CHAPTER 6

COMPARATIVE STATICS: THE TRADITIONAL METHODOLOGY

6.1 INTRODUCTION; PROFIT MAXIMIZATION ONCE MORE

In this chapter we shall begin the general comparative statics analysis of economic models that contain an explicit maximization hypothesis. The focus, as always, will be on discovering the structure that must be imposed on the models so that useful, i.e., refutable, hypotheses are implied. A very powerful methodology, *duality theory,* has been developed for some important models such as profit maximization, constrained cost minimization, and utility maximization subject to a budget constraint. These new methods provide a vast simplification and clarification of the traditional methodology for those models; we shall explore them in the next chapter. In order to analyze models other than the three just mentioned, however, and to better appreciate the newer methods, it is still necessary to understand the traditional methodology. It is to that task that we now turn.

Comparative statics of economic models involving more than one variable requires the solution to simultaneous linear equations in the partial derivatives of the choice variables with respect to the parameters. We shall employ elementary matrix manipulations and Cramer's rule in order to systematically write down the solutions to the first-order equations. In that way, the structure of these models can be most efficiently explored.

Consider again the profit-maximizing firm analyzed in Chap. 4, and recall Eqs. (4-19):

$$\begin{aligned} pf_{11}\frac{\partial x_1^*}{\partial w_1} + pf_{12}\frac{\partial x_2^*}{\partial w_1} &\equiv 1 \\ pf_{21}\frac{\partial x_1^*}{\partial w_1} + pf_{22}\frac{\partial x_2^*}{\partial w_1} &\equiv 0 \end{aligned} \tag{4-19}$$

In matrix form these equations appear as

$$\begin{pmatrix} pf_{11} & pf_{12} \\ pf_{21} & pf_{22} \end{pmatrix} \begin{pmatrix} \dfrac{\partial x_1^*}{\partial w_1} \\ \dfrac{\partial x_2^*}{\partial w_1} \end{pmatrix} \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{6-1}$$

Using Cramer's rule,

$$\frac{\partial x_1^*}{\partial w_1} = \frac{\begin{vmatrix} 1 & pf_{12} \\ 0 & pf_{22} \end{vmatrix}}{H} = \frac{pf_{22}}{H} \qquad \text{where } H = \begin{vmatrix} pf_{11} & pf_{12} \\ pf_{21} & pf_{22} \end{vmatrix} \tag{6-2}$$

This is Eq. (4-20a), which was derived by algebraic manipulations. Notice that the term 1 on the right-hand side of (6-1) will always appear in column i, in the solution for $\partial x_i^*/\partial w_j$. If the numerator is expanded by that column, it is immediately apparent that Eqs. (4-20a–d) can be written as

$$\frac{\partial x_i^*}{\partial w_j} = \frac{H_{ji}}{H} \qquad i, j = 1, 2 \tag{6-3}$$

where H_{ji} is the cofactor (signed, of course) of the element in the jth row and ith column. In this model, $H_{11} = pf_{22}$, $H_{22} = pf_{11}$, $H_{12} = H_{21} = -pf_{12}$. Notice, too, that $H = p^2(f_{11}f_{22} - f_{12}^2)$, and that $H > 0$, from the second-order conditions (4-15). This is in fact indicative of a trend; determinants will play a crucial role in the theory of maxima and minima.

In like fashion, Eqs. (4-21), dealing with changes in the factor utilizations due to output price changes, can be written

$$(pf_{ij}) \begin{pmatrix} \dfrac{\partial x_1^*}{\partial p} \\ \dfrac{\partial x_2^*}{\partial p} \end{pmatrix} \equiv \begin{pmatrix} -f_1 \\ -f_2 \end{pmatrix} \tag{6-4}$$

where the expression (pf_{ij}) stands for the 2×2 matrix in the left-hand side of (6-1). It is obvious from Cramer's rule that the solutions for $\partial x_1^*/\partial p$ and $\partial x_2^*/\partial p$ will involve the "off-diagonal" terms of pf_{12} and pf_{21}. Since the sign of these (equal) terms is not implied by maximization, we immediately suspect that no sign will emerge for $\partial x_1^*/\partial p$, etc., and hence no refutable hypotheses concerning the responses of inputs to output price changes will emerge.

The two-factor, profit-maximizing firm is an example of a maximization model with two choice variables. The most general form of such models is[†]

maximize

$$f(x_1, x_2, \alpha) \tag{6-5}$$

where the choice variables are x_1 and x_2 and α is a parameter, or perhaps a vector of parameters, $\alpha = (\alpha_1, \ldots, \alpha_m)$. The first-order necessary conditions implied by (6-5), usually called the *equilibrium conditions,* are

$$\begin{aligned} f_1(x_1, x_2, \alpha) &= 0 \\ f_2(x_1, x_2, \alpha) &= 0 \end{aligned} \tag{6-6}$$

The sufficient second-order conditions are

$$f_{11} < 0 \quad f_{22} < 0 \quad f_{11}f_{22} - f_{12}^2 > 0 \tag{6-7}$$

Equations (6-6) are two equations in three variables, x_1, x_2, and α. The sufficient second-order conditions imply, by the implicit function theorem, that these equations can be solved for the explicit choice functions

$$\begin{aligned} x_1 &= x_1^*(\alpha) \\ x_2 &= x_2^*(\alpha) \end{aligned} \tag{6-8}$$

It should always be remembered that Eqs. (6-8) are the *simultaneous* solutions of (6-6). As the parameter α changes, *both* x_1 and x_2 will in general change. Substituting (6-8) back into (6-6), the identities from which the comparative statics are derivable are obtained:

$$\begin{aligned} f_1(x_1^*(\alpha), x_2^*(\alpha), \alpha) &\equiv 0 \\ f_2(x_1^*(\alpha), x_2^*(\alpha), \alpha) &\equiv 0 \end{aligned} \tag{6-9}$$

Differentiating this system with respect to α, the following system is obtained.

$$\begin{aligned} f_{11}\frac{\partial x_1^*}{\partial \alpha} + f_{12}\frac{\partial x_2^*}{\partial \alpha} + f_{1\alpha} &\equiv 0 \\ f_{21}\frac{\partial x_1^*}{\partial \alpha} + f_{22}\frac{\partial x_2^*}{\partial \alpha} + f_{2\alpha} &\equiv 0 \end{aligned} \tag{6-10}$$

In matrix form, this system is

$$\begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} \begin{pmatrix} \dfrac{\partial x_1^*}{\partial \alpha} \\ \dfrac{\partial x_2^*}{\partial \alpha} \end{pmatrix} \equiv \begin{pmatrix} -f_{1\alpha} \\ -f_{2\alpha} \end{pmatrix} \tag{6-11}$$

[†]The function f here refers to the whole maximand, not just the production function part of the previous objective function.

Solving by Cramer's rule,

$$\frac{\partial x_1^*}{\partial \alpha} = \frac{\begin{vmatrix} -f_{1\alpha} & f_{12} \\ -f_{2\alpha} & f_{22} \end{vmatrix}}{H} = \frac{-f_{1\alpha} f_{22}}{H} + \frac{f_{2\alpha} f_{12}}{H} \tag{6-12a}$$

and

$$\frac{\partial x_2^*}{\partial \alpha} = \frac{\begin{vmatrix} f_{11} & -f_{1\alpha} \\ f_{21} & -f_{2\alpha} \end{vmatrix}}{H} = \frac{-f_{2\alpha} f_{11}}{H} + \frac{f_{1\alpha} f_{21}}{H} \tag{6-12b}$$

where H is the determinant

$$H = \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix} = f_{11} f_{22} - f_{12}^2 > 0 \tag{6-13}$$

Equations (6-12) represent the most general comparative statics relations for unconstrained maximization models with two choice variables. Not surprisingly at this level of generality, no refutable hypotheses are implied. Certain information is available, though. The denominators H in Eqs. (6-12) are positive. In addition, f_{11}, f_{22} are negative. This information is provided by the sufficient conditions for a maximum.

The other information that is available is provided by the actual structure of the model. Specifically, to be useful, a model must be constructed so that the effects of the parameters on the objective function, and hence the first-order equations, will in general be known. That is, $f_{1\alpha}$ and $f_{2\alpha}$ will have an assumed sign, or else the model is simply not specified well enough to yield any results. In the preceding profit maximization model, for the factor prices (recall, f in that model designates only the production function, not the whole objective function),

$$f_{i\alpha} = \pi_{iw_i} = -1 \quad i = 1, 2 \tag{6-14}$$

and

$$\pi_{1w_2} = \pi_{2w_1} = 0$$

The parameter w_1, for example, appears only in the first first-order equation, $\pi_1 = 0$. That is, $f_{2\alpha} \equiv 0$, in Eqs. (6-10). For that reason, the term involving the cross-partial f_{12} in Eq. (6-12a) is 0. Since $\pi_{1w_1} = -1$, the result $\partial x_1^*/\partial w_1 < 0$ is obtained for the profit maximization model.

Similarly, for w_2, $f_{2\alpha} \equiv -1$, $f_{1\alpha} \equiv 0$. Hence, in Eq. (6-12b), the only remaining term on the right-hand side is $-f_{2\alpha} f_{11}/H$. From the second-order conditions, $\partial x_2^*/\partial w_2 < 0$ is implied.

The situation is different for the parameter p, output price. Output price enters *both* first-order equations (6-10). Therefore, the indeterminate cross-term f_{12} appears in the expressions for $\partial x_1^*/\partial p$ and $\partial x_2^*/\partial p$. As a result, no refutable hypotheses emerge for this parameter with regard to each input.

The preceding analysis suggests that refutable comparative statics theorems will be forthcoming in a maximization model only if a given parameter enters one and only one first-order equation. This result, known as the *conjugate pairs theorem,* will be shown in greater generality in the succeeding sections. From Eqs. (6-12), if some parameter α_i enters only the ith first-order equation, then $\partial x_i^*/\partial \alpha_i$ and $f_{i\alpha_i}$ must have the same sign. This can be expressed as

$$f_{i\alpha_i}\frac{\partial x_i^*}{\partial \alpha_i} > 0 \tag{6-15}$$

Virtually all of the comparative statics results in economics are specific instances of Eq. (6-15), where some parameter α_i enters only the ith first-order equation.

6.2 GENERALIZATION TO n VARIABLES

Let us now investigate how the two-factor, profit maximization model is generalized to n factors. We must first derive the first- and second-order conditions for an unconstrained maximum (and minimum). We will then use the profit maximization model to motivate and illustrate the general methodology of comparative statics.

First-Order Necessary Conditions

As we noted in Chap. 4, the necessary first-order conditions for $y = f(x_1, \ldots, x_n)$ to have a stationary value are that all the first partials of f equal zero; that is, $f_i = 0$, $i = 1, \ldots, n$. This is a straightforward and intuitive generalization of the two-variable case. The second-order conditions, however, are a bit more complex.

Second-Order Sufficient Conditions

Using a Taylor series approach, as was done in the Appendix to Chap. 4, it can be shown that a sufficient condition for $y = f(x_1, \ldots, x_n)$ to have a maximum at some stationary value is that for all curves, $y(t) = f(x_1(t), \ldots, x_n(t))$, $y''(t) < 0$. Using the chain rule, this sufficient condition is

$$\frac{d^2y}{dt^2} = \sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij}\frac{dx_i}{dt}\frac{dx_j}{dt} < 0 \tag{6-16}$$

for all dx_i/dt, dx_j/dt not all equal to 0.

A square matrix (a_{ij}) which has the property that

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}h_ih_j < 0 \tag{6-17}$$

for all nontrivial (not all 0) h_i, h_j is said to be *negative definite*. (If the strict inequality is replaced by "≤ 0," the matrix is called *negative semidefinite*.) Similarly, (a_{ij}) positive definite (semidefinite) means that the sum in (6-17) is strictly positive (nonnegative) for all nontrivial h_i, h_j. Thus, if at a point where $f_i = 0, i = 1, \ldots, n$, the matrix of second partials of f (called the *Hessian* matrix) is negative definite, then $f(x_1, \ldots, x_n)$ has a maximum there. If the Hessian matrix is positive definite there, a minimum exists. If the Hessian is negative semidefinite, then f definitely does *not* have a minimum, but it is not possible to say whether f has a maximum or some sort of saddle point at the stationary value. An expression of the form (6-17), in matrix form $\mathbf{h}'\mathbf{A}\mathbf{h}$, is called a *quadratic form.*

Geometrically, negative definiteness of the Hessian matrix

$$\mathbf{H} = \begin{pmatrix} f_{11} & \cdots & f_{1n} \\ \vdots & & \\ f_{n1} & & f_{nn} \end{pmatrix}$$

ensures that the function f will be *strictly concave* (downward). If $\mathbf{H}$ is positive definite, f is strictly convex.

Example. Consider the function $y = (x_2 - x_1^2)(x_2 - 2x_1^2)$ depicted in Fig. 4-1 of Chap. 4. This is a function that has a minimum at the origin when evaluated along all straight lines through the origin, yet the function itself does *not* have a minimum there. The Hessian matrix of second partials is

$$\mathbf{H} = \begin{pmatrix} 24x_1^2 - 6x_2 & -6x_1 \\ -6x_1 & 2 \end{pmatrix}$$

At the origin, this matrix is

$$\mathbf{H} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

This matrix is clearly positive *semidefinite:*

$$Q = \sum_{j=1}^{2} \sum_{i=1}^{2} f_{ij} h_i h_j = 2h_2^2 \geq 0$$

When h_1 = anything, $h_2 = 0$, this quadratic form $Q = 0$; when $h_2 \neq 0$, $Q > 0$.

In the two-variable case, $y = f(x_1, x_2)$, the sufficient second-order conditions for a maximum, (6-16), imply that $f_{11} < 0$, $f_{22} < 0$, and $f_{11}f_{22} - f_{12}^2 > 0$, as was shown in Chap. 4. Note that this last expression can be stated as the determinant of the cross-partials of the objective function,

$$\begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix} > 0$$

Note also that the conditions $f_{11}, f_{22} < 0$ relate to the diagonal elements of that determinant. The theory of determinants allows a very simple statement of the sufficient second-order conditions for $y = f(x_1, \ldots, x_n)$ to have a maximum. First, consider the following construction:

Definition. Let A_n be some nth-order determinant. By a "principal minor of order k" of A_n we mean that determinant which remains of A_n when any $n - k$ rows and the *same numbered columns* are eliminated from A_n.

For example, if some row, row i, is eliminated, then to form a principal minor of order $n - 1$, *column* i must be eliminated. Since there are n choices of rows (and their corresponding columns) to eliminate, there are clearly n principal minors of order $n-1$ of A_n. If, say, rows 1 and 3 and columns 1 and 3 are eliminated, then a principal minor of order $n-2$ remains. There are $\binom{n}{2} = n(n-1)/2!$ of these, and in general $\binom{n}{k} = n!/k!(n-k)!$ principal minors of order k [or order $(n-k)$]. Note that the first-order principal minors of A_n are simply the diagonal elements of A_n, and the second-order principal minors are the set of 2×2 determinants that look like

$$\begin{vmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{vmatrix}$$

The resemblance of this determinant to the 2×2 determinant of cross-partials of a function $f(x_1, x_2)$ provides the motivation for the following theorem.

Theorem. Consider a function $y = f(x_1, \ldots, x_n)$ that has a stationary value at $\mathbf{x} = \mathbf{x}^0$. Consider the Hessian matrix of cross-partials of f, (f_{ij}). Then if all of the principal minors of $|(f_{ij})|$ of order k have sign $(-1)^k$, for all $k = 1, \ldots, n$ ($k = n$ yields the whole determinant, $|(f_{ij})|$) at $\mathbf{x} = \mathbf{x}^0$, then $f(x_1, \ldots, x_n)$ has a maximum at $\mathbf{x} = \mathbf{x}^0$. If all the principal minors of $|(f_{ij})|$ are positive, for all $k = 1, \ldots, n$, at $\mathbf{x} = \mathbf{x}^0$, then $f(x_1, \ldots, x_n)$ has a minimum value at $\mathbf{x} = \mathbf{x}^0$. If any of the principal minors has a sign strictly opposite to that stated above, the function has a saddle point at $\mathbf{x} = \mathbf{x}^0$. If some or all of the principal minors are 0 and the rest have the appropriate sign given in the preceding conditions, then it is not possible to indicate the shape of the function at $\mathbf{x} = \mathbf{x}^0$. (This corresponds to the 0 second-derivative situation in the calculus of functions of one variable.)

The theorem as stated is the form in which we shall actually use the result. However, it is somewhat overstated. Consider the "naturally ordered" principal minors of an $n \times n$ Hessian,

$$|f_{11}| \quad \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix} \quad \begin{vmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{vmatrix} \quad \cdots$$

Recall that in the two-variable case, $f_{11} < 0$ and $f_{11}f_{22} - f_{12}^2 > 0$ implies $f_{22} < 0$. In fact, if all of these naturally ordered principal minors have the appropriate sign for a maximum or minimum of $f(x_1, \ldots, x_n)$, then all of the other principal minors have the appropriate sign. Thus, the theorem as stated is in some sense "too strong"; i.e., more is assumed than is necessary, but we shall need the sufficient condition that *all* principal minors of order k have sign $(-1)^k$ for a maximum, or that they are *all* positive for a minimum.

There are several inelegant proofs of this theorem, one by completing a rather gigantic square à la the proof used in Chap. 4, and an elegant proof based on matrix theory, a proof that is beyond the level of this book.† Hence, no proof will be offered.

†See George Hadley, *Linear Algebra,* Addison-Wesley Publishing Co., Inc., Reading, MA, 1961.

It is hoped that the discussion of the two-variable case will have at least made the theorem not implausible.

Profit Maximization: *n* Factors

Consider the profit-maximizing firm with n factors of production. The objective function, again, is

maximize

$$\pi = pf(x_1, \ldots, x_n) - \sum w_i x_i$$

The first-order conditions, again, are

$$\pi_i = pf_i - w_i = 0 \qquad i = 1, \ldots, n \tag{6-18}$$

The firm equates the value of marginal product to the wage at every margin, i.e., for every factor input. This is a straightforward generalization of the two-variable case. These equations represent n equations in the n decision variables $x_1, \ldots, x_n$ and $n + 1$ parameters $w_1, \ldots, w_n$, p. If the Jacobian determinant is nonzero, i.e.,

$$J = \left| \frac{\partial \pi_i}{\partial x_j} \right| \neq 0 \tag{6-19}$$

then at this stationary value, these equations can be solved for the explicit choice functions, i.e., the factor demand curves,

$$x_i = x_i^*(w_1, \ldots, w_n, p) \quad i = 1, \ldots, n \tag{6-20}$$

The sufficient conditions for a maximum are that the principal minors of $(\pi_{ij}) = (pf_{ij})$ alternate in sign, i.e., have sign $(-1)^k$, $k = 1, \ldots, n$. Since $p > 0$, this is equivalent to saying that the principal minors of the matrix of second partials of the production function,

$$\begin{vmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & & f_{2n} \\ \vdots & & & \\ f_{n1} & f_{n2} & & f_{nn} \end{vmatrix}$$

alternate in sign. Specifically, this means that, among other things, the diagonal terms are all negative, that is, $f_{ii} < 0$, $i = 1, \ldots, n$. This says that there is diminishing marginal productivity in each factor. In addition, all $n(n-1)/2$ second-order determinants

$$\begin{vmatrix} f_{ii} & f_{ij} \\ f_{ji} & f_{jj} \end{vmatrix} > 0 \quad i, j = 1, \ldots, n \quad i \neq j$$

The "own-effects" dominate cross-effects in the sense that $f_{ii} f_{jj} - f_{ij}^2 > 0$, $i, j = 1, \ldots, n$, $i \neq j$. Then there are all the remaining principal minors to consider; these are not easily given intuitive explanations.

The sufficient second-order conditions say that in a neighborhood of a maximum point, the objective function (in this example, this is equivalent to the production function) must be strictly concave (downward). The conditions $f_{ii} < 0$ ensure that the function is concave in all the two-dimensional planes whose axes are y and some x_i. The second-order principal minors relate to concavity in all possible three-dimensional subspaces y, x_i, x_j. But concavity in all of these lower-order dimensions is not sufficient to guarantee concavity in higher dimensions; hence, all the orders of principal minors, including the whole Hessian determinant itself, must be checked for the appropriate sign.

