.16) can also be regarded as expressions of the chain rule, or the composite-function rule—a more sophisticated version of the chain rule introduced in Sec. 7.3.
2. The chain of derivatives does not have to be limited to only two "links" (two derivatives being multiplied); the concept of total derivative should be extendible to cases where there are three or more links in the composite function.
3. In all cases discussed, total derivatives—including those which have been called *partial total derivatives*—measure rates of change with respect to some *ultimate* variables in the chain or, in other words, with respect to certain variables which are in a sense *exogenous* and which are *not* expressed as functions of some other variables. The essence of the total derivative and of the process of total differentiation is to make due allowance for *all* the channels, indirect as well as direct, through which the effects of a change in an *ultimate* independent variable can possibly be carried to the particular dependent variable under study.

EXERCISE 8.4

1. Find the total derivative dz/dy, given
   (*a*) $z = f(x, y) = 5x + xy - y^2$, where $x = g(y) = 3y^2$
   (*b*) $z = 4x^2 - 3xy + 2y^2$, where $x = 1/y$
   (*c*) $z = (x + y)(x - 2y)$, where $x = 2 - 7y$
2. Find the total derivative dz/dt, given
   (*a*) $z = x^2 - 8xy - y^3$, where $x = 3t$ and $y = 1 - t$

(b) $z = 7u + vt$, where $u = 2t^2$ and $v = t + 1$

(c) $z = f(x, y, t)$, where $x = a - bt$ and $y = c + kt$

3. Find the rate of change of output with respect to time, if the production function is $Q = A(t)K^{\alpha}L^{\beta}$, where $A(t)$ is an increasing function of t, and $K = K_0 + at$, and $L = L_0 + bt$.
4. Find the partial total derivatives $§W/§u$ and $§W/§v$ if

   (a) $W = ax^2 + bxy + cu$, where $x = \alpha u + \beta v$ and $y = \gamma u$

   (b) $W = f(x_1, x_2)$, where $x_1 = 5u^2 + 3v$ and $x_2 = u - 4v^3$
5. Draw a channel map appropriate to the case of (8.15).
6. Derive the expression for $§y/§v$ formally from (8.15) by taking the total differential of y and then dividing through by dv.

8.5 Derivatives of Implicit Functions

The concept of total differentials can also enable us to find the derivatives of so-called implicit functions.

Implicit Functions

A function given in the form of $y = f(x)$, say,

$$y = f(x) = 3x^4 \tag{8.17}$$

is called an *explicit function,* because the variable y is explicitly expressed as a function of x. If this function is written alternatively in the equivalent form

$$y - 3x^4 = 0 \tag{8.17'}$$

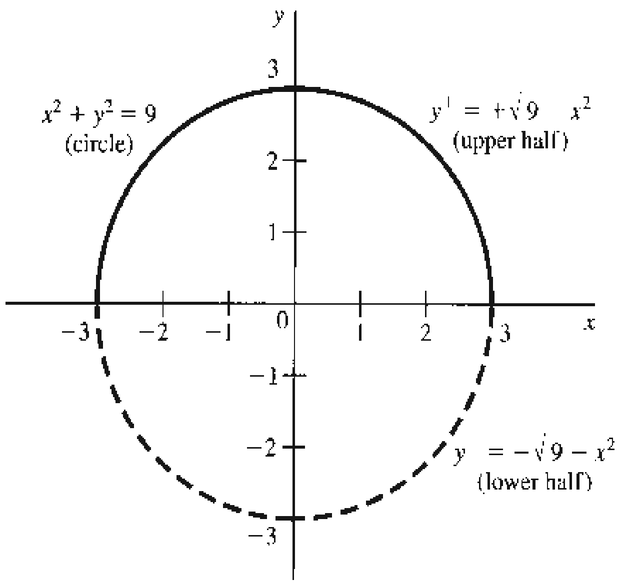
however, we no longer have an explicit function. Rather, the function (8.17) is then only *implicitly* defined by the equation (8.17′). When we are (only) given an equation in the form of (8.17′), therefore, the function $y = f(x)$ which it implies, and whose specific form may not even be known to us, is referred to as an *implicit function.*

An equation in the form of (8.17′) can be denoted in general by $F(y, x) = 0$, because its left side is a function of the two variables y and x. Note that we are using the capital letter F here to distinguish it from the function f; the function F, representing the left-side expression in (8.17′), has two arguments, y and x, whereas the function f, representing the implicit function, has only one argument, x. There may, of course, be more than two arguments in the F function. For instance, we may encounter an equation $F(y, x_1, \ldots, x_m) = 0$. Such an equation *may* also define an implicit function $y = f(x_1, \ldots, x_m)$.