In terms of solving for the factor demand curves, the *sufficient* second-order conditions guarantee that this is possible. The nth-order principal minor, i.e., the determinant of the entire (π_{ij}) matrix, has sign $(-1)^n \neq 0$ by these sufficient conditions. But this determinant is precisely the Jacobian of the system (6-18); hence, applying the implicit function theorem, the choice functions (6-20) are derivable from (6-18).

Substituting the choice functions (6-20) back into (6-18) yields the identities

$$pf_i(x_1^*, \dots, x_n^*) - w_i \equiv 0 \quad i = 1, \dots, n \tag{6-21}$$

To find the responses of the system to a change in some factor price w_j, differentiate (6-21) with respect to w_j. This yields the system of equations

$$pf_{11}\frac{\partial x_1^*}{\partial w_j} + \cdots + pf_{1n}\frac{\partial x_n^*}{\partial w_j} \equiv 0$$

$$\vdots$$

$$pf_{j1}\frac{\partial x_1^*}{\partial w_j} + \cdots + pf_{jn}\frac{\partial x_n^*}{\partial w_j} \equiv 1$$

$$\vdots$$

$$pf_{n1}\frac{\partial x_1^*}{\partial w_j} + \cdots + pf_{nn}\frac{\partial x_n^*}{\partial w_j} \equiv 0$$

In matrix notation, this system is written

$$\begin{pmatrix} pf_{11} & \cdots & pf_{1n} \\ \vdots & & \\ pf_{n1} & & pf_{nn} \end{pmatrix} \begin{pmatrix} \dfrac{\partial x_1^*}{\partial w_j} \\ \vdots \\ \dfrac{\partial x_n^*}{\partial w_j} \end{pmatrix} \equiv \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \tag{6-22}$$

where the 1 on the right-hand side appears in row j. Solving for $\partial x_i^*/\partial w_j$ by Cramer's rule involves putting the right-hand column in column i of the $|(pf_{ij})|$ determinant, in the numerator, i.e.,

$$\frac{\partial x_i^*}{\partial w_j} = \frac{\begin{vmatrix} pf_{11} & 0 & pf_{1n} \\ & \vdots & \\ \vdots & 1 & \\ & \vdots & \\ pf_{n1} & 0 & pf_{nn} \end{vmatrix}}{H} \tag{6-23}$$

where $H = |pf_{ij}|$, the Jacobian determinant of second partials of π. Expanding the numerator by the cofactors of column i,

$$\frac{\partial x_i^*}{\partial w_j} = \frac{H_{ji}}{H} \tag{6-24}$$

where H_{ji} is the cofactor of the element in row j and column i of H.

In general, H has sign $(-1)^n$ by the sufficient second-order conditions for a maximum. For $i \neq j$, however, the sign of H_{ij} is not implied by the maximum conditions. Thus, in general, no refutable implications emerge for the response of any factor to a change in the price of some *other* factor. However, when $i = j$,

$$\frac{\partial x_i^*}{\partial w_i} = \frac{H_{ii}}{H} \tag{6-25}$$

The cofactor H_{ii} is a principal minor; by the maximum conditions it has sign $(-1)^{n-1}$, i.e., opposite to the sign of H. Thus,

$$\frac{\partial x_i^*}{\partial w_i} = \frac{H_{ii}}{H} < 0 \quad i = 1, \ldots, n \tag{6-26}$$

As in the two-factor case, the model does yield a refutable hypothesis concerning the slope of each factor demand curve. The response of any factor to a change in its *own* price is in the opposite direction to the change in its price.

Finally, from the symmetry of H, using Eq. (6-24),

$$\frac{\partial x_i^*}{\partial w_j} = \frac{H_{ji}}{H} = \frac{H_{ij}}{H} = \frac{\partial x_j^*}{\partial w_i} \tag{6-27}$$

The reciprocity conditions thus generalize in a straightforward fashion to the n-factor case. Since the parameter p enters each first-order equation (6-18), no refutable hypotheses emerge for the responses of factor inputs to output price changes. The matrix system of comparative statics relations obtained from differentiating (6-18) with respect to p are [compare Eqs. (4-21), Chap. 4]:

$$\begin{pmatrix} pf_{11} & \cdots & pf_{1n} \\ \vdots & & \\ pf_{n1} & & pf_{nn} \end{pmatrix} \begin{pmatrix} \dfrac{\partial x_1^*}{\partial p} \\ \vdots \\ \dfrac{\partial x_n^*}{\partial p} \end{pmatrix} = \begin{pmatrix} -f_1 \\ \vdots \\ -f_n \end{pmatrix} \tag{6-28}$$

Solving by Cramer's rule for $\partial x_i^*/\partial p$,

$$\frac{\partial x_i^*}{\partial p} = -\sum_{j=1}^{n} \frac{f_j H_{ji}}{H} \gtrless 0 \tag{6-29}$$

It can be shown that if p increases, then at least one factor must increase, but this is precious little information.

Finally, the supply function of this competitive firm is defined as

$$y = f(x_1^*(\mathbf{w}, p), \ldots, x_n^*(\mathbf{w}, p)) = y^*(w_1, \ldots, w_n, p)$$

where $\mathbf{w}$ is the vector of factor prices $(w_1, \ldots, w_n)$. It can be shown that

$$\frac{\partial y^*}{\partial p} > 0 \tag{6-30}$$

and

$$\frac{\partial y^*}{\partial w_i} = -\frac{\partial x_i^*}{\partial p} \qquad i = 1, \ldots, n \tag{6-31}$$

We shall leave these results to a later chapter, as they are difficult to obtain by the present methods and outrageously simple by methods involving what is known as the *envelope theorem,* which will be discussed later.

We now state an interesting theorem without proof and apply it to the profit maximization model.

> **Theorem.** Let H be an $n \times n$ negative definite matrix (whose diagonal elements are all necessarily negative) and whose off-diagonal elements are all positive. Then the inverse matrix H^{-1} consists entirely of negative entries.

This theorem is evident upon inspection for the 2×2 and 3×3 cases; however, we have found no simple proof for the general case. The proof is an application of what are known as the Perron-Frobenius theorems. We refer the reader to A. Takayama's text† for discussion and proof of these propositions.

Consider the application of this theorem to the profit maximization model. For changes in some wage w_j, we get the matrix equation (6-22) above. Let b be the column vector on the right-hand side of this equation; it consists of zeros in every row except row j, in which the element $+1$ appears. The solution to this equation, in matrix form, is $\partial x_i/\partial w_j = H^{-1}b$. Since every element of H^{-1} is negative and b is either 1 or 0, $\partial x_i/\partial w_j < 0$, $i, j = 1, \ldots, n$. In the two-variable model, we showed that the sign of $\partial x_1/\partial w_2$ is the same as the sign of $-f_{12}$. With only two factors, technical complementarity ($f_{12} > 0$) is the same as complementarity defined in terms of the change in the use of one factor as the price of the other factor changes.

†*Mathematical Economics,* 2d ed., Cambridge Press, Cambridge, England, 1985, pp. 392ff.

However, if more than two factors are present, one cannot infer that if, say, $f_{13} > 0$, then $\partial x_1/\partial w_3 < 0$; the signs of the other cross-partials of the production function matter. The above theorem shows, however, that if *all* the factors are complements in the sense of $f_{ij} > 0$, then $\partial x_i/\partial w_j < 0$ for all the factors.

Likewise, consider Eq. (6-28) for the responses to a change in output price. Assuming the marginal products of each factor are positive, the solution of this equation is the matrix product of H^{-1}, which has only negative elements, and the column vector of the negatives of the marginal products of each factor. It therefore follows that $\partial x_i/\partial p > 0$ for all factors; i.e., there are no inferior factors with these assumptions.

6.3 THE THEORY OF CONSTRAINED MAXIMA AND MINIMA: FIRST-ORDER NECESSARY CONDITIONS

In most of the maximization problems encountered in economics, a separate, additional equation appears that constrains the values of the decision variables to some subspace of all real values, i.e., some subspace of what is referred to as Euclidean n-space. For example, in the theory of the consumer, individuals are posited to maximize a utility function, $U(x_1, x_2)$, subject to a constraint that dictates that the consumer not exceed a certain total budgetary expenditure. This problem can be stated more formally as

maximize

$$U(x_1, x_2) = U \tag{6-32}$$

subject to

$$p_1x_1 + p_2x_2 = M \tag{6-33}$$

where x_1 and x_2 are the amounts of two goods consumed, p_1 and p_2 are their respective prices, and M is total money income. This problem can be solved simply by solving for one of the decision variables, say x_2, from the constraint and inserting that solution into the objective function. In that case, an unconstrained problem of one less dimension results: From (6-33),

$$x_2(x_1) = \frac{-p_1}{p_2}x_1 + \frac{M}{p_2} \tag{6-34}$$

Since once x_1 is known, x_2 is known also from the preceding, the problem reduces to maximizing $U(x_1, x_2(x_1))$ over the one decision variable x_1. This yields

$$\begin{aligned}\frac{dU}{dx_1} &= \frac{\partial U}{\partial x_1} + \frac{\partial U}{\partial x_2}\frac{dx_2}{dx_1}\\ &= U_1 + U_2\frac{-p_1}{p_2} = 0\end{aligned}$$

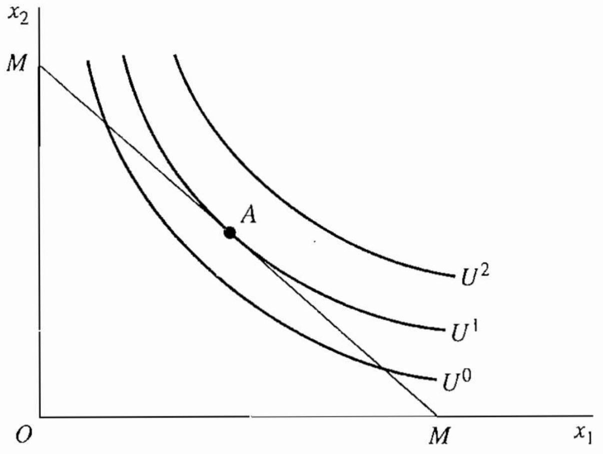

FIGURE 6-1
Utility Maximization. In this diagram, three indifference levels are drawn, with $U^2 > U^1 > U^0$. The line MM represents a consumer's budget constraint. The constrained utility maximum occurs at point A, where the indifference curve is tangent to (has the same slope as) the budget constraint. The second-order conditions for a maximum say that the level curves of the utility function, i.e., the indifference curves, must be convex to the origin; i.e., the utility function must be "quasi-concave" (in addition to strictly increasing).

or

$$\frac{U_1}{U_2} = \frac{p_1}{p_2} \tag{6-35}$$

This is the familiar tangency condition that the marginal rate of substitution ($-U_1/U_2$, the rate at which a consumer is *willing* to trade off x_2 for x_1) is equal to the opportunity to do so in the market ($-p_1/p_2$, the slope of the budget line). The condition is illustrated in Fig. 6-1. Under the right curvature conditions on the utility function (to be guaranteed by the appropriate second-order conditions), point A clearly represents the maximum achievable utility if the consumer is constrained to consume some consumption bundle along the budget line MM.

The more general constrained maximum problem,

maximize

$$f(x_1, \ldots, x_n) = y$$

subject to

$$g(x_1, \ldots, x_n) = 0$$

can be solved in the same way, i.e., by direct substitution, reducing the problem to an unconstrained one in $n - 1$ dimensions. However, a highly elegant solution that preserves the symmetry of the problem, known as the method of Lagrange multipliers (after the French mathematician Lagrange), will be given instead. The proof proceeds along the lines developed earlier for unconstrained maxima.

Consider the behavior of the function $f(x_1, \ldots, x_n)$ along some differentiable curve $\mathbf{x}(t) = (x_1(t), \ldots, x_n(t))$; that is, consider $y(t) = f(x_1(t), \ldots, x_n(t))$. If $y'(t) = 0$ and $y''(t) < 0$ for every feasible curve $\mathbf{x}(t)$, then $f(x_1, \ldots, x_n)$ has a maximum at that point. However, in this case, $\mathbf{x}(t)$ cannot represent *all* curves in n-space. Only those curves that lie in the constraint are admissible. This smaller family of curves comprises those curves for which $g(x_1(t), \ldots, x_n(t)) \equiv 0$. Notice the identity sign—we mean to ensure that $g(x_1, \ldots, x_n)$ is 0 *for every point along a* *given curve* $\mathbf{x}(t)$, not just for some points. The problem can be stated as follows:

maximize

$$f(x_1(t), \ldots, x_n(t)) = y(t) \tag{6-36}$$

subject to

$$g(x_1(t), \ldots, x_n(t)) \equiv 0 \tag{6-37}$$

Setting $y'(t) = 0$ yields

$$f_1 \frac{dx_1}{dt} + \cdots + f_n \frac{dx_n}{dt} = 0 \tag{6-38}$$

for all values of the dx_i/dt that satisfy the constraint. What restriction does $g(x_1(t), \ldots, x_n(t)) \equiv 0$ place on these values? Differentiating g with respect to t yields

$$g_1 \frac{dx_1}{dt} + \cdots + g_n \frac{dx_n}{dt} \equiv 0 \tag{6-39}$$

In the unconstrained case, the expression (6-38) was zero for *all* dx_i/dt; thus, in that case $f_i = 0, i = 1, \ldots, n$, was necessary for a maximum. Here, however, (6-38) *and* (6-39) must hold simultaneously. Hence, the values of dx_i/dt are not completely unrestricted. However, assuming $f_1 \neq 0$, we can write, from (6-38),

$$\frac{dx_1}{dt} = -\frac{f_2}{f_1}\frac{dx_2}{dt} - \cdots - \frac{f_n}{f_1}\frac{dx_n}{dt} \tag{6-40}$$

Similarly, from (6-39), if $g_1 \neq 0$,

$$\frac{dx_1}{dt} = -\frac{g_2}{g_1}\frac{dx_2}{dt} - \cdots - \frac{g_n}{g_1}\frac{dx_n}{dt} \tag{6-41}$$

Subtracting (6-40) from (6-41) yields, after factoring,

$$\left(\frac{f_2}{f_1} - \frac{g_2}{g_1}\right)\frac{dx_2}{dt} + \cdots + \left(\frac{f_n}{f_1} - \frac{g_n}{g_1}\right)\frac{dx_n}{dt} = 0 \tag{6-42}$$

and, what is more, this expression must be 0 for *all* $dx_2/dt, \ldots, dx_n/dt$. By eliminating one of the dx_i/dt's, the remaining dx_i/dt's can have unrestricted values. If $f_1 \neq 0$, $g_1 \neq 0$, then for any values whatsoever of $dx_2/dt, \ldots, dx_n/dt$, a judicious choice of dx_1/dt will allow (6-38) and (6-39) to hold. But since (6-42) holds for any values at all of $dx_2/dt, \ldots, dx_n/dt$, it must be true that the coefficients in parentheses are all 0; i.e., $f_i/f_1 = g_i/g_1, i = 2, \ldots, n$. In the case where all of the f_i, g_i are not 0, these conditions can be expressed simply as

$$\frac{f_i}{f_j} = \frac{g_i}{g_j} \qquad i, j = 1, \ldots, n \tag{6-43}$$

These $n - 1$ conditions say that the level curves of the objective function have to be parallel to the level curves of the constraint. This is the familiar tangency condition, illustrated by the preceding utility maximization problem. The $n-1$ conditions (6-43) and the constraint (6-37) itself constitute the complete set of first-order conditions for a constrained maximum problem with one constraint. Of course, these first-order conditions are necessary for any stationary value—maximum, minimum, or saddle shape.

The above conditions can be given an elegant and useful formulation by constructing a new function $\mathcal{L}$ called a Lagrangian, where

$$\mathcal{L} = f(x_1, \ldots, x_n) + \lambda g(x_1, \ldots, x_n)$$

The variable λ is simply a new, independent variable and is called a Lagrange multiplier.† Note that $\mathcal{L}$ always equals f for values of $x_1, \ldots, x_n$ that satisfy the constraint. Thus, $\mathcal{L}$ can be expected to have a stationary value when f does. Indeed, taking the partials of $\mathcal{L}$ with respect to $x_1, \ldots, x_n$ and λ and setting them equal to 0 yields

$$\begin{aligned} \mathcal{L}_1 &= f_1 + \lambda g_1 = 0 \\ &\vdots \\ \mathcal{L}_n &= f_n + \lambda g_n = 0 \\ \mathcal{L}_\lambda &= g(x_1, \ldots, x_n) = 0 \end{aligned} \tag{6-44}$$

Eliminating λ from the first n equations of (6-44) (by bringing λg_i over to the right-hand side and dividing one equation by another) yields

$$\frac{f_i}{f_j} = \frac{g_i}{g_j}$$

precisely the first-order conditions for a constrained maximum. Hence, the Lagrangian function provides an easy mnemonic for writing the first-order conditions for constrained maximum problems. However, we shall see that this is a most useful construction for the second-order conditions also, and, in the theory of comparative statics, the Lagrange multiplier λ often has an interesting economic interpretation.

Example. Consider again the utility maximization problem analyzed at the beginning of this section. The Lagrangian for this problem is

$$\mathcal{L} = U(x_1, x_2) + \lambda(M - p_1 x_1 - p_2 x_2)$$

Differentiating $\mathcal{L}$ with respect to x_1, x_2, and λ yields

$$\mathcal{L}_1 = U_1 - \lambda p_1 = 0 \tag{6-45a}$$

$$\mathcal{L}_2 = U_2 - \lambda p_2 = 0 \tag{6-45b}$$

$$\mathcal{L}_\lambda = M - p_1 x_1 - p_2 x_2 = 0 \tag{6-45c}$$

The partial $\mathcal{L}_\lambda$ is simply the budget constraint again since $\mathcal{L}$ is linear in λ. The variable λ can be eliminated from (6-45a) and (6-45b) by bringing λp_1, λp_2 over to the

†It is of no consequence whether one writes $\mathcal{L} = f + \lambda g$ or $\mathcal{L} = f - \lambda g$; this merely changes the sign of the Lagrange multiplier.

right-hand side and then dividing one equation by the other. This yields $U_1/U_2 = p_1/p_2$, the tangency conditions (6-35) arrived at by direct substitution.

There are many problems in economics in which more than one constraint appears. For example, a famous general equilibrium model is that of the "small country" which maximizes the value of its output with fixed world prices, subject to constraints which say that the amount of each of several factors of production used cannot exceed a given amount. The general mathematical structure of maximization problems with r constraints is

maximize

$$f(x_1, \ldots, x_n) = y \tag{6-46}$$

subject to

$$\begin{aligned} g^1(x_1, \ldots, x_n) &= 0 \\ &\vdots \\ g^r(x_1, \ldots, x_n) &= 0 \end{aligned} \tag{6-47}$$

These are r equations where, of necessity, $r < n$. (Why?)

The first-order conditions for this problem can be found by generalizing the Lagrange multiplier method previously derived. Multiplying each constraint by its own Lagrange multiplier λ^j, form the Lagrangian

$$\mathcal{L} = f(x_1, \ldots, x_n) + \lambda^1 g^1(x_1, \ldots, x_n) + \cdots + \lambda^r g^r(x_1, \ldots, x_n) \tag{6-48}$$

Then the first partials of $\mathcal{L}$ with respect to the $n + r$ variables x_i, λ^j give the correct first-order conditions:

$$\mathcal{L}_i = f_i + \lambda^1 g_i^1 + \cdots + \lambda^r g_i^r = 0 \quad i = 1, \ldots, n \tag{6-49}$$

$$\mathcal{L}_j = g^j = 0 \quad j = 1, \ldots, r \tag{6-50}$$

where g_i^j means $\partial g^j/\partial x_i$. The proof of this can be obtained only by more advanced methods; it is given in the next section.

6.4 CONSTRAINED MAXIMIZATION WITH MORE THAN ONE CONSTRAINT: A DIGRESSION†

Consider the maximization problem

maximize

$$f(x_1, \ldots, x_n) = y$$

†In order to understand this section, the student must be familiar with some concepts of linear algebra, such as rank of a matrix, etc., developed in the Appendix to Chap. 5. We are indebted to Ron Heiner for demonstrating this approach to the problem to us.

subject to

$$g^1(x_1, \ldots, x_n) = 0$$
$$\vdots$$
$$g^r(x_1, \ldots, x_n) = 0$$

Letting $x_i = x_i(t), i = 1, \ldots, n$, as before, the first-order conditions for a maximum (or any stationary value) are

$$\frac{dy}{dt} = f_1 \frac{dx_1}{dt} + \cdots + f_n \frac{dx_n}{dt} = 0 \tag{6-51}$$

for any $dx_1/dt, \ldots, dx_n/dt$ satisfying

$$\begin{aligned} g_1^1 \frac{dx_1}{dt} + \cdots + g_n^1 \frac{dx_n}{dt} &= 0 \\ &\vdots \\ g_1^r \frac{dx_1}{dt} + \cdots + g_n^r \frac{dx_n}{dt} &= 0 \end{aligned} \tag{6-52}$$

where $g_i^j = \partial g^j / \partial x_i$.

For any function $y = f(x_1, \ldots, x_n)$, the *gradient* of f, written $\nabla \mathbf{f}$, is a vector composed of the first partials of f:

$$\nabla \mathbf{f} = (f_1, \ldots, f_n)$$

The differential of f can be written

$$dy = \nabla \mathbf{f}\, \mathbf{dx}$$

where $\mathbf{dx} = (dx_1, \ldots, dx_n)$. Along a level surface, $dy = 0$, and hence $\nabla \mathbf{f}$ is orthogonal to the direction of the tangent hyperplane. The gradient of f, $\nabla \mathbf{f}$, thus represents the direction of maximum increase of $f(x_1, \ldots, x_n)$.

Note that Eq. (6-51) is the scalar product of the gradient of f, $\nabla \mathbf{f}$, and the vector $\mathbf{h} = (h_1, \ldots, h_n) = (dx_1/dt, \ldots, dx_n/dt)$. Likewise, Eqs. (6-52) are the scalar products of the gradients of the g^j functions, $\nabla \mathbf{g}^j$, and $\mathbf{h}$. Let $\nabla \mathbf{g}$ denote the $r \times n$ matrix whose rows are, respectively, $\nabla \mathbf{g}^1, \ldots, \nabla \mathbf{g}^r$. Then Eqs. (6-51) and (6-52) can be written, respectively,

$$\nabla \mathbf{f} \cdot \mathbf{h} = 0 \tag{6-53}$$

for all $\mathbf{h} \neq 0$ satisfying

$$(\nabla \mathbf{g})\mathbf{h} = \mathbf{0} \tag{6-54}$$

Assume now that the matrix $\nabla \mathbf{g}$ has rank r, equal to the number of constraints. This says that the constraints are independent, i.e., there are no redundant constraints. If the rank of $\nabla \mathbf{g}$ was less than r, say $r - 1$, then one constraint could be dropped and the subspace in which the dx_i/dt could range would not be affected. It is as if a ration-point constraint were imposed with the ration prices proportional to the original money prices. In that case, the additional rationing constraint would either be redundant to or inconsistent with the original budget constraint.

Assuming rank $\nabla \mathbf{g} = r$, the rows of $\nabla \mathbf{g}$, that is, the gradient vectors $\nabla \mathbf{g}^j = (g_1^j, \ldots, g_n^j)$, $j = 1, \ldots, r$, form a basis for an r-dimensional subspace E_r of E_n, Euclidean n-space. From (6-54), the admissible vectors $\mathbf{h}$ are all orthogonal to E_r; hence, they must all lie in the remaining $n - r$ dimensional space, E_r'. However, from (6-53), $\nabla \mathbf{f}$ is orthogonal to all those $\mathbf{h}$'s and hence to E_r'. Hence, $\nabla \mathbf{f}$ must lie in E_r. Since the vectors $\nabla \mathbf{g}^j$ form a basis for E_r, $\nabla \mathbf{f}$ can be written as a unique linear combination of those vectors, or

$$\nabla \mathbf{f} = \lambda^1 \nabla \mathbf{g}^1 + \cdots + \lambda^r \nabla \mathbf{g}^r \tag{6-55}$$

However, this is equivalent to setting the partial derivatives of the Lagrangian expression $\mathcal{L} = f - \sum \lambda^j g^j$ with respect to $x_1, \ldots, x_n$ equal to 0.

6.5 SECOND-ORDER CONDITIONS

In the past two sections, the first-order necessary conditions for a function to achieve a stationary value subject to constraints were derived. Those conditions are implied whenever the function has a maximum, a minimum, or a saddle shape (a minimum in some directions and a maximum in others). We now seek to state sufficient conditions under which the type of stationary position can be specified. The discussion will be largely limited to the two-variable case, with the general theorems stated at the end of this section.

Consider the two-variable problem

maximize

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

subject to

$$g(x_1, x_2) = 0$$

The Lagrangian function is $\mathcal{L}(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2)$. The first-order conditions are, again,

$$\frac{dy}{dt} = f_1 \frac{dx_1}{dt} + f_2 \frac{dx_2}{dt} = 0 \tag{6-56}$$

for all dx_1/dt, dx_2/dt satisfying

$$g_1 \frac{dx_1}{dt} + g_2 \frac{dx_2}{dt} \equiv 0 \tag{6-57}$$

These conditions imply that $\mathcal{L}_1 = f_1 + \lambda g_1 = 0$, $\mathcal{L}_2 = f_2 + \lambda g_2 = 0$. Sufficient conditions for these equations to represent a relative *maximum* are that $d^2y/dt^2 < 0$, for all dx_1/dt, dx_2/dt satisfying (6-57). Similarly, $d^2y/dt^2 > 0$ under those conditions implies a *minimum*. How can these conditions be put into a more useful form? Differentiating (6-56) again with respect to t, the sufficient second-order condition is

$$\frac{d^2y}{dt^2} = f_1\frac{d^2x_1}{dt^2} + f_2\frac{d^2x_2}{dt^2} + f_{11}\left(\frac{dx_1}{dt}\right)^2 + 2f_{12}\frac{dx_1}{dt}\frac{dx_2}{dt} + f_{22}\left(\frac{dx_2}{dt}\right)^2 < 0 \tag{6-58}$$

subject to

$$g_1\frac{dx_1}{dt} + g_2\frac{dx_2}{dt} \equiv 0 \tag{6-57}$$

Since (6-57) is an identity, differentiate it again with respect to t, remembering that g_1 and g_2 are functions of $x_1(t)$, $x_2(t)$. This yields

$$g_1\frac{d^2x_1}{dt^2} + g_2\frac{d^2x_2}{dt^2} + g_{11}\left(\frac{dx_1}{dt}\right)^2 + 2g_{12}\frac{dx_1}{dt}\frac{dx_2}{dt} + g_{22}\left(\frac{dx_2}{dt}\right)^2 \equiv 0 \tag{6-59}$$

Now multiply (6-59) through by λ, the Lagrange multiplier, and add to Eq. (6-58). Since this amounts to adding 0,

$$\frac{d^2y}{dt^2} = (f_1+\lambda g_1)\frac{d^2x_1}{dt^2} + (f_2+\lambda g_2)\frac{d^2x_2}{dt^2} + (f_{11}+\lambda g_{11})\left(\frac{dx_1}{dt}\right)^2$$

$$+2(f_{12}+\lambda g_{12})\frac{dx_1}{dt}\frac{dx_2}{dt} + (f_{22}+\lambda g_{22})\left(\frac{dx_2}{dt}\right)^2 < 0 \tag{6-60}$$

subject to (6-57). However, from the first-order conditions, $\mathcal{L}_1 = f_1 + \lambda g_1 = 0$, $\mathcal{L}_2 = f_2 + \lambda g_2 = 0$. Also, $f_{11} + \lambda g_{11}$ is simply $\mathcal{L}_{11}$, and likewise $\mathcal{L}_{12} = f_{12} + \lambda g_{12}$, etc. If we simplify the notation a bit and write $h_1 = dx_1/dt$, $h_2 = dx_2/dt$, then the sufficient second-order conditions for a maximum are that

$$\mathcal{L}_{11}h_1^2 + 2\mathcal{L}_{12}h_1h_2 + \mathcal{L}_{22}h_2^2 < 0 \tag{6-61}$$

for all h_1, h_2 not both equal to 0, such that

$$g_1h_1 + g_2h_2 \equiv 0 \tag{6-62}$$

For the case of n variables and one constraint, the derivations proceed along similar lines, producing

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\mathcal{L}_{ij}h_ih_j < 0 \tag{6-63}$$

for all h_i, h_j such that

$$\sum_{i=1}^{n} g_ih_i \equiv 0 \tag{6-64}$$

In this case the matrix of terms $(\mathcal{L}_{ij})$ is said to be *negative definite subject to constraint.*

Equations (6-61) and (6-62) can be combined into one useful expression: From (6-62),

$$h_2 = -h_1 \frac{g_1}{g_2}$$

Substituting this into (6-61) yields

$$\mathcal{L}_{11}h_1^2 + 2\mathcal{L}_{12}h_1\left(-h_1\frac{g_1}{g_2}\right) + \mathcal{L}_{22}\left(-h_1\frac{g_1}{g_2}\right)^2 < 0$$

Or, by multiplying by g_2^2,

$$(\mathcal{L}_{11}g_2^2 - 2\mathcal{L}_{12}g_1g_2 + \mathcal{L}_{22}g_1^2)h_1^2 < 0 \tag{6-65}$$

for any value of $h_1 \neq 0$. This implies that the expression in the parentheses must itself be < 0. How can that expression be conveniently remembered? It turns out, fortuitously, that the expression in parentheses in (6-65) is precisely the negative of the determinant

$$H = \begin{vmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} & g_1 \\ \mathcal{L}_{21} & \mathcal{L}_{22} & g_2 \\ g_1 & g_2 & 0 \end{vmatrix} \tag{6-66}$$

as can be immediately verified by expansion of H. *Hence, a sufficient condition for $f(x_1, x_2)$ to have a maximum subject to $g(x_1, x_2) \equiv 0$ is, together with the first-order relations, that $H > 0$.* Likewise, for a minimum subject to constraint, the sufficient second-order condition is that $H < 0$. Also, $H = 0$ corresponds to the case where the second derivatives $d^2y/dt^2 = 0$; hence no statement can be made regarding the type of stationary value in question. Note that $\partial^2\mathcal{L}/\partial x_1 \partial\lambda \equiv \mathcal{L}_{1\lambda} \equiv g_1 \equiv \mathcal{L}_{\lambda 1}$ and $\mathcal{L}_{2\lambda} \equiv \mathcal{L}_{\lambda 2} \equiv g_2$, and $\mathcal{L}_{\lambda\lambda} = 0$, since λ enters the Lagrangian $\mathcal{L} = f + \lambda g$ linearly. Hence, H is simply the determinant of the matrix of cross-partials of $\mathcal{L}$ with respect to x_1, x_2, and λ, that is,

$$H = \begin{vmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} & \mathcal{L}_{1\lambda} \\ \mathcal{L}_{21} & \mathcal{L}_{22} & \mathcal{L}_{2\lambda} \\ \mathcal{L}_{\lambda 1} & \mathcal{L}_{\lambda 2} & \mathcal{L}_{\lambda\lambda} \end{vmatrix}$$

For the n-variable case, the situation is more complicated, but the rules are analogous to the unconstrained case. The Lagrangian is $\mathcal{L} = f(x_1, \ldots, x_n) + \lambda g(x_1, \ldots, x_n)$. Consider the matrix of cross-partials of $\mathcal{L}$ with respect to $x_1, \ldots, x_n$ *and* λ, noting, as before, that $\mathcal{L}_{i\lambda} = g_i$, $\mathcal{L}_{\lambda\lambda} = 0$:

$$\mathbf{H} = \begin{pmatrix} \mathcal{L}_{11} & \cdots & \mathcal{L}_{1n} & g_1 \\ \vdots & & & \\ \mathcal{L}_{n1} & & \mathcal{L}_{nn} & g_n \\ g_1 & & g_n & 0 \end{pmatrix}$$

This matrix is commonly referred to as a *bordered Hessian* matrix, noting how the first partials of the constraint function g *border* the cross-partials of $\mathscr{L}$ with respect to $x_1, \ldots, x_n$.

Consider the following construction: By a "*border-preserving* principal minor of order k" of the preceding matrix, we mean that determinant which remains when any $n - k$ rows and the same numbered columns are deleted, *with the special added proviso that the border itself not be deleted.* Hence, the deletions that can occur must only come from rows 1 through n, not row or column $n + 1$. [Note that a border-preserving principal minor of order k is a $(k + 1) \times (k + 1)$ determinant.]

The second-order sufficient conditions are then:

Theorem. Together with the first-order conditions $\mathscr{L}_i = 0$, $i = 1, \ldots, n$ and $\mathscr{L}_\lambda = g = 0$, if all the border-preserving principal minors of H of order k have sign $(-1)^k$, $k = 2, \ldots, n$, then a maximum position is obtained. If all the border-preserving principal minors are negative, $k = 2, \ldots, n$, then a minimum is obtained.†

Suppose, even more generally, that there are r constraints involved. The Lagrangian function is $\mathscr{L} = f(x_1, \ldots, x_n) + \sum_{j=1}^{r} \lambda_j g^j(x_1, \ldots, x_n)$. The bordered Hessian matrix of this Langrangian is

$$\mathbf{H} = \begin{pmatrix} \mathscr{L}_{11} & \cdots & \mathscr{L}_{1n} & g_1^1 & \cdots & g_1^r \\ \vdots & & & & & \\ \mathscr{L}_{n1} & & \mathscr{L}_{nn} & g_n^1 & & g_n^r \\ g_1^1 & & g_n^1 & 0 & & 0 \\ \vdots & & & & & \\ g_1^r & & g_n^r & 0 & & 0 \end{pmatrix}$$

The sufficient conditions here state that for a *minimum,* the border-preserving principal minors of order $k > r$ (which again must involve deletions only from rows 1 through n) have sign $(-1)^r$, where r is the number of (independent) constraints. For a *maximum,* the border-preserving principal minors of order $k > r$ alternate in sign, beginning with $(-1)^{r+1}$, the second of opposite sign, etc. These principal minors must be of order greater than r, because, as inspection of H reveals (note the $r \times r$ matrix of 0s in the lower right), a determinant involving fewer than r rows and columns from rows and columns 1 through n must equal 0. Note again that with r bordering rows, a border-preserving principal minor of order k has $k + r$ rows and columns. An alternative presentation of the second-order conditions is given in Table 6-1. In this table, $m \geq 2r + 1$ is the size of the whole determinant.

†In fact, if only the "naturally ordered" principal minors have this property, then *all* of the border-preserving principal minors have that property.

TABLE 6-1
Second-order conditions: Sign of all size $m \times m$ (border-preserving) principal minors

	Constraints		
Condition	**0**	**1**	***r***
Maximum	$(-1)^m$ $m = 1, \ldots, n$	$(-1)^{m-1}$ $m = 3, \ldots, n+1$	$(-1)^{m-r}$ $m = 1+2r, \ldots, n+r$
Minimum	$(-1)^0 = +1$	$(-1)^1 = -1$	$(-1)^r$

The Geometry of Constrained Maximization

We visualize an unconstrained maximization in three dimensions as the top of a hill; the surface must be concave there. Constrained maxima (or minima) are somewhat more subtle. Consider the problem in two variables:

maximize

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

subject to

$$g(x_1, x_2) = k$$

The constraint $g(x_1, x_2) = k$ represents a curve in the x_1x_2 plane; we typically think of it as a "frontier," i.e., some sort of boundary that constrains consumption or production. Assume that the first partials g_1 and g_2 are positive so that the frontier has a negative slope $(-g_1/g_2)$, and increases in k move the frontier "northeast" in the x_1x_2 plane. Three such frontiers are represented in Fig. 6-2: in panel *(a)*, the frontier is concave, in panel *(b)* it is linear, and in panel *(c)* it is convex.

Assume that the first partials of $f(x_1, x_2)$ are also positive so that the level curves of f are likewise negatively sloping $(-f_1/f_2)$, and increasing values of f are associated with level curves that are increasingly distant from the origin. It is visually obvious that if the constrained maximum occurs at some interior point along the frontier (i.e., not at a corner, where the constraint intersects an axis), the maximum occurs where a level curve of $f(x_1, x_2)$ is tangent to the frontier. This is the algebraic condition $-f_1/f_2 = -g_1/g_2$, derived earlier. However, this tangency condition is implied by both a maximum and a minimum. If this condition is to represent a maximum, the level curves of the objective function must be either less concave than the constraint frontier, as shown in panel *(a)*, or more convex than the frontier, as shown in panels *(b)* and *(c)*.

If the constraint is linear, the level curves must appear "convex to the origin," the classic shape attributed to consumer's indifference curves and production isoquants. However, this characterization is in fact imprecise. The essential property required of the objective function to guarantee a constrained maximum subject to a linear constraint is that $f(x_1, x_2)$ be strictly increasing and *quasi-concave*. This latter characteristic is defined as follows.

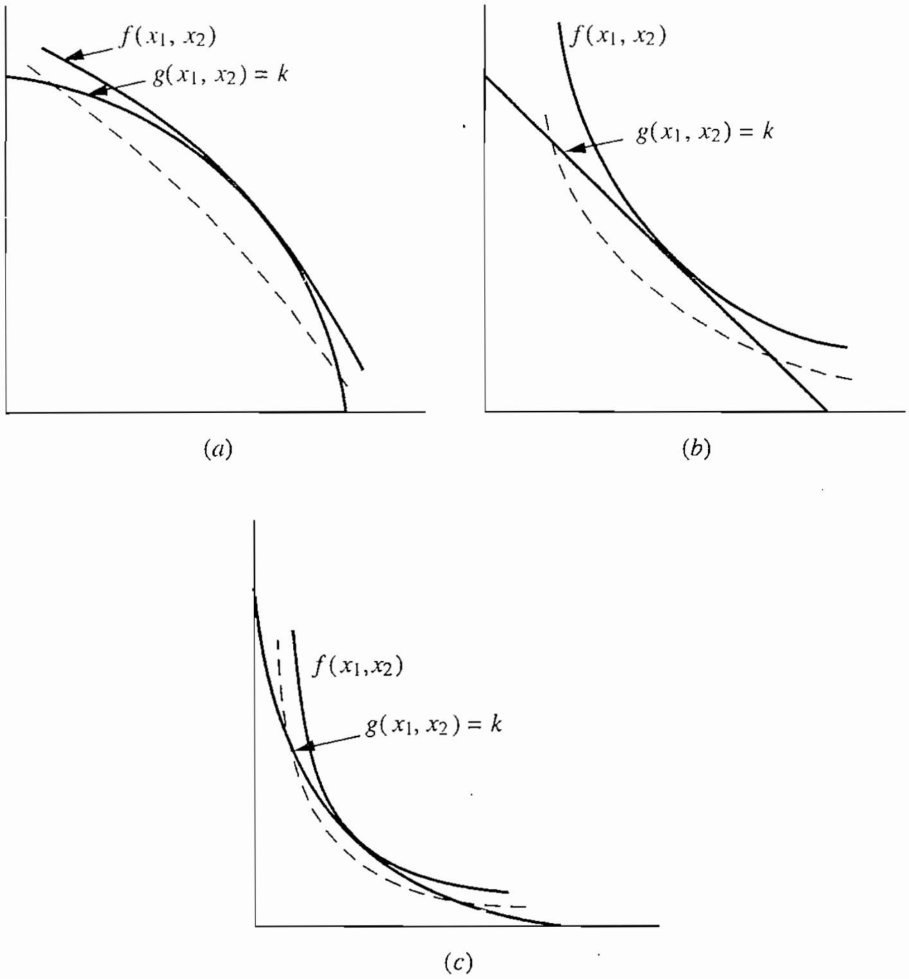

FIGURE 6-2
Constrained Maximization. Constrained maximization requires, for increasing functions, that the level curves of the objective function be either less concave or more convex than the level curves of the constraint. If the constraint is linear, as in panel (*b*), or convex, as in (*c*), the level curves of the objective function must be "convex to the origin"; i.e., the objective function must be quasi-concave.

Consider a typical indifference curve U^0 as shown in Fig. 6-3. Consider the set, call it S, of points that are at least as preferred as a point on U^0, shown as the shaded area. This set has the property that if any two points in S are connected by a straight line, the entire line also lies in S. A set with this property is called a *convex set* (not to be confused with a convex function). (As an example of a set that is not convex, consider the set of consumption bundles that are *less* preferred than those on U^0.) Algebraically, if $\mathbf{x}^0 = (x_1^0, x_2^0)$ and $\mathbf{x}^1 = (x_1^1, x_2^1)$ are any two points in the x_1x_2 plane, $\mathbf{x}^t = t\mathbf{x}^0 + (1-t)\mathbf{x}^1$, $0 \le t \le 1$ represents all points on the straight line joining $\mathbf{x}^0$ and $\mathbf{x}^1$. *A function is called quasi-concave if the set of points for which the function takes on values greater than or equal to some arbitrary value comprises a convex set.* That is, $U(x_1, x_2)$ is quasi-concave if $U(\mathbf{x}^1) \ge U(\mathbf{x}^0)$ implies $U(t\mathbf{x}^0 + (1-t)\mathbf{x}^1) \ge U(\mathbf{x}^0)$, $0 \le t \le 1$. (The definition is generalized in an obvious way for functions of n variables.) We note in passing that if the function *decreases* as the distance from the origin increases, quasi-concavity produces level curves that are "concave to the origin."