The equivocal word *may* in the last sentence was used advisedly. For, whereas an explicit function, say, $y = f(x)$, can always be transformed into an equation $F(y, x) = 0$ by simply transposing the $f(x)$ expression to the left side of the equals sign, the reverse transformation is not always possible. Indeed, in certain cases, a given equation in the form of $F(y, x) = 0$ may not implicitly define a function $y = f(x)$. For instance, the equation $x^2 + y^2 = 0$ is satisfied only at the point of origin (0, 0), and hence yields no meaningful function to speak of. As another example, the equation

$$F(y, x) = x^2 + y^2 - 9 = 0 \tag{8.18}$$

FIGURE 8.6



implies not a function, but a relation, because (8.18) plots as a circle, as shown in Fig. 8.6, so that no unique value of y corresponds to each value of x. Note, however, that if we restrict y to nonnegative values, then we will have the upper half of the circle only, and that does constitute a function, namely, $y = +\sqrt{9 - x^2}$. Similarly, the lower half of the circle, with y values nonpositive, constitutes another function, $y = -\sqrt{9 - x^2}$. In contrast, neither the left half nor the right half of the circle can qualify as a function.

In view of this uncertainty, it becomes of interest to ask whether there are known general conditions under which we can be sure that a given equation in the form of

$$F(y, x_1, \ldots, x_m) = 0 \tag{8.19}$$

does indeed define an implicit function

$$y = f(x_1, \ldots, x_m) \tag{8.20}$$

locally, i.e., around some specific point in the domain. The answer to this lies in the so-called implicit-function theorem, which states that:

> Given (8.19), if (*a*) the function F has continuous partial derivatives $F_y, F_1, \ldots, F_m$, and if (*b*) at a point $(y_0, x_{10}, \ldots, x_{m0})$ satisfying the equation (8.19), F_y is nonzero, then there exists an m-dimensional neighborhood of $(x_{10}, \ldots, x_{m0})$, N, in which y is an implicitly defined function of the variables $x_1, \ldots, x_m$, in the form of (8.20). This implicit function satisfies $y_0 = f(x_{10}, \ldots, x_{m0})$. It also satisfies the equation (8.19) for *every* m-tuple $(x_1, \ldots, x_m)$ in the neighborhood N—thereby giving (8.19) the status of an *identity* in that neighborhood. Moreover, the implicit function f is continuous and has continuous partial derivatives $f_1, \ldots, f_m$.

Let us apply this theorem to the equation of the circle, (8.18), which contains only one x variable. First, we can duly verify that $F_y = 2y$ and $F_x = 2x$ are continuous, as required. Then we note that F_y is nonzero except when $y = 0$, that is, except at the leftmost point $(-3, 0)$ and the rightmost point $(3, 0)$ on the circle. Thus, around any point on the circle except $(-3, 0)$ and $(3, 0)$, we can construct a neighborhood in which the equation (8.18) defines an implicit function $y = f(x)$. This is easily verifiable in Fig. 8.6, where it is indeed possible to draw, say, a rectangle around any point on the circle—except $(-3, 0)$ and $(3, 0)$—such that the portion of the circle enclosed therein will constitute the graph of a function, with a unique y value for each value of x in that rectangle.

Several things should be noted about the implicit-function theorem. First, the conditions cited in the theorem are in the nature of sufficient (but not necessary) conditions. This means that if we happen to find $F_y = 0$ at a point satisfying (8.19), we cannot use the theorem to deny the existence of an implicit function around that point. For such a function may in fact exist (see Exercise 8.5-7).[†] Second, even if an implicit function f is assured to exist, the theorem gives no clue as to the specific form the function f takes. Nor, for that matter, does it tell us the exact size of the neighborhood N in which the implicit function is defined. However, despite these limitations, this theorem is one of great importance. For whenever the conditions of the theorem are satisfied, it now becomes meaningful to talk about and make use of a function such as (8.20), even if our model may contain an equation (8.19) which is difficult or impossible to solve explicitly for y in terms of the x variables. Moreover, since the theorem also guarantees the existence of the partial derivatives $f_1, \ldots, f_m$, it is now also meaningful to talk about these derivatives of the implicit function.

Derivatives of Implicit Functions

If the equation $F(y, x_1, \ldots, x_m) = 0$ can be solved for y, we can explicitly write out the function $y = f(x_1, \ldots, x_m)$, and find its derivatives by the methods learned before. For instance, (8.18) can be solved to yield two separate functions

$$\begin{aligned} y^+ &= +\sqrt{9 - x^2} \qquad \text{[upper half of circle]} \\ y^- &= -\sqrt{9 - x^2} \qquad \text{[lower half of circle]} \end{aligned} \tag{8.18'}$$

and their derivatives can be found as follows:

$$\begin{aligned} \frac{dy^+}{dx} &= \frac{d}{dx}(9 - x^2)^{1/2} = \tfrac{1}{2}(9 - x^2)^{-1/2}(-2x) \\ &= \frac{-x}{\sqrt{9 - x^2}} = \frac{-x}{y^+} \qquad (y^+ \neq 0) \\ \frac{dy^-}{dx} &= \frac{d}{dx}[-(9 - x^2)^{1/2}] = -\tfrac{1}{2}(9 - x^2)^{-1/2}(-2x) \\ &= \frac{x}{\sqrt{9 - x^2}} = \frac{-x}{y^-} \qquad (y^- \neq 0) \end{aligned} \tag{8.21}$$