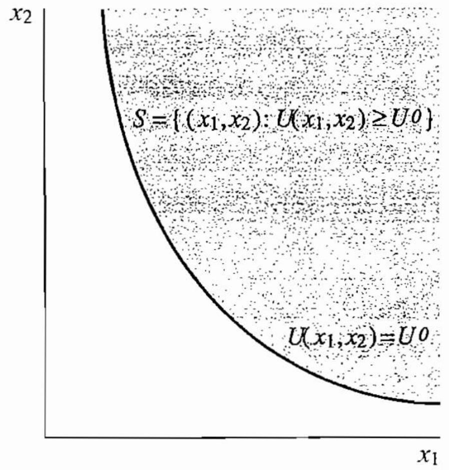

FIGURE 6-3
Quasi-Concavity. A function is said to be quasi-concave if the set of points for which the function takes on values greater than or equal to some arbitrary amount, say, U^0, is a convex set. These points are represented by the shaded area. This is the property generally assumed for utility and production functions.

Recall from Chap. 2 that a concave function is one for which $f(t\mathbf{x}^0 + (1-t)\mathbf{x}^1) \geq tf(\mathbf{x}^0) + (1-t)f(\mathbf{x}^1), 0 \leq t \leq 1$. Concavity clearly implies quasi-concavity: assuming $f(\mathbf{x}^1) \geq f(\mathbf{x}^0)$, $f(t\mathbf{x}^0 + (1-t)\mathbf{x}^1) \geq tf(\mathbf{x}^0) + (1-t)f(\mathbf{x}^1) \geq tf(\mathbf{x}^0) + (1-t)f(\mathbf{x}^0) = f(\mathbf{x}^0)$. The converse, however, is not true. Quasi-concavity is a weaker restriction than concavity. Concavity is required for an unconstrained maximum; quasi-concavity is all that is required for maximization subject to a linear constraint. In the preceding theorem, the second-order conditions dealing with the signs of the border-preserving principal minors define algebraically the geometric properties of the objective and constraint functions required for a constrained maximum (or minimum). If the constraint is linear, these second-order conditions for a maximum can be used to define algebraically the property of quasi-concavity of the objective function. (This requires the additional step of using the first-order conditions to replace the first partials of g with those of f in the bordering row and column.) If a linear objective function is *minimized* subject to constraint, these second-order conditions likewise describe quasi-concavity of the constraint function. This situation is encountered in Chap. 8, dealing with the minimization of cost subject to an output constraint. These concepts will be applied in the following chapters. Lastly, it is true, but not easy to prove, that if a function $f(\mathbf{x})$, $\mathbf{x} = (x_1, \ldots, x_n)$ is quasi-concave and linear homogeneous, it is (weakly) concave. Also, if f is strictly quasi-concave and homogeneous of degree r, $0 < r < 1$, it is strictly concave. The proofs are left as exercises.

Example. Consider again the basic consumer theory model, maximize $U(x_1, x_2)$ subject to $p_1x_1 + p_2x_2 = M$. (See Fig. 6-1 again.) Assuming more is preferred to less, the ordinal indifference levels must be indexed such that $U^2 > U^1 > U^0$. The condition that a point of tangency of an indifference curve and the budget constraint actually represents a maximum rather than a minimum of utility subject to a linear budget constraint is that the utility function be strictly increasing and quasi-concave. In this two-variable model, these conditions imply the usual shape, "convex to the origin." These assumptions compose the law of diminishing marginal rate of substitution, i.e., in two dimensions, that the slope of the level (indifference) curve increases (becomes less negative) as x_1 increases. We showed in Chap. 3 that the algebraic expression of this shape is [see Sec. 3.5, Eq. (3-24)]

$$\frac{d^2x_2}{dx_1^2} \equiv \left[-U_2^2U_{11} + 2U_{12}U_1U_2 - U_1^2U_{22} \right] \frac{1}{U_2^3} > 0 \tag{6-67}$$

If this is to be positive, the square-bracketed term must be positive, assuming that $U_2 > 0$, i.e., the consumer is not sated in good 2. But by inspection, the term in brackets is equal to the following determinant, which must therefore itself be positive:

$$H' = \begin{vmatrix} U_{11} & U_{12} & -U_1 \\ U_{21} & U_{22} & -U_2 \\ -U_1 & -U_2 & 0 \end{vmatrix} > 0 \tag{6-68}$$

However, from the first-order conditions for utility maximization (6-45), $U_1 = \lambda p_1$, $U_2 = \lambda p_2$. Substituting this into H' and then dividing the last row and column by λ (and, hence, H' by λ^2, which is positive), the condition $H' > 0$ is equivalent to

$$H = \begin{vmatrix} U_{11} & U_{12} & -p_1 \\ U_{21} & U_{22} & -p_2 \\ -p_1 & -p_2 & 0 \end{vmatrix} > 0 \tag{6-69}$$

But H is seen to be the determinant of the bordered Hessian matrix, the cross-partials of $\mathcal{L}$ with respect to x_1, x_2, and λ. This is in accordance with the general theorem of this section.

6.6 GENERAL METHODOLOGY

At the beginning of this chapter, we considered the general economic model that was characterized by being an *unconstrained* maximization. Let us now explore models that have a constraint as an added feature.

Consider some economic agent that behaves in accordance with the following general model:

maximize

$$f(x_1, x_2, \alpha) = y \tag{6-70}$$

subject to

$$g(x_1, x_2, \alpha) = 0 \tag{6-71}$$

where x_1 and x_2 are the decision variables and α is some parameter (or vector of parameters) over which the agent has no control. What will be the response to autonomous changes in the environment, i.e., to changes in the parameter α?

The first-order conditions for a maximum are derived by setting the partials of the Lagrangian function $\mathcal{L} = f(x_1, x_2, \alpha) + \lambda g(x_1, x_2, \alpha)$, with respect to x_1, x_2, and λ, equal to zero:

$$\begin{aligned} \mathcal{L}_1 &= f_1(x_1, x_2, \alpha) + \lambda g_1(x_1, x_2, \alpha) = 0 \\ \mathcal{L}_2 &= f_2(x_1, x_2, \alpha) + \lambda g_2(x_1, x_2, \alpha) = 0 \\ \mathcal{L}_\lambda &= g(x_1, x_2, \alpha) = 0 \end{aligned} \tag{6-72}$$

Equations (6-72) represent three equations in the four unknowns x_1, x_2, λ, and α. Assuming the implicit function theorem (as was discussed previously) is applicable, these equations can be solved, in principle at least, for the choice functions

$$\begin{aligned} x_1 &= x_1^*(\alpha) \\ x_2 &= x_2^*(\alpha) \\ \lambda &= \lambda^*(\alpha) \end{aligned} \tag{6-73}$$

Substituting these values back into Eqs. (6-72) from which they were derived yields the *identities*

$$\begin{aligned} f_1(x_1^*, x_2^*, \alpha) + \lambda^* g_1(x_1^*, x_2^*, \alpha) &\equiv 0 \\ f_2(x_1^*, x_2^*, \alpha) + \lambda^* g_2(x_1^*, x_2^*, \alpha) &\equiv 0 \\ g(x_1^*, x_2^*, \alpha) &\equiv 0 \end{aligned} \tag{6-74}$$

Since we are interested in *changes* in the x_i^*'s (i.e., marginal values) as α changes, we differentiate (6-74) with respect to α, using the chain rule. The first equation then yields

$$f_{11}\frac{\partial x_1^*}{\partial \alpha} + f_{12}\frac{\partial x_2^*}{\partial \alpha} + f_{1\alpha} + \lambda^* g_{11}\frac{\partial x_1^*}{\partial \alpha} + \lambda^* g_{12}\frac{\partial x_2^*}{\partial \alpha} + \lambda^* g_{1\alpha} + g_1\frac{\partial \lambda^*}{\partial \alpha} \equiv 0$$

Noting that $\mathscr{L}_{11} = f_{11} + \lambda^* g_{11}, \ldots$, this equation can be more conveniently written

$$\mathscr{L}_{11}\frac{\partial x_1^*}{\partial \alpha} + \mathscr{L}_{12}\frac{\partial x_2^*}{\partial \alpha} + g_1\frac{\partial \lambda^*}{\partial \alpha} \equiv -\mathscr{L}_{1\alpha} \tag{6-75}$$

Similarly, differentiating the second and third equations of (6-74) yields

$$\mathscr{L}_{21}\frac{\partial x_1^*}{\partial \alpha} + \mathscr{L}_{22}\frac{\partial x_2^*}{\partial \alpha} + g_2\frac{\partial \lambda^*}{\partial \alpha} \equiv -\mathscr{L}_{2\alpha} \tag{6-76}$$

$$g_1\frac{\partial x_1^*}{\partial \alpha} + g_2\frac{\partial x_2^*}{\partial \alpha} \equiv -g_\alpha \tag{6-77}$$

In matrix notation, this system of three linear equations can be written

$$\begin{pmatrix} \mathscr{L}_{11} & \mathscr{L}_{12} & g_1 \\ \mathscr{L}_{21} & \mathscr{L}_{22} & g_2 \\ g_1 & g_2 & 0 \end{pmatrix} \begin{pmatrix} \dfrac{\partial x_1^*}{\partial \alpha} \\ \dfrac{\partial x_2^*}{\partial \alpha} \\ \dfrac{\partial \lambda^*}{\partial \alpha} \end{pmatrix} = \begin{pmatrix} -\mathscr{L}_{1\alpha} \\ -\mathscr{L}_{2\alpha} \\ -g_\alpha \end{pmatrix} \tag{6-78}$$

Notice that the coefficient matrix on the left of (6-78) is the matrix of second partials of the Lagrangian function. In unconstrained maximization models, this coefficient matrix was the matrix of second partials of the objective function. The manipulation of the model is formally identical in the constrained and unconstrained cases; the only difference is the conditions imposed on the principal minors of the coefficient matrix by the sufficient second-order conditions.

The reason why the coefficient matrix comes out to be the second partials of $\mathcal{L}$ is that identities (6-74) are precisely the first partials of $\mathcal{L}$,

$$\begin{aligned} \mathcal{L}_1(x_1^*, x_2^*, \lambda^*, \alpha) &\equiv 0 \\ \mathcal{L}_2(x_1^*, x_2^*, \lambda^*, \alpha) &\equiv 0 \\ \mathcal{L}_\lambda(x_1^*, x_2^*, \alpha) &\equiv 0 \end{aligned} \tag{6-79}$$

[Notice that λ^* does not appear in $\mathcal{L}_\lambda = g(x_1^*, x_2^*, \alpha) \equiv 0$.] Differentiating the first identity with respect to α yields

$$\mathcal{L}_{11}\frac{\partial x_1^*}{\partial \alpha} + \mathcal{L}_{12}\frac{\partial x_2^*}{\partial \alpha} + \mathcal{L}_{1\lambda}\frac{\partial \lambda^*}{\partial \alpha} + \mathcal{L}_{1\alpha} \equiv 0$$

This is precisely Eq. (6-75), noting again that $\mathcal{L}_{1\lambda} = g_1$. In like fashion, Eqs. (6-76) and (6-77) are derivable directly from $\mathcal{L}_2 \equiv 0$, $\mathcal{L}_\lambda \equiv 0$.

Since the Jacobian determinant J needed to ensure solution of Eqs. (6-72) for the explicit choice functions (6-73) is formed from the matrix of first partials of (6-72), J is in fact the determinant of second partials of the Lagrangian $\mathcal{L}$ with respect to x_1, x_2, and λ, that is, the determinant of the coefficient matrix in (6-78). This determinant is denoted by H below. The *sufficient* second-order conditions imply, among other things, that this determinant is nonzero, and thus the explicit relations (6-73) are valid. And this determinant forms the denominator in the solution by Cramer's rule for $\partial x_i^*/\partial \alpha$ and $\partial \lambda^*/\partial \alpha$. Let us now proceed, in the same manner as for the unconstrained models.

Solving for $\partial x_1^*/\partial \alpha$ by Cramer's rule,

$$\frac{\partial x_1^*}{\partial \alpha} = \frac{\begin{vmatrix} -\mathcal{L}_{1\alpha} & \mathcal{L}_{12} & g_1 \\ -\mathcal{L}_{2\alpha} & \mathcal{L}_{22} & g_2 \\ -g_\alpha & g_2 & 0 \end{vmatrix}}{H} = \frac{-\mathcal{L}_{1\alpha}H_{11}}{H} - \frac{\mathcal{L}_{2\alpha}H_{21}}{H} - \frac{g_\alpha H_{31}}{H} \tag{6-80}$$

where H is the bordered Hessian determinant of the coefficient matrix. Solutions for $\partial x_2^*/\partial \alpha$ and $\partial \lambda^*/\partial \alpha$ are, likewise,

$$\frac{\partial x_2^*}{\partial \alpha} = \frac{\begin{vmatrix} \mathcal{L}_{11} & -\mathcal{L}_{1\alpha} & g_1 \\ \mathcal{L}_{21} & -\mathcal{L}_{2\alpha} & g_2 \\ g_1 & -g_\alpha & 0 \end{vmatrix}}{H} = \frac{-\mathcal{L}_{1\alpha}H_{12}}{H} - \frac{\mathcal{L}_{2\alpha}H_{22}}{H} - \frac{g_\alpha H_{32}}{H} \tag{6-81}$$

$$\frac{\partial \lambda^*}{\partial \alpha} = \frac{\begin{vmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} & -\mathcal{L}_{1\alpha} \\ \mathcal{L}_{21} & \mathcal{L}_{22} & -\mathcal{L}_{2\alpha} \\ g_1 & g_2 & -g_\alpha \end{vmatrix}}{H} = \frac{-\mathcal{L}_{1\alpha}H_{13}}{H} - \frac{\mathcal{L}_{2\alpha}H_{23}}{H} - \frac{g_\alpha H_{33}}{H} \tag{6-82}$$

It is clear that at this level of generality, no prediction as to the sign of $\partial x_i^*/\partial \alpha$

or $\partial\lambda^*/\partial\alpha$ is forthcoming. There simply is not enough information in the system. All we know is that the denominators in these expressions are positive, but we have no information regarding the numerators. The signs of the off-diagonal cofactors are not implied by the maximum conditions.

Suppose now that the parameter α did not appear in either the second or third first-order relations (6-72). Then $\mathcal{L}_{2\alpha} \equiv 0$ and $g_\alpha \equiv 0$, and

$$\frac{\partial x_1^*}{\partial \alpha} = \frac{-\mathcal{L}_{1\alpha}\begin{vmatrix}\mathcal{L}_{22} & g_2 \\ g_2 & 0\end{vmatrix}}{H} = \frac{+\mathcal{L}_{1\alpha}g_2^2}{H} \tag{6-83}$$

The partial $\partial x_1^*/\partial\alpha$ now has a predictable sign: Since $H > 0$ and $H_{11} < 0$, by the second-order conditions (here, $H_{11} = -g_2^2 < 0$ always), $\partial x_1^*/\partial\alpha$ will have the same sign as the direction of "disturbance" of the first equation. That is, if an increase in α has the effect of shifting the marginal curve $\mathcal{L}_1$ to the right ($\mathcal{L}_{1\alpha} > 0$), then the response will be to increase the utilization of x_1. Hence, if it is possible to make statements like, "an increase in income will *shift* a demand curve to the right," or "a change in technology will lower (shift down) such and such marginal cost curve," then if that income or technology parameter enters only one first-order relation, it will in general be possible to predict the direction of change of the associated variable (the one for which that first-order equation is the first partial of the Lagrangian). More succinctly, if α enters the ith first-order equation only, then $\partial x_i^*/\partial\alpha$ and $\mathcal{L}_{i\alpha}$ have the same sign. However, since $g_\alpha \equiv 0$, $\mathcal{L}_{i\alpha} = f_{i\alpha} + \lambda g_{i\alpha} = f_{i\alpha}$, and thus, just as in the case of maximization models without constraints, $\partial x_i/\partial\alpha$ and $f_{i\alpha}$ must have the same sign, or

$$f_{i\alpha}\frac{\partial x_i}{\partial \alpha} > 0 \tag{6-84}$$

This result holds for the case of n variables as well as for just two variables; its precise statement is given in the problems following. The result follows because of the conditions on the principal minors imposed by the second-order conditions for a constrained maximum.

In the case of $\partial\lambda^*/\partial\alpha$, however, a sign is *never* implied by the sufficient second-order conditions alone, no matter how the parameter α enters the first-order equations. Suppose, for example, α enters only the constraint, i.e., the third first-order equation. Then $-\mathcal{L}_{1\alpha} = -\mathcal{L}_{2\alpha} = 0$, and

$$\frac{\partial \lambda^*}{\partial \alpha} = \frac{-g_\alpha H_{33}}{H} = \frac{-g_\alpha}{H}(\mathcal{L}_{11}\mathcal{L}_{22} - \mathcal{L}_{12}^2) \gtrless 0 \tag{6-85}$$

The cofactor H_{33}, while a principal minor, is not a *border-preserving* principal minor. The border row and column of H are deleted when forming H_{33}. Hence, no sign is implied for $\partial\lambda^*/\partial\alpha$. If α enters any of the other equations, then the off-diagonal cofactors H_{31} and H_{32} will enter the expressions. These expressions are likewise not signed by the maximum conditions.

For the same reasons, it is apparent that any time the parameter α enters the constraint, off-diagonal cofactors will be present in the expressions for $\partial x_i/\partial\alpha$. Thus no refutable implications are possible in models for a parameter that appears in the constraint.

Example. To illustrate the principles just developed, let us return to the profit maximization model, slightly modified. Consider a firm with production $y = f(x_1, x_2)$ selling output y at price p. The firm hires input x_1 at wage w_1; x_2, however, represents the entrepreneur's input and is fixed at some level x_2^0. The firm seeks to maximize net rents R, the difference between total revenue and the total factor cost of x_1. Algebraically, the model is

$$\underset{x_1, x_2}{\text{maximize}} \quad R = pf(x_1, x_2) - w_1x_1$$

subject to

$$x_2 = x_2^0$$

Although we have essentially solved this model in Chap. 4, by directly substituting the constraint into the objective function, we shall analyze it here as a constrained maximization model. Even though in this particular example the constraint says that x_2 is fixed, we treat x_2 as a variable, maintaining the structure of the Lagrangian analysis.

Using the Lagrangian

$$\mathcal{L} = pf(x_1, x_2) - w_1x_1 + \lambda\left(x_2^0 - x_2\right)$$

the first-order conditions are

$$\mathcal{L}_1 = pf_1(x_1, x_2) - w_1 = 0 \tag{6-86a}$$

$$\mathcal{L}_2 = pf_2(x_1, x_2) - \lambda = 0 \tag{6-86b}$$

$$\mathcal{L}_\lambda = x_2^0 - x_2 = 0 \tag{6-86c}$$

Equation (6-86*a*) says that the firm will hire x_1 until the value of its marginal product of that factor equals its wage, as previously derived. Equation (6-86*b*) identifies the Lagrange multiplier λ as the value of the marginal product of the entrepreneurial input. Whereas the wage of factor 1 is exogenously set by the competitive labor market, the wage of factor 2 is endogenously "imputed." If a competitive market existed for entrepreneurial services, another firm would be willing to pay λ for this owner's services.

The sufficient second-order condition is that the bordered Hessian determinant formed from the second partials of $\mathcal{L}$ is positive:

$$H = \begin{vmatrix} pf_{11} & pf_{12} & 0 \\ pf_{21} & pf_{22} & -1 \\ 0 & -1 & 0 \end{vmatrix} > 0 \tag{6-87}$$

Evaluating this determinant (say, by the third row, which has two zeros in it) yields $pf_{11} < 0$. Note that no restriction is placed on f_{22}; since only x_1 is really variable (even though we treat x_2 as variable in the constrained model), the only margin on which the firm adjusts is how much x_1 to hire. Only diminishing marginal product of x_1 is thus required for an interior maximum.

Assuming the sufficient second-order condition holds, the first-order equations can be solved simultaneously for the explicit choice functions:

$$
\begin{aligned}
x_1 &= x_1^*\left(w_1, p, x_2^0\right) \\
x_2 &= x_2^*\left(w_1, p, x_2^0\right) \\
\lambda &= \lambda^*\left(w_1, p, x_2^0\right)
\end{aligned}
\tag{6-88}
$$

These choice functions represent the factor demands for x_1 and x_2 (trivial, in the case of x_2; $x_2^* = x_2^0$) and the profit-maximizing imputed value of entrepreneurial input.

Multiplying Eq. (6-86*a*) by x_1^*, (6-86*b*) by x_2^*, and adding,

$$
p(f_1 x_1^* + f_2 x_2^*) \equiv w_1 x_1^* + \lambda^* x_2^* \tag{6-89}
$$

If the production function is homogeneous of degree 1 (constant returns to scale), then from Euler's theorem, the left-hand side of this identity is py^*. In that case, (6-89) can be interpreted as Total Revenue = Total Cost, where the total factor cost of x_2 is its imputed opportunity cost $\lambda^* x_2^*$. Thus with constant returns to scale, the product is "exhausted"; i.e., the revenue received by the firm is exactly accounted for by the total factor cost. Incidentally, (6-89) is an identity in w_1, x_2^0, and p, not in x_1 and x_2. This relation holds only for values of the factors satisfying the first-order equations, assuming the sufficient second-order conditions are also satisfied.

Let us now investigate the comparative statics of this model. Note that the parameter w_1 enters only the objective function, whereas x_2^0 enters the constraint. Substituting the solutions (6-88) back into the first-order equations yields the identities

$$
\begin{aligned}
pf_1(x_1^*, x_2^*) - w_1 &\equiv 0 \\
pf_2(x_1^*, x_2^*) - \lambda^* &\equiv 0 \\
x_2^0 - x_2^* &\equiv 0
\end{aligned}
\tag{6-90}
$$

Since the parameter w_1 enters only the first first-order equation, we expect therefore to be able to derive a refutable implication for this parameter. The parameter x_2^0, on the other hand, appears in the constraint; we expect no refutable implication for this parameter. Differentiating these identities first with respect to w_1 produces the matrix equation

$$
\begin{pmatrix}
pf_{11} & pf_{12} & 0 \\
pf_{21} & pf_{22} & -1 \\
0 & -1 & 0
\end{pmatrix}
\begin{pmatrix}
\dfrac{\partial x_1^*}{\partial w_1} \\[2ex]
\dfrac{\partial x_2^*}{\partial w_1} \\[2ex]
\dfrac{\partial \lambda^*}{\partial w_1}
\end{pmatrix}
=
\begin{pmatrix}
1 \\ 0 \\ 0
\end{pmatrix}
\tag{6-91}
$$

Solving for $\partial x_1^*/\partial w_1$,

$$
\frac{\partial x_1^*}{\partial w_1} = \frac{H_{11}}{H} = \frac{-1}{-pf_{11}} < 0 \tag{6-92a}
$$

Also, as expected, since x_2 is fixed,

$$\frac{\partial x_2^*}{\partial w_1} = \frac{H_{12}}{H} = \frac{0}{-pf_{11}} = 0 \tag{6-92b}$$

and

$$\frac{\partial \lambda^*}{\partial w_1} = \frac{H_{13}}{H} = \frac{-pf_{21}}{-pf_{11}} \gtrless 0 \tag{6-92c}$$

As we showed earlier, a sign is never implied for rates of change of the Lagrange multiplier with respect to any parameter. However, Eq. (6-92c) shows that if the marginal product of x_1 increases with an increase in the entrepreneurial input (meaning, in the two-factor case, that the two factors are complements), the imputed marginal value of the entrepreneurial input moves in the opposite direction as the wage of x_1. (If elevators are fixed in supply, an increase in the wages of elevator operators will lower the imputed marginal value of elevators.)

Differentiating Eqs. (6-90) with respect to x_2^0 produces the matrix equation

$$\begin{pmatrix} pf_{11} & pf_{12} & 0 \\ pf_{21} & pf_{22} & -1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} \dfrac{\partial x_1^*}{\partial x_2^0} \\ \dfrac{\partial x_2^*}{\partial x_2^0} \\ \dfrac{\partial \lambda^*}{\partial x_2^0} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \tag{6-93}$$

Solving,

$$\frac{\partial x_1^*}{\partial x_2^0} = \frac{-H_{31}}{H} = \frac{pf_{12}}{-pf_{11}} \gtrless 0 \tag{6-94a}$$

Also, since $x_2 = x_2^0$,

$$\frac{\partial x_2^*}{\partial x_2^0} = \frac{-H_{32}}{H} = \frac{-pf_{11}}{-pf_{11}} = 1 \tag{6-94b}$$

and

$$\frac{\partial \lambda^*}{\partial x_2^0} = \frac{-H_{33}}{H} = \frac{-p^2\left(f_{11}f_{22} - f_{12}^2\right)}{-pf_{11}} \gtrless 0 \tag{6-94c}$$

Note the curious "reciprocity" result $\partial x_1^*/\partial x_2^0 = -\partial \lambda^*/\partial w_1$, since $H_{13} = H_{31}$. We shall have more to say about these types of relations in the next chapter. Note also from Eq. (6-94c) that an increase in the parametric entrepreneurial input level has an unpredictable effect on the imputed marginal value of the entrepreneurial input. Only if we assume, additionally, that the production function is concave so that $f_{11}f_{22} - f_{12}^2 > 0$ is $\partial \lambda^*/\partial x_2^0 < 0$. In that case, as in ordinary profit maximization, x_2 exhibits diminishing marginal product, lowering its marginal value as more x_2 is utilized. Thus, assuming information in addition to the maximization hypothesis leads to additional results.

To sum up, for parameters entering only the objective function, refutable implications are possible. Because such a parameter, w_1, enters one and only one first-order condition, a sign can be determined for $\partial x_1^*/\partial w_1$. For parameters entering the constraint, such as x_2^0 in this model, refutable implications are not possible on the basis of the maximization hypothesis alone, though additional assumptions may yield useful propositions.

PROBLEMS

1. Consider the constrained maximum problem

maximize
$$f(x_1, \ldots, x_n, \alpha_1, \ldots, \alpha_m) = y$$

subject to
$$g(x_1, \ldots, x_n, \alpha_1, \ldots, \alpha_m) = 0$$

Prove that if some parameter α_i enters the ith first-order relation and that equation only, then
$$\mathcal{L}_{i\alpha_i} \frac{\partial x_i^*}{\partial \alpha_i} > 0$$

2. Prove the same result if there is more than one constraint.
3. Show that diminishing marginal utility in each good neither implies nor is implied by convexity of the indifference curves.
4. Find the maximum or minimum values of the following functions $f(x_1, x_2)$ subject to the constraints $g(x_1, x_2) = 0$ by the method of direct substitution and by Lagrange multipliers. Be sure to check the second-order conditions to see if a maximum or minimum (if either) is achieved.
   (a) $f(x_1, x_2) = x_1x_2$; $g(x_1, x_2) = 2 - (x_1 + x_2)$.
   (b) $f(x_1, x_2) = x_1 + x_2$; $g(x_1, x_2) = 1 - x_1x_2$.
   (c) $f(x_1, x_2) = x_1x_2$; $g(x_1, x_2) = M - p_1x_1 - p_2x_2$, where p_1, p_2, and M are parameters.
   (d) $f(x_1, x_2) = p_1x_1 + p_2x_2$; $g(x_1, x_2) = U^0 - x_1x_2$.
5. Show that the second-order conditions for Probs. 4(*a*) and 4(*b*) are equivalent; also that the second-order conditions for Probs. 4(*c*) and 4(*d*) are equivalent.
6. Consider the class of models

maximize
$$y = f(x_1, x_2) + \alpha x_1$$

subject to
$$g(x_1, x_2) + \beta x_2 = 0$$

where x_1 and x_2 are choice variables and α and β are parameters. Using the Lagrangian
$$\mathcal{L} = f(x_1, x_2) + \alpha x_1 + \lambda(g(x_1, x_2) + \beta x_2)$$

   (a) Prove that $\partial x_1^*/\partial\alpha > 0$ but that no refutable comparative statics result is available for β.
   (b) Prove that $\partial x_1^*/\partial\beta = \lambda^*(\partial x_2^*/\partial\alpha) + x_2^*(\partial\lambda^*/\partial\alpha)$.

7. Consider a general maximization problem

maximize

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

subject to

$$g(x_1, x_2) = k$$

where x_1 and x_2 are choice variables, and α and k are parameters. Using the Lagrangian

$$\mathcal{L} = f(x_1, x_2, \alpha) + \lambda(k - g(x_1, x_2))$$

(*a*) Prove that $f_{1\alpha}(\partial x_1^*/\partial k) + f_{2\alpha}(\partial x_2^*/\partial k) = \partial\lambda^*/\partial\alpha$.

(*b*) What functional forms of the objective function and constraint would lead to the simple reciprocity result $\partial x_1^*/\partial k = \partial\lambda^*/\partial\alpha$?

8. Consider a firm that hires two inputs x_1 and x_2 at factor prices w_1 and w_2, respectively. If this firm is one of many identical firms, then in the long run, the profit-maximizing position will be at the minimum of its average cost (AC) curve. Analyze the comparative statics of this firm in the long run by asserting the behavioral postulate

minimize

$$\text{AC} = \frac{w_1 x_1 + w_2 x_2}{f(x_1, x_2)}$$

where $y = f(x_1, x_2)$ is the firm's production function.

(*a*) Show that the first-order necessary conditions for min AC are $w_i - \text{AC}^* f_i = 0$, $i = 1, 2$, where AC* is min AC. Interpret.

(*b*) Show that the sufficient second-order conditions for min AC are the same as for profit maximization in the short run (fixed-output price), that is,

$$f_{11} < 0 \quad f_{22} < 0 \quad f_{11}f_{22} - f_{12}^2 > 0$$

(*Hint:* In differentiating the product $\text{AC}^* f_i$, remember that $\partial\text{AC}^*/\partial x_i = 0$ by the first-order conditions.)

(*c*) Find all partials of the form $\partial x_i^*/\partial w_j$. (Remember that w_1 and w_2 appear in AC.) Show that $\partial x_i^*/\partial w_i < 0$ is *not* implied by this model, nor is $\partial x_i^*/\partial w_j = \partial x_j^*/\partial w_i$.

(*d*) Show that $f_1 x_1^* + f_2 x_2^* \equiv y^*$. Is this Euler's theorem? (If it is, you have just proved that all production functions are linear homogeneous!)

9. Consider a firm with the production function $y = f(x_1, x_2)$, which sells its output in a competitive output market at price p. It is, however, a monopsonist in the input market, i.e., it faces rising factor supply curves, in which the unit factor prices w_1 and w_2 rise with increasing factor usage, that is, $w_1 = k_1 x_1$, $w_2 = k_2 x_2$. The firm is asserted to be a profit maximizer.

(*a*) How might one represent algebraically a decrease in the supply of factor 1?

(*b*) If the supply of x_1 decreases, will the use of factor 1 decrease? Demonstrate.

(*c*) What will happen to the usage of factor 2 if the supply of x_1 decreases?

(*d*) Explain, in about one sentence, why factor demand curves for this firm do *not* exist.

(*e*) Suppose the government holds the firm's use of x_2 constant at the previous profit-maximizing level. If the supply of x_1 decreases, will the use of x_1 change by more or less, absolutely, than previously?

10. Prove the propositions stated at the end of Sec. 6.5 that if a function $f(\mathbf{x})$, $\mathbf{x} = (x_1, \ldots, x_n)$ is quasi-concave and linear homogeneous, it is (weakly) concave, and if f is strictly quasi-concave and homogeneous of degree r, $0 < r < 1$, it is strictly concave.

SELECTED REFERENCES

Allen, R. G. D.: *Mathematical Analysis for Economists,* Macmillan & Co., Ltd., London, 1938. Reprinted by St. Martin's Press, Inc., New York, 1967.

Apostol, T.: *Mathematical Analysis,* Addison-Wesley Publishing Company, Inc., Reading, MA, 1957.

Courant, R.: *Differential and Integral Calculus* (Trans.), Interscience Publishers, Inc., New York, 1947.

Hadley, G.: *Nonlinear and Dynamic Programming,* Addison-Wesley Publishing Company, Inc., Reading, MA, 1964.

Hancock, H.: *Theory of Maxima and Minima,* Ginn and Company, Boston, MA, 1917. Reprinted by Dover Publications, Inc., New York, 1960.

Panik, M. J.: *Classical Optimization: Foundations and Extensions,* North-Holland Publishing Company, Amsterdam, 1976.

Samuelson, P. A.: *Foundations of Economic Analysis,* Harvard University Press, Cambridge, MA, 1947.

CHAPTER

7

THE ENVELOPE THEOREM AND DUALITY

7.1 HISTORY OF THE PROBLEM

In the early 1930s, a very distinguished economist, Jacob Viner, was analyzing the behavior of firms in the short and long run. Viner defined the "short run" as a time period in which one factor of production, presumably capital, was fixed, while the other factor, labor, was variable. He posited a series of short-run cost curves, whose minimum points (for successively larger capital inputs) first fall and then rise. Viner reasoned that if both inputs were variable, the resulting "long-run" average cost would always be less than or equal to the corresponding short-run cost. He therefore concluded that the long-run average cost curve should be drawn as an "envelope" to all the short-run curves. The eventual diagram, pictured in Fig. 7-1, now appears in virtually all intermediate price theory texts.

However, Viner also was puzzled by the fact that the resulting long-run curve did not pass through the minimum points of the short-run curves, since reducing unit costs seemed to increase available profits. Moreover, at the points of tangency, the slopes of the long-run and short-run curves were the same, indicating that average cost was falling (or rising) at the same rate, irrespective of whether capital was being held constant. Viner therefore apparently asked his draftsman, Wong, to draw a long-run average cost curve that was both an envelope curve to the short-run curves and that also passed through the minimum points of the short-run curves. When Wong indicated the impossibility of this joint occurrence, Viner opted to draw the long-run average cost curve through the minimum points of the short-run average cost curves,

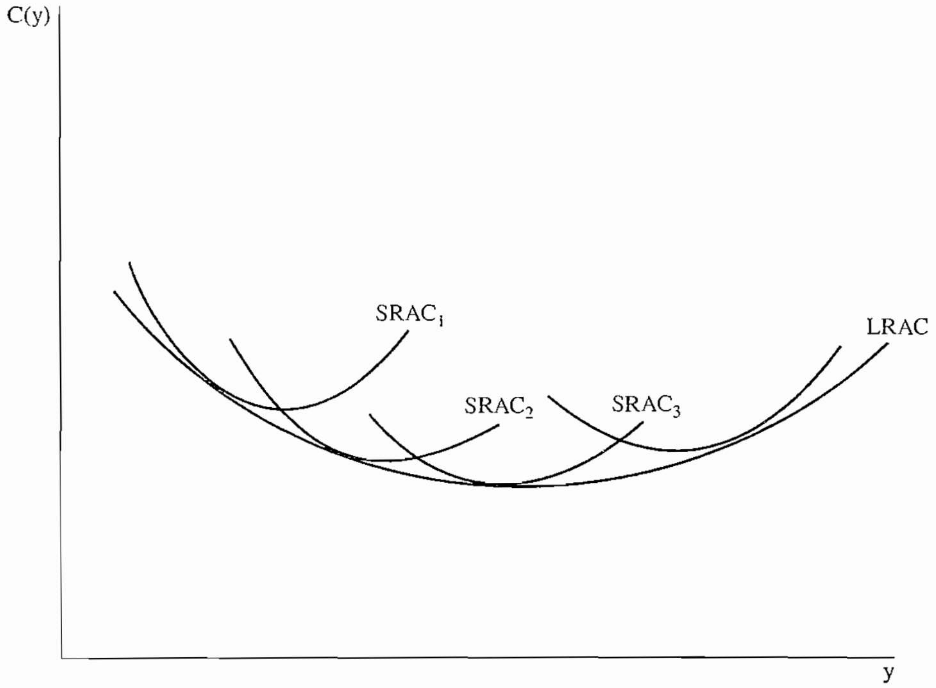

FIGURE 7-1
The modern Viner-Wong diagram shows the long-run average cost curve as an *envelope* to the short-run average cost curves.

rather than as an envelope curve.[†] The egos of many succeeding economists have been soothed by that decision.

The problem was soon analyzed algebraically by Paul Samuelson, who demonstrated the correctness of the tangency of such long- and short-run curves.[‡] However, it remained somewhat of a puzzle that the rate of change of an objective function should be the same whether or not one variable is held constant. Perhaps most surprising, as economists investigated this puzzle further, was the discovery that the relationships that underlie this "envelope theorem" also reveal the basic theorems about the existence of refutable comparative statics theorems. It is to this larger issue that we now turn.

7.2 THE PROFIT FUNCTION

Samuelson began his analysis as follows. Consider a general maximization model with two decision variables, x_1 and x_2, and one parameter, α:

[†]See Jacob Viner, "Cost Curves and Supply Curves," *Zeitschrift fur Nationalokonomie,* **3:**1931. Reprinted in *AEA Readings in Price Theory,* Irwin, Homewood, IL, 1952.

[‡]See Paul Samuelson, *Foundations of Economic Analysis,* Harvard University Press, Cambridge, MA, 1947.

maximize

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

(The generalization to n variables is trivial; we will later consider models with multiple parameters.) The first-order necessary conditions are, of course, $f_1 = f_2 = 0$; assuming the sufficient second-order conditions hold, the explicit choice functions $x_i = x_i^*(\alpha)$ are derived as the solutions to the first-order equations. If we now substitute these solutions *into the objective function,* we obtain the function

$$\phi(\alpha) = f(x_1^*(\alpha), x_2^*(\alpha), \alpha) \tag{7-1}$$

The function $\phi(\alpha)$ is the value of the objective function f when the x_i's that maximize f (for given α) are used. Therefore, $\phi(\alpha)$ represents the maximum value of f, for any specified α. *We call $\phi(\alpha)$ the indirect objective function.*

How does ϕ vary (as compared to f) when α varies? Differentiating with respect to α,

$$\phi_\alpha(\alpha) = f_1 \frac{\partial x_1^*}{\partial \alpha} + f_2 \frac{\partial x_2^*}{\partial \alpha} + f_\alpha$$

However, from the first-order conditions, $f_1 = f_2 = 0$; hence the first two terms on the right-hand side vanish. Therefore,

$$\phi_\alpha(\alpha) = f_\alpha \tag{7-2}$$

Equation (7-2) says that as α changes, the rate of change of *the maximum value of* f, where x_1 and x_2 vary optimally as α varies, equals the rate of change of f as α varies, holding x_1 and x_2 constant! This result has puzzled many economists long after the publication of Viner's original article.

Before we study the geometry of Eq. (7-2), let us verify the result for the profit maximization model. The explicit choice functions (factor demand functions) that result from the hypothesis, maximize $\pi = pf(x_1, x_2) - w_1x_1 - w_2x_2$ are, again, $x_1 = x_1^*(w_1, w_2, p)$, $x_2 = x_2^*(w_1, w_2, p)$. If these profit-maximizing levels of input are substituted into the objective function, the resulting profit level, by definition, must be the maximum profits attainable at those factor and output prices. Algebraically,

$$\pi^*(w_1, w_2, p) = pf(x_1^*, x_2^*) - w_1x_1^* - w_2x_2^* \tag{7-3}$$

The function $\pi^*(w_1, w_2, p)$ is called the *profit function;* it is the indirect objective function for this model. Its value is always the maximum value of profits for given w_1, w_2, and p.