But what if the given equation, $F(y, x_1, \ldots, x_m) = 0$, cannot be solved for y explicitly? In this case, if under the terms of the implicit-function theorem an implicit function is known to exist, we can still obtain the desired derivatives without having to solve for y first. To do this, we make use of the so-called implicit-function rule—a rule that can give us the derivatives of *every* implicit function defined by the given equation. The development of this rule depends on the following basic facts: (1) if two expressions are *identically*

[†] On the other hand, if $F_y = 0$ in an entire neighborhood, then it can be concluded that no implicit function is defined in that neighborhood. By the same token if $F_y = 0$ identically, then no implicit function exists anywhere.

equal, their respective total differentials must be equal;† (2) differentiation of an expression that involves $y, x_1, \ldots, x_m$ will yield an expression involving the differentials $dy, dx_1, \ldots, dx_m$; and (3) the differential of y, dy, can be substituted out, so the fact that we cannot solve for y does not matter.

Applying these facts to the equation $F(y, x_1, \ldots, x_m) = 0$—which, we recall, has the status of an *identity* in the neighborhood N in which the implicit function is defined—we can write $dF = d0$, or

$$F_y\, dy + F_1\, dx_1 + F_2\, dx_2 + \cdots + F_m\, dx_m = 0 \qquad \textbf{(8.22)}$$

Since the implicit function $y = f(x_1, x_2, \ldots, x_m)$ has the total differential

$$dy = f_1\, dx_1 + f_2\, dx_2 + \cdots + f_n\, dx_n$$

we can substitute this dy expression into (8.22) to get (after collecting terms)

$$(F_y\, f_1 + F_1)\, dx_1 + (F_y\, f_2 + F_2)\, dx_2 + \cdots + (F_y\, f_m + F_m)\, dx_m = 0 \qquad \textbf{(8.22')}$$

The fact that all the dx_i can vary independently from one another means that, for the equation (8.22') to hold, each parenthesized expression must individually vanish; i.e., we must have

$$F_y\, f_i + F_i = 0 \qquad \text{(for all } i\text{)}$$

Dividing through by F_y, and solving for f_i, we obtain the so-called implicit-function rule for finding the partial derivative f_i of the implicit function $y = f(x_1, x_2, \ldots, x_m)$:

$$f_i \equiv \frac{\partial y}{\partial x_i} = -\frac{F_i}{F_y} \qquad (i = 1, 2, \ldots, m) \qquad \textbf{(8.23)}$$

In the simple case where the given equation is $F(y, x) = 0$, the rule gives

$$\frac{dy}{dx} = -\frac{F_x}{F_y} \qquad \textbf{(8.23')}$$

† Take, for example, the identity

$$x^2 - y^2 \equiv (x + y)(x - y)$$

This is an identity because the two sides are equal for *any* values of x and y that one may assign. Taking the total differential of each side, we have

$$\begin{aligned} d(\text{left side}) &= 2x\, dx - 2y\, dy \\ d(\text{right side}) &= (x - y)\, d(x + y) + (x + y)\, d(x - y) \\ &= (x - y)(dx + dy) + (x + y)(dx - dy) \\ &= 2x\, dx - 2y\, dy \end{aligned}$$

The two results are indeed equal. If two expressions are *not* identically equal, but are equal only for certain specific values of the variables, however, their total differentials will *not* be equal. The equation

$$x^2 - y^2 = x^2 + y^2 - 2$$

for instance, is valid only for $y = \pm 1$. The total differentials of the two sides are

$$\begin{aligned} d(\text{left side}) &= 2x\, dx - 2y\, dy \\ d(\text{right side}) &= 2x\, dx + 2y\, dy \end{aligned}$$

which are not equal. Note, in particular, that they are not equal even at $y = \pm 1$.

What this rule states is that, even if the specific form of the implicit function is not known to us, we can nevertheless find its derivative(s) by taking the *negative* of the ratio of a pair of partial derivatives of the F function which appears in the given equation that defines the implicit function. Observe that F_y always appears in the denominator of the ratio. This being the case, it is not admissible to have $F_y = 0$. Since the implicit-function theorem specifies that $F_y \neq 0$ at the point around which the implicit function is defined, the problem of a zero denominator is automatically taken care of in the relevant neighborhood of that point.