How do profits vary when, say, w_1 changes? One could simply differentiate the objective function with respect to w_1, holding not only other prices constant, but the input levels x_1 and x_2 constant as well. In that case, we would find

$$\frac{\partial \pi}{\partial w_1} = -x_1$$

No assumption of profit maximization is invoked here. This relation simply says, for example, that if a firm employed 100 workers, and if wages increased by, say, \$1, profits would start to decrease (note the minus sign) by \$100 (100 workers times \$1, the change in the wage rate). However, a profit-maximizing firm would start to reduce the number of its workers as wages increased. If we want to evaluate how *maximum* profit varies when w_1 changes, we must differentiate the *indirect* profit function. Differentiating (7-3) with respect to w_1,

$$\frac{\partial \pi^*}{\partial w_1} = p\left(f_1 \frac{\partial x_1^*}{\partial w_1} + f_2 \frac{\partial x_2^*}{\partial w_1}\right) - w_1 \frac{\partial x_1^*}{\partial w_1} - x_1^* - w_2 \frac{\partial x_2^*}{\partial w_1}$$

Combining the terms involving $\partial x_1^*/\partial w_1$, etc., yields

$$\frac{\partial \pi^*}{\partial w_1} = (pf_1 - w_1)\left(\frac{\partial x_1^*}{\partial w_1}\right) + (pf_2 - w_2)\left(\frac{\partial x_2^*}{\partial w_1}\right) - x_1^*$$

However, the terms in parentheses on the right-hand side are zero at profit-maximizing values of x_1 and x_2 Therefore,

$$\frac{\partial \pi^*}{\partial w_1} = -x_1^* = \frac{\partial \pi}{\partial w_1} \qquad (7\text{-}4)$$

where the latter term must be evaluated at x_1^*. Equation (7-4) says that starting at some profit-maximizing input levels, the *instantaneous* rate of change of profits with respect to a change in a factor price is the same whether or not the factors are held fixed or whether they in principle can vary as that factor price changes. Moreover, the value of this instantaneous rate of change is simply the negative of the factor demand function for x_1, $x_1 = x_1^*(w_1, w_2, p)$, evaluated at the particular prices for which the input levels are in fact profit-maximizing.

We can get a better understanding of what is going on here by considering the geometry more closely. Suppose the factor and output prices have the specific values w_1^0, w_2^0, p^0. Some values of x_1^* and x_2^* are implied:

$$x_1^0 = x_1^*(w_1^0, w_2^0, p^0) \quad x_2^0 = x_2^*(w_1^0, w_2^0, p^0)$$

Let us vary w_1 only, holding w_2 and p fixed at the above values, and observe how the level of profit varies. In particular, we shall initially hold x_1 and x_2 fixed at x_1^0 and x_2^0. In Fig. 7-2, the "constrained" profit function

$$\pi(w_1, w_2^0, p^0, x_1^0, x_2^0) = p^0 f(x_1^0 x_2^0) - w_1 x_1^0 - w_2^0 x_2^0 \qquad (7\text{-}5)$$

shows the level of profits as w_1 varies, holding everything else constant, i.e., for given w_2^0 and p^0, with $x_1 = x_1^0$, $x_2 = x_2^0$. [Note that every variable in Eq. (7-5) has a superscript 0 except w_1.] Note also that $\pi(w_1, w_2^0, p^0, x_1^0, x_2^0)$ is a linear function in w_1. Its slope is $\partial \pi/\partial w_1 = -x_1^0$.

Now consider where the profit function $\pi^*(w_1, w_2^0, p^0)$ lies in relation to this line. Since $\pi^*(w_1, w_2^0, p^0)$ is by definition the *maximum* profits for given factor and output prices, it must in general lie above the straight line defined by $\pi(w_1, w_2^0, p^0, x_1^0, x_2^0)$. However, when $w_1 = w_1^0$, exactly the correct input levels are used, since x_1^0 and x_2^0 were *defined* as the profit-maximizing input levels when $w_1 = w_1^0$. Thus,

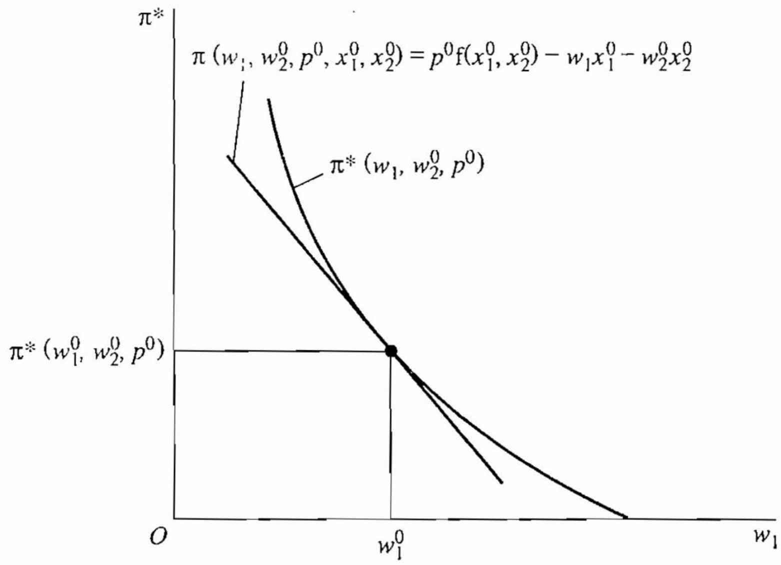

FIGURE 7-2
The profit function $\pi^*(w_1, w_2^0, p^0)$ and the profit function $\pi(w_1, w_2^0, p^0, x_1^0, x_2^0)$, where x_1^0 and x_2^0 are those levels that maximize profits when $w_1 = w_1^0$.

at $w_1 = w_1^0$, $\pi^*(w_1, w_2^0, p^0) = \pi(w_1, w_2^0, p^0, x_1^0, x_2^0)$. When $w_1 \neq w_1^0$, the input levels x_1^0 and x_2^0 are "wrong," i.e., non-profit-maximizing. Hence $\pi^*(w_1, w_2^0, p^0) > \pi(w_1, w_2^0, p^0, x_1^0, x_2^0)$ *on both sides* of w_1^0. But observe the geometric consequences of this in Fig. 7-2. Assuming π^* and π are both differentiable, π^* and π must be tangent to each other at w_1^0. Tangency means that π^* and π have the same slope at w_1^0. This is precisely Eq. (7-4), $\partial\pi^*/\partial w_1 = \partial\pi/\partial w_1 = -x_1^*$.

Suppose we had started at some other level of w_1, say w_1^1. In that case we would have held x_1 and x_2 fixed at the levels implied by that wage, $x_1^1 = x_1^*(w_1^1, w_2^0, p^0)$, $x_2^1 = x_2^*(w_1^1, w_2^0, p^0)$. The resulting constrained profit function would be some other straight line tangent to π^* at this different value of w_1; their common slope at this point would be $-x_1^*(w_1^1, w_2^0, p^0)$. We can see the reason for the name "envelope" theorem: the profit function $\pi^*(w_1, w_2, p)$ is the envelope of all the possible constrained profit lines as w_1 is varied.

However, we have more information than just the equality of slope of π and π^*. Since π^* lies above π on both sides of w_1^0, $\pi^*(w_1, w_2^0, p^0)$ must be more convex (or less concave) than $\pi(w_1, w_2^0, p^0, x_1^0, x_2^0)$. But in this model, π is linear, and therefore $\pi^*(w_1, w_2^0, p^0)$ must be convex in w_1, as shown in Fig. 7-2. That the indirect function is convex (we assume *strictly* convex) has major consequences for the comparative statics of this model. Convexity in w_1 means $\partial^2\pi^*/\partial w_1^2 \geq 0$. But from Eq. (7-4), $\partial\pi^*/\partial w_1 = -x_1^*(w_1, w_2, p)$. Differentiating both sides therefore yields

$$\frac{\partial^2\pi^*}{\partial w_1^2} = -\frac{\partial x_1^*}{\partial w_1} \geq 0 \tag{7-6}$$

Since in this model the factor demand function $x_1^*(w_1, w_2, p)$ is in fact the negative of the first partial of $\pi^*(w_1, w_2, p)$ with respect to w_1, the slope of the factor demand function (its first partial with respect to w_1) is the negative second partial derivative of π^* with respect to w_1. Since this second partial of π^* is positive (nonnegative), the slope of the factor demand function must be negative. Thus (in this model at least), the curvature of the indirect objective function (the profit function, here) directly implies an important comparative statics result.

By symmetry, it follows obviously that $\pi^*(w_1, w_2, p)$ is convex in w_2, yielding the same comparative statics result for that factor. It is also the case that $\pi^*(w_1, w_2, p)$ is convex in output price p, and that therefore $\partial^2\pi^*/\partial p^2 = \partial y^*/\partial p \geq 0$. The proof and geometrical explanation of this are left as an exercise. We now turn to an examination of the general maximization model. Can the preceding results be derived without resort to visual geometry?

7.3 GENERAL COMPARATIVE STATICS ANALYSIS: UNCONSTRAINED MODELS

Consider any two-variable model, maximize $y = f(x_1, x_2, \alpha)$, where x_1 and x_2 are the choice variables and, for the moment, α is a single parameter representing some constraint on the maximizing agent's behavior. The first-order equations are $f_1 = f_2 = 0$. By solving the first-order equations simultaneously, assuming unique solutions, explicit choice functions $x_1 = x_1^*(\alpha)$, $x_2 = x_2^*(\alpha)$ are implied. Again, the refutable propositions consist of the implications of maximization regarding the directions of change in some or all x_i's as α changes. The "indirect objective function" is, again, $\phi(\alpha) = f(x_1^*(\alpha), x_2^*(\alpha), \alpha)$. By definition, $\phi(\alpha)$ gives the maximum value of f for given α. At what rates do $\phi(\alpha)$ and $f(x, \alpha)$ vary (both first- and second-order rates of change) as α changes?

In Fig. 7-3, $\phi(\alpha)$ is plotted for various α's. For an arbitrary α^0 some $x_1^0 = x_1^*(\alpha^0)$ and $x_2^0 = x_2^*(\alpha^0)$ are implied. Consider the behavior of $f(x_1, x_2, \alpha)$ when x_1 and x_2 are held fixed at x_1^0 and x_2^0 as opposed to when they are variable. Since $\phi(\alpha)$ is the *maximum* value of f for given α, in general, $f \leq \phi$. When $\alpha = \alpha^0$, the "correct" x_i's are chosen, and therefore $\phi(\alpha) = f(x_1, x_2, \alpha)$ at that one point. On both sides of α^0, the "wrong" (i.e., nonmaximizing) x_i's are used, and thus by definition,

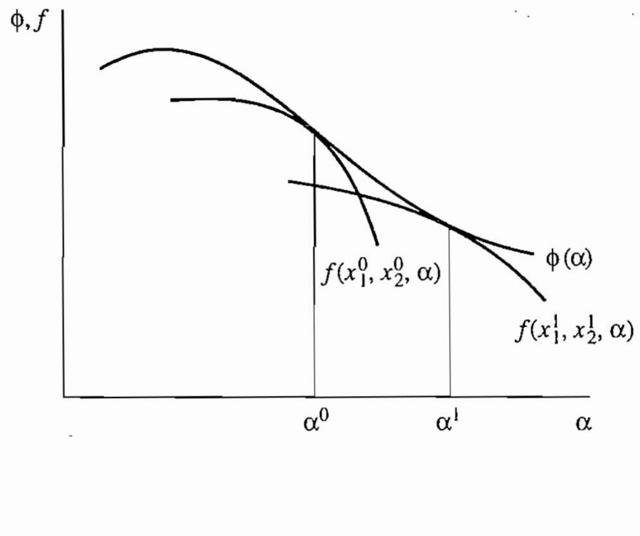

FIGURE 7-3
The indirect objective function $\phi(\alpha)$ is an envelope curve to the direct objective functions for various α's.

$f(x_1^0, x_2^0, \alpha) < \phi(\alpha)$ in any neighborhood around α^0. Unless f has some sort of nondifferentiable corner at α^0, ϕ and f must be tangent at α^0, and, moreover, f must be either more concave or less convex than ϕ there. Since this must happen for arbitrary α, similar tangencies occur at other values of α. It is apparent from the diagram that $\phi(\alpha)$ is the *envelope* of the $f(x_1, x_2, \alpha)$'s for each α. How do we derive these properties algebraically?

Consider a new function, the difference between the actual and the maximum value of f for given α,

$$F(x_1, x_2, \alpha) = f(x_1, x_2, \alpha) - \phi(\alpha)$$

called the *primal-dual* objective function. Since $f \leq \phi$ for $x \neq x^*$ and $f = \phi$ for $x_i = x_i^*$, F has a maximum (of zero) when $x_i = x_i^*(\alpha)$.† Moreover, we can consider $F(x_1, x_2, \alpha)$ as a function of *three* independent variables, x_1, x_2, and α. That is, just as for a given α there are values of x_1 and x_2 that maximize f, for *given* x_1 and x_2, there is some value of α which makes those x_i's the "correct" (i.e., maximizing) values. For example, for a given amount of labor and capital, there is some set of factor and output prices for which those input levels would be the profit-maximizing values.

This maximum position of $F(x_1, x_2, \alpha)$ can be described by the usual first- and second-order conditions. The first-order conditions are that $f(x_1, x_2, \alpha) - \phi(\alpha)$ has zero partial derivatives with respect to the original choice variables x_1 and x_2, *and also* α:

$$F_i = f_i = 0 \quad i = 1, 2 \tag{7-7}$$

and

$$F_\alpha = f_\alpha - \phi_\alpha = 0 \tag{7-8}$$

Equations (7-7) are simply the original maximum conditions. Equation (7-8) is the "envelope" result, $\phi_\alpha = f_\alpha$. These first-order conditions hold whenever $x_i = x_i^*(\alpha)$, $i = 1, 2$.

The sufficient second-order conditions state that the Hessian matrix of second partials of $F(x_1, x_2, \alpha)$ (with respect to x_1, x_2, *and* α) is negative definite, or that its principal minors alternate in sign. By inspection, $F_{11} = f_{11}$, etc., and $F_{\alpha\alpha} = f_{\alpha\alpha} - \phi_{\alpha\alpha}$. Thus,

$$H = \begin{vmatrix} F_{11} & F_{12} & F_{1\alpha} \\ F_{21} & F_{22} & F_{2\alpha} \\ F_{\alpha 1} & F_{\alpha 2} & F_{\alpha\alpha} \end{vmatrix} = \begin{vmatrix} f_{11} & f_{12} & f_{1\alpha} \\ f_{21} & f_{22} & f_{2\alpha} \\ f_{\alpha 1} & f_{\alpha 2} & f_{\alpha\alpha} - \phi_{\alpha\alpha} \end{vmatrix} \tag{7-9}$$

†If we think of x_1^* and x_2^* as an "efficient" allocation of resources, and x_1 and x_2 as any other allocation, then this says that efficient allocation occurs when "waste" equals zero.

These second-order conditions include the original ones ($f_{11} < 0$, $f_{11}f_{22} - f_{12}^2 > 0$, etc.) in the top left corner. In addition, the sufficient second-order conditions also imply $F_{\alpha\alpha} < 0$, or $f_{\alpha\alpha} - \phi_{\alpha\alpha} < 0$. *Moreover, it is from this inequality that all known comparative statics results (in maximization models) flow.*

The first-order envelope result (7-8), with the functional dependence noted, is $\phi_\alpha(\alpha) \equiv f_\alpha(x_1^*(\alpha), x_2^*(\alpha), \alpha)$. Differentiating both sides with respect to α yields

$$\phi_{\alpha\alpha} \equiv f_{\alpha x_1}\frac{\partial x_1^*}{\partial \alpha} + f_{\alpha x_2}\frac{\partial x_2^*}{\partial \alpha} + f_{\alpha\alpha}$$

From the sufficient second-order conditions, therefore, and using Young's theorem,

$$\phi_{\alpha\alpha} - f_{\alpha\alpha} = f_{1\alpha}\frac{\partial x_1^*}{\partial \alpha} + f_{2\alpha}\frac{\partial x_2^*}{\partial \alpha} > 0$$

This analysis is readily generalized to the n-variable case, producing the condition

$$\sum_{i=1}^{n} f_{i\alpha}\frac{\partial x_i^*}{\partial \alpha} > 0 \tag{7-10}$$

Equation (7-10) is the general and fundamental comparative statics equation for all unconstrained maximization models. As it stands, however, it is too general to be of much use. In order for a model to have refutable implications, it must contain more structure than just a general maximization problem. Suppose therefore that some α enters only one first-order condition $f_i = 0$, i.e., $f_{j\alpha} = 0$ for $j \neq i$. Then Eq. (7-10) reduces to a single term,

$$f_{i\alpha}\frac{\partial x_i^*}{\partial \alpha} > 0 \tag{7-11}$$

This is Samuelson's famous "conjugate pairs" result. In maximization models, if some parameter α enters only the ith first-order equation, the response of the ith choice variable x_i to a change in that parameter is in the same direction as the effect α has on the first-order equation.

The significance of this theorem lies in its application to some important models. For example, in the profit maximization model, the parameter w_1 enters only the first first-order equation $\pi_1 = pf_1 - w_1 = 0$; it enters with a negative sign: $\partial\pi_1/\partial w_1 = -1$. Thus, the conjugate pairs theorem states that the response of x_1^* to an increase in w_1 will be negative, and similarly for x_2^*. The theorem also applies to the constrained cost minimization model, as we shall presently see.

In the more general case where x is a vector of decision variables $(x_1, \ldots, x_n)$, and α is a vector of parameters $\alpha = (\alpha_1, \ldots, \alpha_m)$, the second-order conditions for maximizing $F(x, \alpha) = f(x, \alpha) - \phi(\alpha)$ with respect to α are that the matrix $F_{\alpha\alpha} = f_{\alpha\alpha} - \phi_{\alpha\alpha}$ is negative semidefinite. The usual comparative statics results follow from the negativity of the diagonal elements of this matrix. However, a richer set of theorems is also available from the other properties of negative semidefinite matrices: The principal minors of the terms in $f_{\alpha\alpha} - \phi_{\alpha\alpha}$ alternate in sign.

The envelope theorem also reveals the origins of the nonintuitive "reciprocity" conditions that appear in maximization models. Recall that in the profit maximization model, we derived $\partial x_1^*/\partial w_2 = \partial x_2^*/\partial w_1$. This result can be more clearly shown by first noting that each factor demand is the negative first partial of π^* with respect to its factor price, i.e., $\pi_1^* = -x_1^*(w_1, w_2, p)$, $\pi_2^* = -x_2^*(w_1, w_2, p)$. Applying Young's theorem on invariance of cross-partials to the order of differentiation to $\pi^*(w_1, w_2, p)$ therefore yields $\pi_{12}^* = -\partial x_1^*/\partial w_2 = -\partial x_2^*/\partial w_1 = \pi_{21}^*$. Thus this curious result is no more curious than Young's theorem itself.

All reciprocity theorems are in fact simply the application of Young's theorem to the indirect objective function. Suppose there are two parameters α and β so that the model is maximize $y = f(x_1, x_2, \alpha, \beta)$. The implied choice functions are then $x_i = x_i^*(\alpha, \beta)$, $i = 1, 2$, and the indirect objective function is $\phi(\alpha, \beta) = f(x_1^*(\alpha, \beta), x_2^*(\alpha, \beta), \alpha, \beta)$. Then noting that $\phi_\alpha(\alpha, \beta) = f_\alpha$,

$$\phi_{\alpha\beta}(\alpha, \beta) = f_{\alpha 1}\frac{\partial x_1^*}{\partial \beta} + f_{\alpha 2}\frac{\partial x_2^*}{\partial \beta} + f_{\alpha\beta}$$

Similarly,

$$\phi_{\beta\alpha}(\alpha, \beta) = f_{\beta 1}\frac{\partial x_1^*}{\partial \alpha} + f_{\beta 2}\frac{\partial x_2^*}{\partial \alpha} + f_{\beta\alpha}$$

Since $\phi_{\alpha\beta} = \phi_{\beta\alpha}$,

$$f_{1\alpha}\frac{\partial x_1^*}{\partial \beta} + f_{2\alpha}\frac{\partial x_2^*}{\partial \beta} = f_{1\beta}\frac{\partial x_1^*}{\partial \alpha} + f_{2\beta}\frac{\partial x_2^*}{\partial \alpha} \tag{7-12}$$

For the general case of n decision variables,

$$\sum f_{i\alpha}\frac{\partial x_i^*}{\partial \beta} = \sum f_{i\beta}\frac{\partial x_i^*}{\partial \alpha} \tag{7-13}$$

However, these relations are most interesting when each parameter enters only one first-order equation. In that case, Eq. (7-13) reduces to one term on each side, as in the profit maximization model.

7.4 MODELS WITH CONSTRAINTS

Most models in economics involve one or more side constraints. A particularly important model, for example, is

minimize

$$C = \sum w_i x_i$$

subject to

$$f(x_1, \ldots, x_n) = y^0$$

If f is a production function of n inputs, $x_1, \ldots, x_n$, and the w_i's are factor prices, this famous model, which we shall presently analyze in detail, describes achieving some output level y^0 at minimum cost.