Example 1

Find dy/dx for the implicit function defined by (8.17′). Since $F(y, x)$ takes the form of $y - 3x^4$, we have, by (8.23′),

$$\frac{dy}{dx} = -\frac{F_x}{F_y} = -\frac{-12x^3}{1} = 12x^3$$

In this particular case, we can easily solve the given equation for y to get $y = 3x^4$. Thus the correctness of the derivative is easily verified.

Example 2

Find dy/dx for the implicit functions defined by the equation of the circle (8.18). This time we have $F(y, x) = x^2 + y^2 - 9$; thus $F_y = 2y$ and $F_x = 2x$. By (8.23′), the desired derivative is

$$\frac{dy}{dx} = -\frac{2x}{2y} = -\frac{x}{y} \qquad (y \neq 0)^\dagger$$

Earlier, it was asserted that the implicit-function rule gives us the derivative of *every* implicit function defined by a given equation. Let us verify this with the two functions in (8.18′) and their derivatives in (8.21). If we substitute y^+ for y in the implicit-function-rule result $dy/dx = -x/y$, we will indeed obtain the derivative dy^+/dx as shown in (8.21); similarly, the substitution of y^- for y will yield the other derivative in (8.21). Thus our earlier assertion is duly verified.

Example 3

Find $\partial y/\partial x$ for any implicit function(s) that may be defined by the equation $F(y, x, w) = y^3x^2 + w^3 + yxw - 3 = 0$. This equation is not easily solved for y. But since F_y, F_x, and F_w are all obviously continuous, and since $F_y = 3y^2x^2 + xw$ is indeed nonzero at a point such as (1, 1, 1) which satisfies the given equation, an implicit function $y = f(x, w)$ assuredly exists around that point at least. It is thus meaningful to talk about the derivative $\partial y/\partial x$. By (8.23), moreover, we can immediately write

$$\frac{\partial y}{\partial x} = -\frac{F_x}{F_y} = -\frac{2y^3x + yw}{3y^2x^2 + xw}$$

At the point (1, 1, 1), this derivative has the value $-\frac{3}{4}$.

Example 4

Assume that the equation $F(Q, K, L) = 0$ implicitly defines a production function $Q = f(K, L)$. Let us find a way of expressing the marginal physical products MPP_K and MPP_L in relation to the function F. Since the marginal products are simply the partial derivatives $\partial Q/\partial K$ and $\partial Q/\partial L$, we can apply the implicit-function rule and write

$$\text{MPP}_K \equiv \frac{\partial Q}{\partial K} = -\frac{F_K}{F_Q} \qquad \text{and} \qquad \text{MPP}_L \equiv \frac{\partial Q}{\partial L} = -\frac{F_L}{F_Q}$$

† The restriction $y \neq 0$ is of course perfectly consistent with our earlier discussion of the equation (8.18) that follows the statement of the implicit-function theorem.

Aside from these, we can obtain yet another partial derivative,

$$\frac{\partial K}{\partial L} = -\frac{F_L}{F_K}$$

from the equation $F(Q, K, L) = 0$. What is the economic meaning of $\partial K/\partial L$? The partial sign implies that the other variable, Q, is being held constant; it follows that the changes in K and L described by this derivative are in the nature of "compensatory" changes designed to keep the output Q constant at a specified level. These are therefore the type of changes pertaining to movements *along* a production *isoquant* drawn with the K variable on the vertical axis and the L variable on the horizontal axis. As a matter of fact, the derivative $\partial K/\partial L$ is the measure of the slope of such an isoquant, which is negative in the normal case. The absolute value of $\partial K/\partial L$, on the other hand, is the measure of the *marginal rate of technical substitution* between the two inputs.

Extension to the Simultaneous-Equation Case

The implicit-function theorem also comes in a more general and powerful version that deals with the conditions under which a set of simultaneous equations

$$\begin{aligned} F^1(y_1, \ldots, y_n; x_1, \ldots, x_m) &= 0 \\ F^2(y_1, \ldots, y_n; x_1, \ldots, x_m) &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ F^n(y_1, \ldots, y_n; x_1, \ldots, x_m) &= 0 \end{aligned} \tag{8.24}$$

will assuredly define a set of implicit functions†

$$\begin{aligned} y_1 &= f^1(x_1, \ldots, x_m) \\ y_2 &= f^2(x_1, \ldots, x_m) \\ \cdots\cdots\cdots\cdots\cdots \\ y_n &= f^n(x_1, \ldots, x_m) \end{aligned} \tag{8.25}$$

The generalized version of the theorem states that:

> Given the equation system (8.24), if (*a*) the functions $F^1, \ldots, F^n$ all have continuous partial derivatives with respect to all the y and x variables, and if (*b*) at a point $(y_{10}, \ldots, y_{n0}; x_{10}, \ldots, x_{m0})$ satisfying (8.24), the following Jacobian determinant is nonzero:

$$|J| \equiv \left|\frac{\partial(F^1, \ldots, F^n)}{\partial(y_1, \ldots, y_n)}\right| \equiv \begin{vmatrix} \frac{\partial F^1}{\partial y_1} & \frac{\partial F^1}{\partial y_2} & \cdots & \frac{\partial F^1}{\partial y_n} \\ \frac{\partial F^2}{\partial y_1} & \frac{\partial F^2}{\partial y_2} & \cdots & \frac{\partial F^2}{\partial y_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial F^n}{\partial y_1} & \frac{\partial F^n}{\partial y_2} & \cdots & \frac{\partial F^n}{\partial y_n} \end{vmatrix} \neq 0$$

† To view it another way, what these conditions serve to do is to assure us that the n equations in (8.24) can *in principle* be solved for the n variables—$y_1, \ldots, y_n$—even if we may not be able to obtain the solution (8.25) in an explicit form.

then there exists an m-dimensional neighborhood of $(x_{10}, \ldots, x_{m0})$, N, in which the variables $y_1, \ldots, y_n$ are functions of the variables $x_1, \ldots, x_m$ in the form of (8.25). These implicit functions satisfy

$$
\begin{aligned}
y_{10} &= f^1(x_{10}, \ldots, x_{m0}) \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
y_{n0} &= f^n(x_{10}, \ldots, x_{m0})
\end{aligned}
$$

They also satisfy (8.24) for *every* m-tuple $(x_1, \ldots, x_m)$ in the neighborhood N thereby giving (8.24) the status of a set of *identities* as far as this neighborhood is concerned. Moreover, the implicit functions $f^1, \ldots, f^n$ are continuous and have continuous partial derivatives with respect to all the x variables.

As in the single-equation case, it is possible to find the partial derivatives of the implicit functions directly from the n equations in (8.24), without having to solve them for the y variables. Taking advantage of the fact that, in the neighborhood N, the equations in (8.24) have the status of identities, we can take the total differential of each of these, and write $dF^j = 0$ $(j = 1, 2, \ldots, n)$. The result is a set of equations involving the differentials $dy_1, \ldots, dy_n$ and $dx_1, \ldots, dx_m$. Specifically, after transposing the dx_i terms to the right of the equals signs, we have

$$
\begin{aligned}
\frac{\partial F^1}{\partial y_1}dy_1 + \frac{\partial F^1}{\partial y_2}dy_2 + \cdots + \frac{\partial F^1}{\partial y_n}dy_n &= -\left(\frac{\partial F^1}{\partial x_1}dx_1 + \cdots + \frac{\partial F^1}{\partial x_m}dx_m\right) \\
\frac{\partial F^2}{\partial y_1}dy_1 + \frac{\partial F^2}{\partial y_2}dy_2 + \cdots + \frac{\partial F^2}{\partial y_n}dy_n &= -\left(\frac{\partial F^2}{\partial x_1}dx_1 + \cdots + \frac{\partial F^2}{\partial x_m}dx_m\right) \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\frac{\partial F^n}{\partial y_1}dy_1 + \frac{\partial F^n}{\partial y_2}dy_2 + \cdots + \frac{\partial F^n}{\partial y_n}dy_n &= -\left(\frac{\partial F^n}{\partial x_1}dx_1 + \cdots + \frac{\partial F^n}{\partial x_m}dx_m\right)
\end{aligned}
\tag{8.26}
$$

Moreover, from (8.25), we can write the differentials of the y_j variables as

$$
\begin{aligned}
dy_1 &= \frac{\partial y_1}{\partial x_1}dx_1 + \frac{\partial y_1}{\partial x_2}dx_2 + \cdots + \frac{\partial y_1}{\partial x_m}dx_m \\
dy_2 &= \frac{\partial y_2}{\partial x_1}dx_1 + \frac{\partial y_2}{\partial x_2}dx_2 + \cdots + \frac{\partial y_2}{\partial x_m}dx_m \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
dy_n &= \frac{\partial y_n}{\partial x_1}dx_1 + \frac{\partial y_n}{\partial x_2}dx_2 + \cdots + \frac{\partial y_n}{\partial x_m}dx_m
\end{aligned}
\tag{8.27}
$$

and these can be used to eliminate the dy_j expressions in (8.26). But since the result of substitution would be unmanageably messy, let us simplify matters by considering only what would happen when x_1 alone changes while all the other variables $x_2, \ldots, x_m$ remain constant. Letting $dx_1 \neq 0$, but setting $dx_2 = \cdots = dx_m = 0$ in (8.26) and (8.27), then substituting (8.27) into (8.26) and dividing through by $dx_1 \neq 0$, we obtain the equation system