The extension of the results for unconstrained maximization models to models involving one or more side conditions (constraints) depends critically on whether the parameters enter only the objective function or whether they enter the constraints also (or exclusively). Note that in the preceding cost minimization model, the prices enter only the objective function, whereas the specified output level enters only the constraint. We shall show that if the parameters enter only the objective function, the comparative statics results are the same as for unconstrained models. However, if a parameter enters a constraint, as that parameter changes, the constraint space also changes, destroying the relation $\phi_{\alpha\alpha} \geq f_{\alpha\alpha}$. Let us investigate these more general models.

The traditional derivation of the envelope theorem for models with one constraint proceeds as follows.

Consider

maximize

$$f(x_1, \ldots, x_n, \alpha) = y$$

subject to

$$g(x_1, \ldots, x_n, \alpha) = 0$$

The Lagrangian is $\mathscr{L} = f + \lambda g$. Setting the first partials of $\mathscr{L}$ equal to 0,

$$\mathscr{L}_i = f_i + \lambda g_i = 0 \quad i = 1, \ldots, n \tag{7-14}$$

$$\mathscr{L}_\lambda = g = 0 \tag{7-15}$$

Solving these equations for

$$x_i = x_i^*(\alpha) \quad i = 1, \ldots, n$$

$$\lambda = \lambda^*(\alpha)$$

we define

$$y^* = f(x_1^*, \ldots, x_n^*, \alpha) = \phi(\alpha) \tag{7-16}$$

as before. Here, $\phi(\alpha)$ is the maximum value of y for any α, for x_i's that satisfy the constraint.

How does $\phi(\alpha)$ change when α changes? Differentiating (7-16) with respect to α

$$\frac{\partial \phi}{\partial \alpha} = \sum f_i \frac{\partial x_i^*}{\partial \alpha} + f_\alpha \tag{7-17}$$

Here, however, $f_i \neq 0$. Differentiating the constraint

$$g(x_1^*(\alpha), \ldots, x_n^*(\alpha), \alpha) \equiv 0$$

with respect to α,

$$\sum g_i \frac{\partial x_i^*}{\partial \alpha} + g_\alpha \equiv 0 \tag{7-18}$$

Multiply Eq. (7-18) by λ, and add to Eq. (7-17). (This adds zero to that expression.) Then

$$\begin{aligned}\frac{\partial \phi}{\partial \alpha} &= \sum f_i \frac{\partial x_i^*}{\partial \alpha} + f_\alpha + \sum \lambda g_i \frac{\partial x_i^*}{\partial \alpha} + \lambda g_\alpha \\ &= \sum (f_i + \lambda g_i) \frac{\partial x_i^*}{\partial \alpha} + f_\alpha + \lambda g_\alpha\end{aligned}$$

Using the first-order conditions (7-14),

$$\frac{\partial \phi}{\partial \alpha} = f_\alpha + \lambda g_\alpha = \mathcal{L}_\alpha \tag{7-19}$$

where $\mathcal{L}_\alpha$ is the partial derivative of the Lagrangian function with respect to α, holding the x_i's fixed. Thus, in evaluating the response of the indirect objective function to a change in a parameter in a *constrained* maximization model, the Lagrangian function plays an analogous role to the objective function in an unconstrained model.

We can derive the envelope theorem for constrained maximization models more conveniently using primal-dual analysis. It is still the case in these models that $\phi(\alpha) \geq f(x_1, \ldots, x_n, \alpha)$, but in this case, the variables must also satisfy the constraint. The primal-dual model is therefore

maximize

$$f(x_1, \ldots, x_n, \alpha) - \phi(\alpha)$$

subject to

$$g(x_1, \ldots, x_n, \alpha) = 0$$

treating α as a (vector of) decision variables as well as the x_i's. The Lagrangian for the primal-dual problem is

$$\mathcal{L} = f(x_1, \ldots, x_n, \alpha) - \phi(\alpha) + \lambda g(x_1, \ldots, x_n, \alpha)$$

Setting the first partials of $\mathcal{L}$ with respect to the x_i's and λ equal to zero produces the ordinary first-order equations (7-14) and (7-15) for a constrained maximum; setting the first partial of $\mathcal{L}$ with respect to α equal to zero produces the envelope relation (7-19) above.

Comparative Statics: Primal-Dual Analysis

We now investigate, using primal-dual analysis, the conditions under which refutable propositions appear in constrained maximization models. We already know from traditional methods developed in Chap. 6 that no refutable propositions appear for parameters that appear in the constraint. We refer the reader to Silberberg's 1974 comparative statics paper for the general results. We can demonstrate the nature of the more likely useful results using the following simple model. Consider

maximize

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

subject to

$$g(x_1, x_2, \beta) = 0$$

In this model, a single parameter α enters the objective function only, and another parameter, β, enters the constraint only. Using Lagrangian techniques, the first-order equations are solved for the explicit choice equations

$$x_1 = x_1^*(\alpha, \beta)$$

$$x_2 = x_2^*(\alpha, \beta)$$

Substituting these solutions into the objective function yields the maximum value of $f(x_1, x_2, \alpha)$ for given α and β, for x_1 and x_2 that satisfy the constraint:

$$\phi(\alpha, \beta) = f(x_1^*(\alpha, \beta), x_2^*(\alpha, \beta), \alpha)$$

Since $\phi(\alpha, \beta)$ is the maximum value of f for given α and β, $\phi(\alpha, \beta) \geq f(x_1, x_2, \alpha)$ for any x_i's that satisfy the constraint. Thus, the function $F(x_1, x_2, \alpha, \beta) = f(x_1, x_2, \alpha) - \phi(\alpha, \beta)$ has a maximum (of zero) for any x_i's that satisfy the constraint. However, $F(x_1, x_2, \alpha, \beta)$ is a function of four independent variables, one of which, α, does not enter the constraint. Therefore, starting with values of x_1, x_2, and β which satisfy the constraint, and holding them fixed at those values, the constraint does not further impinge on the choice of α that maximizes $F(x_1, x_2, \alpha, \beta)$. The constraint affects the values of x_1 and x_2 that can be chosen, but not the maximizing value of α. *In the α dimension(s), therefore, $F(x_1, x_2, \alpha, \beta)$ has an unconstrained maximum.* (Consider, for example, what happens when some good, say, air, enters a person's utility function, but not the budget constraint, there being no price paid for breathing. In that case, we breathe until the marginal utility of air is zero; i.e., we consume in the manner of an unconstrained maximum in that dimension.)

The Lagrangian for this primal-dual problem is

$$\mathcal{L} = f(x_1, \ldots, x_n, \alpha) - \phi(\alpha, \beta) + \lambda g(x_1, \ldots, x_n, \beta)$$

The envelope relations are obtained by setting the first partials of $\mathcal{L}$ with respect to α and β equal to zero, yielding

$$f_\alpha - \phi_\alpha = 0 \tag{7-20a}$$

$$-\phi_\alpha + \lambda g_\beta = 0 \tag{7-20b}$$

Equation (7-20*a*) is just Eq. (7-8), the same envelope relation for unconstrained models. Moreover, since this primal-dual model is an unconstrained maximum in α, $F_{\alpha\alpha} = f_{\alpha\alpha} - \phi_{\alpha\alpha} < 0$, assuming, as always, the sufficient second-order conditions.

The fundamental comparative statics result (7-10) follows as before:

$$f_{1\alpha}\frac{\partial x_1^*}{\partial \alpha} + f_{2\alpha}\frac{\partial x_2^*}{\partial \alpha} > 0 \tag{7-10}$$

If α represents a vector of parameters that enter the objective function only, then the matrix of terms $(f_{\alpha\alpha} - \phi_{\alpha\alpha})$ must be negative semidefinite; Eq. (7-10) then follows from the fact that the diagonal elements are nonpositive.

No such easy relationships exist with regard to changes in β. To best see this, try to construct a diagram like Fig. 7-3 for the parameter β. Plot β on the horizontal axis and f and $\phi(\alpha, \beta)$ on the vertical axis. Hold α constant throughout. At some value β^0, $x_1^0 = x_1^*(\alpha^0, \beta^0)$, $x_2^0 = x_2^*(\alpha^0, \beta^0)$ are implied. The next step is to vary the parameter in question, holding x_1 and x_2 constant. However, it is impossible to do that for β. In the first place, since β is not a variable in the objective function f, it is impossible to plot f against β. Second, if x_1 and x_2 are held constant, β cannot be changed without violating the constraint! Thus the procedure for showing the greater relative concavity of f vs. ϕ breaks down for parameters entering the constraint: One cannot change only one variable in an equation without destroying the equality. As a result, no refutable hypotheses are implied by the maximization hypothesis alone, for parameters that enter the constraint.

In the case where β is a vector of two or more parameters $(\beta_1, \ldots, \beta_m)$, it is possible to hold x_1, x_2 and α constant and characterize the β_j's that solve the primal-dual problem. Since the original objective function does not contain any of the β_j's, the primal-dual problem reduces to

maximize
β

$$-\phi(\alpha, \beta)$$

subject to

$$g(x, \beta) = 0$$

where $x = (x_1, x_2)$ (or, for that matter, a general n-dimensional vector of decision variables). Of course, maximizing $-\phi(\alpha, \beta)$ is the same as minimizing $\phi(\alpha, \beta)$; thus in this case, the indirect objective function is *convex in the β parameters, subject to constraint,* i.e., in the parameters that enter the constraint exclusively. If the constraint is linear in the β_j's, then the indirect objective function must be *quasi-convex* in these parameters (though linearity is not a necessary condition for quasi-convexity).

Example. In the important consumer model, utility of goods is maximized subject to a linear budget constraint:

maximize

$$U(x_1, x_2)$$

subject to

$$p_1x_1 + p_2x_2 = M$$

Using Lagrangian methods, the implied choice functions are the Marshallian demands $x_i = x_i^*(p_1, p_2, M)$, $i = 1, 2$. Substituting these functions into the objective function yields the *indirect utility function* $U^*(p_1, p_2, M) = U(x_1^*(p_1, p_2, M), x_2^*(p_1, p_2, M))$. The primal-dual problem is thus

maximize

$$U(x_1, x_2) - U^*(p_1, p_2, M)$$

subject to

$$p_1 x_1 + p_2 x_2 = M$$

where the maximization runs over x_1, x_2, *and* the parameters p_1, p_2, and M. Since all the parameters are in the constraint exclusively, the maximization problem with respect to the prices and money income is simply

maximize

$$-U^*(p_1, p_2, M)$$

subject to

$$p_1 x_1 + p_2 x_2 = M$$

This says that choosing goods x_1 and x_2 so as to maximize utility (subject to the budget constraint) is equivalent to choosing prices and money income so as to minimize the indirect utility function, also, of course, subject to the budget constraint. Since the budget constraint is *linear* in prices and money income, this implies that the indirect utility function is quasi-convex in prices and money income. The result generalizes immediately to the case of n goods.

Reciprocity relations can be derived in these models using the envelope relations (7-20). Writing these relations as identities and showing the functional dependencies using the explicit choice functions, we have

$$\phi_\alpha(\alpha, \beta) \equiv f_\alpha(x_1^*(\alpha, \beta), x_2^*(\alpha, \beta), \alpha) \tag{7-21a}$$

$$\phi_\beta(\alpha, \beta) \equiv \lambda^*(\alpha, \beta) g_\beta(x_1^*(\alpha, \beta), x_2^*(\alpha, \beta), \beta) \tag{7-21b}$$

Identity (7-21a) is the same as in the case of models without constraints, because the α parameters enter only the objective function. For two such parameters α_1 and α_2, we derive the reciprocity conditions displayed in Eqs. (7-12) and (7-13) in the same manner as before. In addition, since $\phi_{\alpha\beta} \equiv \phi_{\beta\alpha}$, we derive, using the product as well as the chain rule on the right-hand side of (7-21b),

$$f_{1\alpha} \frac{\partial x_1^*}{\partial \beta} + f_{2\alpha} \frac{\partial x_2^*}{\partial \beta} = \lambda^* \left(g_{1\beta} \frac{\partial x_1^*}{\partial \alpha} + g_{2\beta} \frac{\partial x_2^*}{\partial \alpha} \right) + g_\beta \frac{\partial \lambda^*}{\partial \alpha} \tag{7-22}$$

An additional set of reciprocity relations is available in the case of two parameters β_1 and β_2 that both enter the constraint only; these relationships necessarily involve the partial derivatives of λ^* as well, as is apparent from (7-21b). We leave these derivations as an exercise for the student. At this level of generality, these reciprocity relations are not very interesting, but in many more specialized models, (7-22) reduces to interesting expressions. Last, very general reciprocity relations can be derived in models in which the parameters enter both the objective function and the constraint, but there are no known instances of any interesting ones.

An Important Special Case

Most of the useful models encountered in economics involve expressions that are linear in at least some of the parameters, typically the prices of goods or factors. Consider, therefore, models in which the objective function involves the expression $\sum \alpha_i x_i$:

maximize

$$y = f(x, \alpha) = \theta(x_1, \ldots, x_n) + \sum \alpha_i x_i \tag{7-23}$$

subject to

$$g(x_1, \ldots, x_n, \beta) = 0 \tag{7-24}$$

where $x = (x_1, \ldots, x_n)$, the vector of decision variables, $\alpha = (\alpha_1, \ldots, \alpha_n)$, and β is any vector of parameters entering the constraint only. Parameters that enter the constraint are assumed to be absent from the objective function.

Denote the indirect objective function $\phi(\alpha, \beta)$. We know from the preceding analysis that the function $f(x, \alpha) - \phi(\alpha, \beta)$ must be concave in α and that the matrix $(f_{\alpha\alpha} - \phi_{\alpha\alpha})$ must therefore be negative semidefinite. The parameters β and the functional form of g are irrelevant, as long as the first- and second-order conditions are satisfied. However, since f is linear in the α_i's, $f_{\alpha\alpha} \equiv 0$, and thus f has no effect on the curvature of the primal-dual function. Therefore, for these models, $-\phi$ is concave (or, alternatively, ϕ is convex) in α, and the matrix $[-\phi_{\alpha\alpha}]$ must be symmetric (by Young's theorem) and negative semidefinite (or, $[\phi_{\alpha\alpha}]$ is positive semidefinite). In the case of *minimization* models with these properties, $\phi(\alpha)$ is *concave,* and $[\phi_{\alpha\alpha}]$ is a negative semidefinite matrix.

Even more important than these curvature properties are the implications for deriving useful comparative statics theorems. By the envelope theorem, $\phi_{\alpha_i} = f_{\alpha_i} = x_i^*$ in these models. Therefore, the matrix $[\phi_{\alpha\alpha}]$ consists of the terms $\partial x_i^* / \partial \alpha_j$. From symmetry, $\partial x_i^* / \partial \alpha_j = \partial x_j^* / \partial \alpha_i$. The properties of positive semidefinite matrices include nonnegative diagonal terms, i.e., $\partial x_i^* / \partial \alpha_i \geq 0$, and positive principal minors of higher order. These results comprise the useful theorems in economics.

The profit function derived above exhibited these properties (but note that the α_i's are the negative prices). In the next chapter we will study the cost minimization model, which has a similar structure. We shall show that the cost function associated with production functions with the usual properties must be concave, and the demand functions implied by that model are negatively sloped.

Interpretation of the Lagrange Multiplier

The Lagrange multiplier λ has been carried along thus far mainly as a convenient way of stating the first- and second-order conditions for maximization. In fact, the main reason for the use of Lagrangian techniques in economics (and also other sciences) is that λ often has an interesting interpretation of its own.

Consider the constrained maximization model

maximize

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

subject to

$$g(x_1, x_2) = k$$

Usually we set the constraint equation equal to zero; here, it equals some arbitrary value k. By stating the constraint in this manner, we can consider parametric changes in the value of the g function. Using the Lagrangian

$$\mathscr{L} = f(x_1, x_2) + \lambda(k - g(x_1, x_2))$$

the usual first-order equations are

$$\mathscr{L}_1 = f_1(x_1, x_2) - \lambda g_1(x_1, x_2) = 0 \qquad (7\text{-}25a)$$

$$\mathscr{L}_2 = f_2(x_1, x_2) - \lambda g_2(x_1, x_2) = 0 \qquad (7\text{-}25b)$$

$$\mathscr{L}_\lambda = k - g(x_1, x_2) = 0 \qquad (7\text{-}25c)$$

From Eqs. (7-25a) and (7-25b),

$$\lambda = \frac{f_1}{g_1} = \frac{f_2}{g_2} \qquad (7\text{-}26)$$

However, a more revealing expression for λ can be obtained using the envelope theorem.

By solving Eqs. (7-25) simultaneously, we obtain the explicit choice functions $x_1^*(k)$, $x_2^*(k)$, and $\lambda^*(k)$. Substituting these solutions into $f(x_1, x_2)$ yields the indirect objective function

$$\phi(k) = f(x_1^*(k), x_2^*(k))$$

By the envelope theorem for constrained maximization models, Eq. (7-19),

$$\phi_k(k) = \frac{\partial \mathscr{L}}{\partial k} = \lambda^*(k) \qquad (7\text{-}27)$$

That is, the Lagrange multiplier λ equals the rate of change of the maximum (or minimum, as the case may be) value of the objective function with respect to parametric changes in the value of the constraint.

Consider Eq. (7-19) again, $\phi_\alpha(\alpha) = f_\alpha + \lambda g_\alpha$. We can understand this relation by using (7-27). Think of (7-19) as

$$\frac{\partial \phi}{\partial \alpha} = \frac{\partial f}{\partial \alpha} + \frac{\partial f}{\partial g}\frac{\partial g}{\partial \alpha}$$

When a parameter that enters both the objective function and the constraint changes, it produces two separate effects. First, the objective function is affected directly, as indicated by the term $\partial f/\partial \alpha$. In addition, the value of the constraint is affected, by the amount $\partial g/\partial \alpha$. This is then converted to units of the objective function f by multiplying by $\lambda (= \partial f/\partial g)$. The sum of these two effects is the total impact of a change in α on the maximum value of y.

A common application of Eq. (7-27) concerns models in which the objective function is some sort of value of output function, which is maximized subject to a resource constrained to some level k. If an additional increment of resource, Δk, became available, output would increase by some amount $\Delta y^* \approx \lambda^* \Delta k$; in other words, λ^* *is the marginal value of that resource.* In a competitive economy, firms would be willing to pay λ^* for each increment in the resource. In the mathematical programming literature, λ^* is called a *shadow price* of the resource. In a model in which output of society is maximized subject to constraints of parametric labor and capital constraints, the Lagrange multipliers associated with those constraints impute shadow factor prices, i.e., a wage and rental rate to labor and capital. In the next chapter, in a model in which total cost is minimized subject to producing some parametric output level, λ^* measures the change in total cost if output is changed, i.e., marginal cost. We shall explore these relationships in the chapters following.

Consider again the model

maximize

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

subject to

$$g(x_1, x_2) = k$$

Since the parameter k enters the constraint, we know that in general, the sign of $\partial\lambda^*/\partial k$ is indeterminate. However, in some important models, additional assumptions provide a sign for this term. Differentiating Eqs. (7-25) with respect to k,

$$\begin{pmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} & g_1 \\ \mathcal{L}_{21} & \mathcal{L}_{22} & g_2 \\ -g_1 & -g_2 & 0 \end{pmatrix} \begin{pmatrix} \dfrac{\partial x_1^*}{\partial k} \\ \dfrac{\partial x_2^*}{\partial k} \\ \dfrac{\partial \lambda^*}{\partial k} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \tag{7-28}$$

Solving for $\partial\lambda^*/\partial k$,

$$\frac{\partial \lambda^*}{\partial k} = -\frac{H_{33}}{H}$$

where H is the bordered Hessian determinant of the Lagrangian $\mathscr{L}$. From the sufficient second-order conditions, $H > 0$. Suppose now that f and g are strictly increasing functions so that $\lambda^* > 0$ (why?). Suppose in addition that f is a concave and g is a convex function. Then $-g$ must be concave, and thus $\mathscr{L}$ is concave. In this case, then, $H_{33} = \mathscr{L}_{11}\mathscr{L}_{22} - \mathscr{L}_{12}^2 \geq 0$, and thus $\partial\lambda^*/\partial k \leq 0$. If g is linear, these conditions are met as long as f is concave. It is also possible to show, via primal-dual methods, that if $\partial\lambda^*/\partial k < 0$, $\mathscr{L}$ must be a strictly concave function; the proof is left as an exercise.