$$\begin{aligned}
\frac{\partial F^1}{\partial y_1}\left(\frac{\partial y_1}{\partial x_1}\right) + \frac{\partial F^1}{\partial y_2}\left(\frac{\partial y_2}{\partial x_1}\right) + \cdots + \frac{\partial F^1}{\partial y_n}\left(\frac{\partial y_n}{\partial x_1}\right) &= -\frac{\partial F^1}{\partial x_1} \\
\frac{\partial F^2}{\partial y_1}\left(\frac{\partial y_1}{\partial x_1}\right) + \frac{\partial F^2}{\partial y_2}\left(\frac{\partial y_2}{\partial x_1}\right) + \cdots + \frac{\partial F^2}{\partial y_n}\left(\frac{\partial y_n}{\partial x_1}\right) &= -\frac{\partial F^2}{\partial x_1} \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\frac{\partial F^n}{\partial y_1}\left(\frac{\partial y_1}{\partial x_1}\right) + \frac{\partial F^n}{\partial y_2}\left(\frac{\partial y_2}{\partial x_1}\right) + \cdots + \frac{\partial F^n}{\partial y_n}\left(\frac{\partial y_n}{\partial x_1}\right) &= -\frac{\partial F^n}{\partial x_1}
\end{aligned} \qquad \textbf{(8.28)}$$

Even this result—for the case where x_1 alone changes—looks formidably complex, because it is full of derivatives. But its structure is actually quite easy to comprehend, once we learn to distinguish between the two types of derivatives that appear in (8.28). One type, which we have parenthesized for visual distinction, consists of the partial derivatives of the implicit functions with respect to x_1 that we are seeking. These, therefore, should be viewed as the "variables" to be solved for in (8.28). The other type, on the other hand, consists of the partial derivatives of the F^j functions given in (8.24). Since they would all take specific values when evaluated at the point $(y_{10}, \ldots, y_{n0}; x_{10}, \ldots, x_{m0})$—the point around which the implicit functions are defined—they appear here not as derivative functions but as derivative values. As such, they can be treated as given constants. This fact makes (8.28) a *linear equation system,* with a structure similar to (4.1). What is interesting is that such a linear system has arisen during the process of analysis of a problem that is not necessarily linear in itself, since no linearity restrictions have been placed on the equation system (8.24). Thus we have here an illustration of how linear algebra can come into play even in nonlinear problems.

Being a linear equation system, (8.28) can be written in matrix notation as

$$\begin{bmatrix}
\frac{\partial F^1}{\partial y_1} & \frac{\partial F^1}{\partial y_2} & \cdots & \frac{\partial F^1}{\partial y_n} \\
\frac{\partial F^2}{\partial y_1} & \frac{\partial F^2}{\partial y_2} & \cdots & \frac{\partial F^2}{\partial y_n} \\
\cdots & \cdots & \cdots & \cdots \\
\frac{\partial F^n}{\partial y_1} & \frac{\partial F^n}{\partial y_2} & \cdots & \frac{\partial F^n}{\partial y_n}
\end{bmatrix}
\begin{bmatrix}
\left(\frac{\partial y_1}{\partial x_1}\right) \\
\left(\frac{\partial y_2}{\partial x_1}\right) \\
\vdots \\
\left(\frac{\partial y_n}{\partial x_1}\right)
\end{bmatrix}
=
\begin{bmatrix}
-\frac{\partial F^1}{\partial x_1} \\
-\frac{\partial F^2}{\partial x_1} \\
\vdots \\
-\frac{\partial F^n}{\partial x_1}
\end{bmatrix} \qquad \textbf{(8.28')}$$

Since the determinant of the coefficient matrix in (8.28′) is nothing but the particular Jacobian determinant $|J|$ which is known to be nonzero under conditions of the implicit-function theorem, and since the system must be nonhomogeneous (why?), there should be a unique nontrivial solution to (8.28′). By Cramer's rule, this solution may be expressed analytically as follows:

$$\left(\frac{\partial y_j}{\partial x_1}\right) = \frac{|J_j|}{|J|} \qquad (j = 1, 2, \ldots, n) \qquad \text{[see (5.18)]} \qquad \textbf{(8.29)}$$

By a suitable adaptation of this procedure, the partial derivatives of the implicit functions with respect to the other variables, $x_2, \ldots, x_m$, can also be obtained. It is a nice feature of this procedure that, each time we allow a particular x_i variable to change, we can obtain in one fell swoop the partial derivatives of all the implicit functions $f^1, \ldots, f^n$ with respect to that particular x_i variable.

Similarly, to the implicit-function rule (8.23) for the single-equation case, the procedure just described calls only for the use of the partial derivatives of the F functions—evaluated at the point $(y_{10}, \ldots, y_{n0}; x_{10}, \ldots, x_{m0})$—in the calculation of the partial derivatives of the implicit functions $f^1, \ldots, f^n$. Thus the matrix equation (8.28′) and its analytical solution (8.29) are in effect a statement of the simultaneous-equation version of the implicit-function rule.