These results generalize in a straightforward manner to maximization models with multiple constraints,

maximize

$$f(x) = y$$

subject to

$$g(x) \leq k$$

where $x = (x_1, \ldots, x_n)$, $g(x) = g^j(x_1, \ldots, x_n)$, and $k = (k_1, \ldots, k_m)$, $j = 1, \ldots, m$. The choice functions $x = x^*(k)$ and the Lagrange multipliers $\lambda^*(k)$ implied by this model are obtained by simultaneous solution of the first-order Lagrangian conditions, assuming the sufficient second-order conditions hold. The indirect objective function is $\phi(k) = f(x^*(k))$. By the envelope theorem, $\lambda^{j*}(k) = \partial\phi/\partial k^j$, the marginal value of relaxing the jth "resource constraint" k^j, measured by the resulting increase in the value of the objective function. *If $f(x)$ is concave and $g^j(x)$ is convex for $j = 1, \ldots, m$, $\phi(k)$ is concave in k, and thus $(\phi_{kk}) = (\partial\lambda^*/\partial k)$ is negative semidefinite.* Since the diagonal elements of (ϕ_{kk}) would then be nonpositive, this implies that $\partial\lambda^*/\partial k \leq 0$. In many important models, the constraints are linear; such a specification satisfies the conditions of this theorem.

The proof relies on the definitions of concave and convex functions. Let k^1 and k^2 be two arbitrary values of the k vectors, and denote the implied choice vectors as $x^1 = x^*(k^1)$, $x^2 = x^*(k^2)$. Let $k^t = tk^1 + (1-t)k^2$, $x^t = tx^1 + (1-t)x^2$, $0 \leq t \leq 1$. By convexity of the constraints,

$$g(x^t) \leq tg(x^1) + (1-t)g(x^2) \leq tk^1 + (1-t)k^2 = k^t$$

Therefore, x^t is a feasible choice for, or solution to, this model; it satisfies the constraints when $k = k^t$.

Since $f(x)$ is concave,

$$f(x^t) \geq tf(x^1) + (1-t)f(x^2) = t\phi(k^1) + (1-t)\phi(k^2)$$

But by the definition of ϕ, $\phi(k^t) \geq f(x^t)$. Therefore,

$$\phi(k^t) \geq t\phi(k^1) + (1-t)\phi(k^2)$$

Therefore, $\phi(k)$ is concave in k. Assuming differentiability, the Hessian matrix ϕ_{kk} is of course negative semidefinite, yielding the usual comparative statics results in those cases. An important application of this result occurs in the "small country" models of international trade, where total output of an economy is maximized subject to endowment constraints. The factor prices are the associated Lagrange multipliers of those endowment constraints. If the production functions are concave, this theorem implies that an increase in the endowment of some factor cannot increase that factor's price. This model will be developed more fully in the chapters on general equilibrium.

Le Châtelier Effects

We now consider the responses of decision variables to a change in some parameter when an additional just-binding constraint is added to the model. We investigated these Le Châtelier effects in Chap. 4 for the profit maximization model. We showed in that model that if one factor is held constant at its profit-maximizing level, then in a neighborhood of that equilibrium the demand for the remaining factor becomes less elastic. We now consider more general models, and, as always, we are most interested in discovering the structure of models that yield predictable differences in the responses of the choice variables to parameter changes when a just-binding constraint is added. Since no refutable results are available in models in which parameters enter the constraint as well as the objective function, we limit the discussion to models in which some parameters α enter the objective function only and other parameters β enter the constraint only. To save notational clutter, we shall use vector notation throughout this section. Thus, recapitulating, consider

maximize

$$y = f(x, \alpha)$$

subject to

$$g(x, \beta) = 0$$

where $x = (x_1, \ldots, x_n)$ and α and β are vectors of parameters that appear only in the objective function and constraint, respectively. Assuming the first-order necessary and second-order sufficient conditions hold, we derive the explicit choice functions $x^*(\alpha, \beta)$ and $\lambda^*(\alpha, \beta)$. The indirect objective function is $\phi(\alpha, \beta) = f(x^*(\alpha, \beta), \alpha)$. Since the expression $f(x, \alpha) - \phi(\alpha, \beta)$ has an unconstrained maximum in α, we were able to derive the general comparative statics result for any particular scalar α

$$\sum_i f_{i\alpha} \frac{\partial x_i}{\partial \alpha} \geq 0 \tag{7-10}$$

Suppose now an additional constraint, $h(x) = 0$, that is consistent with the original equilibrium is added to the model. That is, defining $x^0 = x^*(\alpha^0, \beta^0)$, we require that $h(x^0) = 0$. We say this constraint is *just binding,* because it does not disturb the original maximum position. However, it does affect the rates of change of the decision variables as the parameter changes. Let us denote the new choice functions, which are solutions to the original first-order conditions and $h(x) = 0$ also,

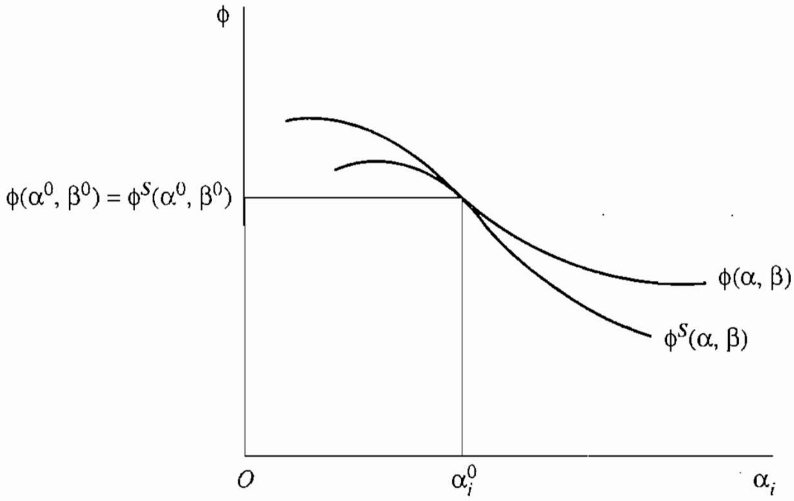

FIGURE 7-4
The indirect objective functions $\phi(\alpha, \beta)$ and $\phi^S(\alpha, \beta)$ plotted against α, $\phi^S(\alpha, \beta)$ being the indirect objective function when the just binding constraint $h(x) = 0$ is added. The constraint $h(x) = 0$ is added so as not to disturb the solution $x^0 = x^*(\alpha^0, \beta^0)$. By this construction, $\phi(\alpha, \beta) = \phi^S(\alpha, \beta)$ when $\alpha = \alpha^0$, and $\phi(\alpha, \beta) > \phi^S(\alpha, \beta)$ in any neighborhood of α^0. Therefore, the function $F(\alpha, \beta) = \phi(\alpha, \beta) - \phi^S(\alpha, \beta)$ has an unconstrained minimum with respect to α at α^0. It follows that $\phi(\alpha, \beta)$ is tangent to $\phi^S(\alpha, \beta)$ at α^0, and $\phi(\alpha, \beta)$ is relatively more convex or less concave than $\phi^S(\alpha, \beta)$ in a neighborhood of α^0. This implies that $\phi_{\alpha_i\alpha_i} \geq \phi^S_{\alpha_i\alpha_i}$ in a neighborhood of α^0.

as $x^s(\alpha, \beta)$ (s for "short run"), and the new indirect objective function as $\phi^s(\alpha, \beta)$. We show these curves in Fig. 7-4. By construction, when $\alpha = \alpha^0$ and $\beta = \beta^0$, $\phi = \phi^s$, but for $\alpha \neq \alpha^0$ or $\beta \neq \beta^0$, $\phi > \phi^s$. Equivalently, the function $F(\alpha, \beta) = \phi - \phi^s$ has an unconstrained minimum value (of zero) at (α^0, β^0) with respect to *both* α and β as long as these parameters are not in the auxiliary constraint. Assuming differentiability of these functions, this means that ϕ^s is relatively more concave than ϕ. The implied necessary first-order conditions are

$$F_\alpha = \phi_\alpha - \phi^s_\alpha = 0 \tag{7-29a}$$

$$F_\beta = \phi_\beta - \phi^s_\beta = 0 \tag{7-29b}$$

The necessary second-order condition is that the matrix $F_{\alpha\beta}$ of second partials with respect to α and β is positive semidefinite. This condition implies that the submatrices $F_{\alpha\alpha}$ and $F_{\beta\beta}$ are positive semidefinite as well, and thus the diagonal elements of those matrices are nonnegative. Thus for any particular scalar parameter α,

$$F_{\alpha\alpha} = \phi_{\alpha\alpha} - \phi^s_{\alpha\alpha} \geq 0 \tag{7-30}$$

Using the analysis leading up to (7-10), this yields

$$\sum_{i=1}^{n} f_{i\alpha}\left(\frac{\partial x_i^*}{\partial \alpha} - \frac{\partial x_i^s}{\partial \alpha}\right) \geq 0 \tag{7-31}$$

Although (7-31) summarizes the available comparative statics Le Châtelier results for the α parameters, the most useful results occur when the conditions of the conjugate pairs theorem hold, i.e., when some particular α enters only the ith first-order equation. In that case, (7-31) reduces to one term, yielding

$$f_{i\alpha}\frac{\partial x_i^*}{\partial\alpha} \geq f_{i\alpha}\frac{\partial x_i^s}{\partial\alpha} \geq 0 \tag{7-32}$$

Since $f_{i\alpha}$ can be negative, we cannot simply cancel this term out. However, since $\partial x_i^*/\partial\alpha$ and $\partial x_i^s/\partial\alpha$ have the same sign as $f_{i\alpha}$, the response of x_i to a change in α is always greater in *absolute value* in the absence of an auxiliary constraint:

$$\left|\frac{\partial x_i^*}{\partial\alpha}\right| \geq \left|\frac{\partial x_i^s}{\partial\alpha}\right| \tag{7-33}$$

The Le Châtelier results are usually stated in terms of the effects of holding one of the choice variables constant. We see here that this is unnecessarily restrictive. The only important restriction on the auxiliary constraint is that it cannot incorporate the parameters in question. The Le Châtelier results thus hold for constraints more complicated than simply $x_n = x_n^0$. Moreover, the process can be repeated as additional just-binding constraints are added.

The β parameters generally do not yield a simple result such as (7-32), since an expression in the Lagrange multiplier is always present. Consider, however, the important special case of models in which the constraint takes the form $g(x) = k$. Define the Lagrangian for this model as $\mathcal{L} = f(x,\alpha) + \lambda(k - g(x))$ and assume unique interior solutions $x^*(\alpha, k)$ and $\lambda^*(\alpha, k)$. Let $\phi(\alpha, k)$ be the indirect objective function. From (7-27), $\phi_k = \lambda^*(\alpha, k)$. We know from general comparative statics analysis that $\partial\lambda^*/\partial k \gtrless 0$. Curiously enough, however, a systematic prediction is available for the Le Châtelier effects.

Add an additional nonbinding constraint $h(x) = 0$ as before. Let $\phi^s(\alpha, k)$ be the indirect objective function when this new constraint is added, and let $\lambda^s(\alpha, k)$ be the resulting solution for the Lagrange multiplier for the constraint $g(x) = k$. The function $\phi - \phi^s$ has an unconstrained minimum with respect to k. The necessary first-order conditions are $\phi_k - \phi_k^s = 0$, i.e., that $\lambda^* = \lambda^s$. The second-order condition says that $\phi_{kk} - \phi_{kk}^s \geq 0$, and so

$$\frac{\partial\lambda^*}{\partial k} \geq \frac{\partial\lambda^s}{\partial k} \tag{7-34}$$

Thus even at this rather general level, even though both terms in (7-34) are unsigned by maximization, it is still the case that a smaller change in λ occurs when k changes when an auxiliary constraint is added to the model. In the next chapter we study the cost minimization model; the Lagrange multiplier turns out to be the marginal cost function. This result says that even though minimization does not imply a sign for the slope of the marginal cost function, it is nonetheless true that the marginal cost function either rises faster or falls slower in the short run than in the long run.

PROBLEMS

1. Consider maximization models with the specification

maximize

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

subject to

$$g(x_1, x_2) = k$$

with Lagrangian $\mathcal{L} = f(x_1, x_2, \alpha) + \lambda[k - g(x_1, x_2)]$, where x_1 and x_2 are choice variables and α and k are parameters.

(a) Define $\phi(\alpha, k)$ = maximum value of y for given α and k in this model. On a graph with α on the horizontal axis and ϕ and f on the vertical axis, explain *geometrically* the envelope results $\phi_\alpha = f_\alpha$ and $\phi_{\alpha\alpha} > f_{\alpha\alpha}$.

(b) On a similar graph, explain why it is *not* possible to carry out a similar procedure for the parameter k. How does this result relate to the appearance of refutable comparative statics theorems in economics?

(c) Using the results of *(a)*, prove that

$$f_{1\alpha}\frac{\partial x_1}{\partial \alpha} + f_{2\alpha}\frac{\partial x_2}{\partial \alpha} > 0$$

(d) Using the primal-dual methodology, prove algebraically the envelope theorem results:

(i) $\phi_\alpha = f_\alpha$

(ii) $\phi_{\alpha\alpha} > f_{\alpha\alpha}$

(iii) $\phi_k = \lambda^*$

(e) Prove that $f_{1\alpha}(\partial x_1^*/\partial k) + f_{2\alpha}(\partial x_2^*/\partial k) = \partial \lambda^*/\partial \alpha$.

(f) Assume that the objective function f measures the net value of some activity, and the constraint represents a restriction on some resource. Using result *(iii)* in part *(d)*, explain why the Lagrange multiplier imputes a shadow price to the resource, i.e., a marginal value of that resource in terms of the objective specified in the model. Also, in these models, what can be said, if anything, about how this marginal evaluation of the resource changes as the constraint eases, i.e., as k increases?

(g) Suppose now that the objective function is linear in α, i.e., $f(x_1, x_2, \alpha) = h(x_1, x_2) + \alpha x_1$. Prove that $\phi(\alpha)$ is convex in α, and, assuming the sufficient second-order conditions hold, $\phi_{\alpha\alpha} > 0$.

2. Consider models with the specification

maximize

$$y = f(x_1, x_2) + h(x_1, \alpha)$$

subject to

$$g(x_1, x_2, \beta) = 0$$

where x_1 and x_2 are choice variables and α and β are parameters that enter only the functions shown.

(a) Derive a refutable comparative statics result for α, and show that no such result exists for β.

(b) Let $\phi(\alpha, \beta)$ = maximum value of y for given α and β in this model. Using the primal-dual methodology, prove the envelope theorem results:

(i) $\phi_\alpha = h_\alpha(x_1^*, \alpha)$

(ii) $\phi_\beta = \lambda^* g_\beta(x_1^*, x_2^*, \beta)$, where λ^* is the Lagrange multiplier.

(c) Prove the "reciprocity" theorem

$$h_{1\alpha}\left(\frac{\partial x_1^*}{\partial \beta}\right) = \lambda^*\left[g_{1\beta}\left(\frac{\partial x_1^*}{\partial \alpha}\right) + g_{2\beta}\left(\frac{\partial x_2^*}{\partial \alpha}\right)\right] + g_\beta\left(\frac{\partial \lambda^*}{\partial \alpha}\right)$$

(d) On a graph with y on the vertical axis and α on the horizontal axis, sketch possible curves $\phi(\alpha, \beta)$ and $f(x_1^0, x_2^0) + h(x_1^0, \alpha)$ where x_1^0, x_2^0 and β^0 are some fixed values of those variables. Demonstrate graphically that $\phi_\alpha = h_\alpha$ and also that $\phi_{\alpha\alpha} > h_{\alpha\alpha}$.

(e) Explain why it is not possible to carry out a similar procedure for the parameter β, and thus why no refutable comparative statics theorems are available for this parameter from maximization alone.

3. Consider the model,

minimize

$$\text{AC} = \frac{w_1 x_1 + w_2 x_2}{y}$$

where x_1 and x_2 are factor inputs, w_1 and w_2 are factor prices, and $y = g(x_1, x_2)$ is a production function. Let $\text{AC}^*(w_1, w_2)$ be the minimum average cost for given factor prices.

(a) Explain how the factor demands $x_i^*(w_1, w_2)$ and the indirect objective function are derived. Prove that the factor demands are homogeneous of degree 0 and that AC* is homogeneous of degree 1 in the factor prices.

(b) On a graph with AC and AC* on the vertical axis, and w_1 on the horizontal axis, plot a typical AC and AC*. Show graphically that AC* is necessarily concave in w_1 (and, of course, w_2 also.)

(c) What is the slope of AC* at any given w_1?

(d) Using this graphical analysis, show that $\partial(x_i^*/y^*)/\partial w_i < 0$.

(e) Show that the elasticity of demand for factor 1 is less than the elasticity of output supply with respect to w_1.

(f) Set up the primal-dual model, minimize AC − AC*, and derive the above results algebraically.

(g) Contrast the factor demands derived from this model, $x_i^*(w_1, w_2)$, with the factor demands $x_i^p(w_1, w_2, p)$ derived from, maximize $pf(x_1, x_2) - w_1 x_1 - w_2 x_2$, where output price p is parametric. Display the first-order conditions for both models, and explain the relation between the models by explaining the following identity, where $p^* = \text{AC}^*(w_1, w_2)$:

$$x_1^*(w_1, w_2) \equiv x_1^p(w_1, w_2, p^*(w_1, w_2))$$

(h) From this identity, show that the elasticity of demand for x_1 derived from min AC, $[(w_1/x_1^*)(\partial x_1^*/\partial w_1)]$ is equal to the elasticity of demand derived from profit maximization, plus an output effect which equals the share spent on x_1 times the output price elasticity of x_1.

4. Consider a profit-maximizing firm employing two factors. Define the short run as the condition where the firm behaves as if it were under a total expenditure constraint; i.e., in the short run, *total expenditures are fixed* (at the long-run profit-maximizing level). The long run is the situation where no additional constraints are placed on the firm.

(a) Are these short-run demands necessarily downward-sloping?

(b) Show that the short-run factor demand curves for this model are not necessarily less elastic than the long-run factor demand curves. Why does this anomalous result arise for this model?

(c) Show that if a factor is *inferior* in terms of its response to a change in total expenditure, the slope of the long-run factor demand is necessarily more negative than the short-run demand for that factor.

5. Consider models with the specification

maximize

$$y = f(x_1, \dots, x_n)$$

subject to

$$g(x_1, \dots, x_n) = k$$

Let $\phi(k)$ = maximum value of f for given k. Assuming an interior solution exists, prove that if f and g are both homogeneous of the same degree r, then $\phi(k)$ is linear in k, i.e., $\phi(k) = ak$, where a is an arbitrary constant, and thus the Lagrange multiplier for such models is a constant.

BIBLIOGRAPHY

Samuelson, P. A.: "The Le Châtelier Principle in Linear Programming," *RAND Corporation Monograph,* August 4, 1949 (Chap. 43 in *Scientific Papers* below).

———: "An Extension of the Le Châtelier Principle," *Econometrica,* pp. 368–379, April 1960 (Chap 42 in *Scientific Papers* below).

———: "Structure of a Minimum Equilibrium System," In R. W. Pfouts (ed.), *Essays in Economics and Econometrics: A Volume in Honor of Harold Hotelling,* The University of North Carolina Press, Chapel Hill, 1960 (Chap. 44 in *Scientific Papers* below).

These three articles have all been reprinted in J. Stiglitz (ed.): *The Collected Scientific Papers of Paul A. Samuelson,* The M.I.T. Press, Cambridge, MA, 1966.

Samuelson, P. A.: *Foundations of Economic Analysis,* Harvard University Press, Cambridge, MA, 1947.

Silberberg, E.: "A Revision of Comparative Statics Methodology in Economics, or, How to Do Economics on the Back of an Envelope," *Journal of Economic Theory,* **7**:159–172, February 1974.

———: "The Le Châtelier Principle as a Corollary to a Generalized Envelope Theorem," *Journal of Economic Theory,* **3**:146–155, June 1971.

Viner, J.: "Cost Curves and Supply Curves," *Zeitschrift fur Nationalokonomie,* **3**:1932. Reprinted in American Economic Association, *Readings in Price Theory,* Richard D. Irwin, Chicago, 1952.