Note that the requirement $|J| \neq 0$ rules out a zero denominator in (8.29), just as the requirement $F_y \neq 0$ did in the implicit-function rule (8.23) and (8.23′). Also, the role played by the condition $|J| \neq 0$ in guaranteeing a unique (albeit implicit) solution (8.25) to the general (possibly *nonlinear*) system (8.24) is very similar to the role of the nonsingularity condition $|A| \neq 0$ in a *linear* system $Ax = d$.

Example 5

The following three equations

$$\begin{aligned} xy - w &= 0 \qquad & F^1 &= (x, y, w; z) = 0 \\ y - w^3 - 3z &= 0 \qquad & F^2 &= (x, y, w; z) = 0 \\ w^3 + z^3 - 2zw &= 0 \qquad & F^3 &= (x, y, w; z) = 0 \end{aligned}$$

are satisfied at point P: $(x, y, w; z) = (\frac{1}{4}, 4, 1, 1)$. The F^i functions obviously possess continuous derivatives. Thus, if the Jacobian $|J|$ is nonzero at point P, we can use the implicit-function theorem to find the comparative-static derivative $(\partial x/\partial z)$.

To do this, we can first take the total differential of the system:

$$\begin{aligned} y\,dx + x\,dy - dw &= 0 \\ dy - 3w^2\,dw - 3\,dz &= 0 \\ (3w^2 - 2z)\,dw + (3z^2 - 2w)\,dz &= 0 \end{aligned}$$

Moving the exogenous differential (and its coefficients) to the right-hand side and writing in matrix form, we get

$$\begin{bmatrix} y & x & -1 \\ 0 & 1 & -3w^2 \\ 0 & 0 & (3w^2 - 2z) \end{bmatrix} \begin{bmatrix} dx \\ dy \\ dw \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 2w - 3z^2 \end{bmatrix} dz$$

where the coefficient matrix on the left-hand side is the Jacobian

$$|J| = \begin{vmatrix} F_x^1 & F_y^1 & F_w^1 \\ F_x^2 & F_y^2 & F_w^2 \\ F_x^3 & F_y^3 & F_w^3 \end{vmatrix} = \begin{vmatrix} y & x & -1 \\ 0 & 1 & -3w^2 \\ 0 & 0 & (3w^2 - 2z) \end{vmatrix} = y(3w^2 - 2z)$$

At the point P, the Jacobian determinant $|J| = 4\ (\neq 0)$. Therefore, the implicit-function rule applies and

$$\begin{bmatrix} y & x & -1 \\ 0 & 1 & -3w^2 \\ 0 & 0 & (3w^2 - 2z) \end{bmatrix} \begin{bmatrix} \left(\dfrac{\partial x}{\partial z}\right) \\ \left(\dfrac{\partial y}{\partial z}\right) \\ \left(\dfrac{\partial w}{\partial z}\right) \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 2w - 3z^2 \end{bmatrix}$$

Using Cramer's rule to find an expression for $(\partial x/\partial z)$, we obtain

$$\left(\frac{\partial x}{\partial z}\right) = \frac{\begin{vmatrix} 0 & x & -1 \\ 3 & 1 & -3w^2 \\ 2w-3z^2 & 0 & (3w^2-2z) \end{vmatrix}}{|J|} = \frac{\begin{vmatrix} 0 & \frac{1}{4} & -1 \\ 3 & 1 & -3 \\ -1 & 0 & 1 \end{vmatrix}}{4}$$

$$= 0 + (-3)\frac{\begin{vmatrix} \frac{1}{4} & -1 \\ 0 & 1 \end{vmatrix}}{4} + (-1)\frac{\begin{vmatrix} \frac{1}{4} & -1 \\ 1 & -3 \end{vmatrix}}{4}$$

$$= \frac{-3}{16} + \frac{-1}{16}$$

$$= -\frac{1}{4}$$

Example 6 Let the national-income model (7.17) be rewritten in the form

$$\begin{aligned} Y - C - I_0 - G_0 &= 0 \\ C - \alpha - \beta(Y - T) &= 0 \\ T - \gamma - \delta Y &= 0 \end{aligned} \tag{8.30}$$

If we take the endogenous variables (Y, C, T) to be (y_1, y_2, y_3), and take the exogenous variables and parameters $(I_0, G_0, \alpha, \beta, \gamma, \delta)$ to be $(x_1, x_2, \ldots, x_6)$, then the left-side expression in each equation can be regarded as a specific F function, in the form of $F^j(Y, C, T; I_0, G_0, \alpha, \beta, \gamma, \delta)$. Thus (8.30) is a specific case of (8.24), with $n = 3$ and $m = 6$. Since the functions F^1, F^2, and F^3 do have continuous partial derivatives, and since the relevant Jacobian determinant (the one involving only the endogenous variables),

$$|J| = \begin{vmatrix} \frac{\partial F^1}{\partial Y} & \frac{\partial F^1}{\partial C} & \frac{\partial F^1}{\partial T} \\ \frac{\partial F^2}{\partial Y} & \frac{\partial F^2}{\partial C} & \frac{\partial F^2}{\partial T} \\ \frac{\partial F^3}{\partial Y} & \frac{\partial F^3}{\partial C} & \frac{\partial F^3}{\partial T} \end{vmatrix} = \begin{vmatrix} 1 & -1 & 0 \\ -\beta & 1 & \beta \\ -\delta & 0 & 1 \end{vmatrix} = 1 - \beta + \beta\delta \tag{8.31}$$

is always nonzero (both β and δ being restricted to be positive fractions), we can take Y, C, and T to be implicit functions of $(I_0, G_0, \alpha, \beta, \gamma, \delta)$ *at* and *around* any point that satisfies (8.30). But a point that satisfies (8.30) would be an equilibrium solution, relating to Y^*, C^* and T^*. Hence, what the implicit-function theorem tells us is that we are justified in writing

$$\begin{aligned} Y^* &= f^1(I_0, G_0, \alpha, \beta, \gamma, \delta) \\ C^* &= f^2(I_0, G_0, \alpha, \beta, \gamma, \delta) \\ T^* &= f^3(I_0, G_0, \alpha, \beta, \gamma, \delta) \end{aligned}$$

indicating that the equilibrium values of the endogenous variables are implicit functions of the exogenous variables and the parameters.

The partial derivatives of the implicit functions, such as $\partial Y^*/\partial I_0$ and $\partial Y^*/\partial G_0$, are in the nature of comparative-static derivatives. To find these, we need only the partial derivatives of the F functions, evaluated at the equilibrium state of the model. Moreover, since $n = 3$, three of these can be found in one operation. Suppose we now hold all exogenous variables and parameters fixed except G_0. Then, by adapting the result in (8.28′), we may write the equation

$$\begin{bmatrix} 1 & -1 & 0 \\ -\beta & 1 & \beta \\ -\delta & 0 & 1 \end{bmatrix} \begin{bmatrix} \partial Y^*/\partial G_0 \\ \partial C^*/\partial G_0 \\ \partial T^*/\partial G_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

from which three comparative-static derivatives (all with respect to G_0) can be calculated. The first one, representing the government-expenditure multiplier, will for instance come out to be

$$\frac{\partial Y^*}{\partial G_0} = \frac{\begin{vmatrix} 1 & -1 & 0 \\ 0 & 1 & \beta \\ 0 & 0 & 1 \end{vmatrix}}{|J|} = \frac{1}{1-\beta+\beta\delta} \qquad \text{[by (8.31)]}$$

This is, of course, nothing but the result obtained earlier in (7.19). Note, however, that in the present approach we have worked only with implicit functions, and have completely bypassed the step of solving the system (8.30) explicitly for Y^*, C^*, and T^*. It is this particular feature of the method that will now enable us to tackle the comparative statics of general-function models which, by their very nature, can yield no explicit solution.

EXERCISE 8.5

1. For each $F(x, y) = 0$, find dy/dx for each of the following:
   (*a*) $y - 6x + 7 = 0$
   (*b*) $3y + 12x + 17 = 0$
   (*c*) $x^2 + 6x - 13 - y = 0$
2. For each $F(x, y) = 0$ use the implicit-function rule to find dy/dx:
   (*a*) $F(x, y) = 3x^2 + 2xy + 4y^3 = 0$
   (*b*) $F(x, y) = 12x^5 - 2y = 0$
   (*c*) $F(x, y) = 7x^2 + 2xy^2 + 9y^4 = 0$
   (*d*) $F(x, y) = 6x^3 - 3y = 0$
3. For each $F(x, y, z) = 0$ use the implicit-function rule to find $\partial y/\partial x$ and $\partial y/\partial z$:
   (*a*) $F(x, y, z) = x^2y^3 + z^2 + xyz = 0$
   (*b*) $F(x, y, z) = x^3z^2 + y^3 + 4xyz = 0$
   (*c*) $F(x, y, z) = 3x^2y^3 + xz^2y^2 + y^3zx^4 + y^2z = 0$
4. Assuming that the equation $F(U, x_1, x_2, \ldots, x_n) = 0$ implicitly defines a utility function $U = f(x_1, x_2, \ldots, x_n)$:
   (*a*) Find the expressions for $\partial U/\partial x_2$, $\partial U/\partial x_n$, $\partial x_3/\partial x_2$, and $\partial x_4/\partial x_n$.
   (*b*) Interpret their respective economic meanings.
5. For each of the given equations $F(y, x) = 0$, is an implicit function $y = f(x)$ defined around the point $(y = 3, x = 1)$?
   (*a*) $x^3 - 2x^2y + 3xy^2 - 22 = 0$
   (*b*) $2x^2 + 4xy - y^4 + 67 = 0$