Producer Theory

Jonathan Levin and Paul Milgrom October 2004

1 Competitive Producer Behavior

Since Marshall, the standard approach to developing a theory of competitive markets is to separate demand behavior ("consumer theory") from supply behavior ("producer theory") and then use the notion of market equilibrium to reconcile demand and supply. This note studies producer theory and a separate one studies consumer theory.

The standard model has the following features. Firms are described by fixed and exogenously given technologies that allow them to convert inputs (in simple models, these are land, labor, capital and raw materials) into outputs (products). "Competitive" producers take both input and output prices as given, and choose a production plan (a technologically feasible set of inputs and outputs) to maximize profits.

Before we get into the details, let's remark on a few key features of the model.

1. Firms are price takers. This "competitive firm" assumption applies to both input and output markets and makes it reasonable to ask questions about (1) what happens to the firm's choices when a price changes and (2) what can be inferred about a firm's technology from its choices at various price levels. For output markets, the assumption fits best when each firm has many competitors who produce perfectly substitutable products, and a parallel condition applies to input markets. Of course, even the most casual empiricism suggests that many firms sell differentiated products and have at

least some flexibility in setting prices, and even small firms may have market power in buying local inputs, such as hiring workers who live near a mine or factory, so the results of the theory need to be applied with care. Even so, the pattern of analysis established in this way is often partially extendable to situations in which firms are not price takers.

- 2. Technology is exogenously given. This assumption is sometimes criticized as too narrow to be useful in a world of technical change, product innovations, and consumer marketing, but it is more flexible and encompassing than most critics acknowledge. The exogenous technology model formally includes the possibility of investing in technical change, provided these investments are themselves treated as inputs into a production process. Similarly, the model formally includes advertising and branding that alter consumer's perceptions, provided that we represent these activities as transforming the output into a different product. It allows managerial effort and talent to be inputs as well, if they, too, are treated as simple inputs into production.
- 3. The firm maximizes profits. Since the time of Adam Smith, if not earlier, many observers have emphasized that corporations are characterized by a separation between ownership (the stockholders) and control (management), and that this separation weakens the incentives of managers to maximize profits. The problem of motivating managers to act on behalf of owners has been a main concern for the economics (and law) of agency theory.
- 4. The Marshallian approach of separating the household, where consumption takes place, from the firm, in which all production takes place, is intensely criticized by some economists. An alternative approach treats households as both consumers and producers. In this alternative view, an essential feature of all economic development is the change in household productive behavior associated with the development of markets. As markets develop, households move away from the pattern of producing for themselves only the goods they plan to consume toward a pattern in which each household specializes, devoting most of its productive efforts to producing one particular good, which it sells or trades for other goods.

Students sometimes wonder about the role of assumptions such as these, particularly when they are contrary to the facts of the situation. Economists have taken a range of positions concerning how to think about simplifying assumptions, and there is no consensus about the "correct" view. One extreme position is to deny the relevance of any inference based on such models, because the premises of the model are false. At the opposite extreme, some practicing economists seem willing to accept "standard" or "customary" assumptions uncritically. Both of these extreme positions are rejected by thoughtful people.

All economic modeling abstracts from reality by making simplifying but untrue assumptions. Experience in economics and other fields shows that such assumptions models can serve useful purposes. One purpose is to support tractable models that isolate and highlight important effects for analysis by suppressing other effects. Another purpose is to serve as a basis for numerical calculations, possibly for use in estimating magnitudes, deciding economic policies, or designing economic institutions. For example, one might want to estimate the effect of a tax policy change on overall investment or hiring. The initial calculations based on a simplified model might then be adjusted to account for the effects suppressed in the model.

For a model to serve these practical purposes, its relevant predictions must be reasonably accurate. The accuracy of predictions can sometimes be checked by testing using data. Sometimes, the "robustness" of predictions can be evaluated partly by theoretical analyses. In no case, however, should models or assumptions be regarded as adequate merely because they are "usual" or "standard." Although this seems to be an obvious point, it needs to be emphasized because the temptation to skip the validation step can be a powerful one. Standard assumptions often make the theory fall into easy, recognizable patterns, while checking the suitability of the assumptions can be much harder. The validation step is not dispensable.

2 Production Sets, Technology

We start by describing the technological possibilities of the firm. Suppose there are n commodities in the economy. A production plan is a vector $y = (y_1, ..., y_n) \in \mathbb{R}^n$, where an output will have $y_k > 0$ and an input will have $y_k < 0$. If the firm has

nothing to do with good k, then $y_k = 0$. The production possibilities of the firm are described by a set $Y \subset \mathbb{R}^n$, where any $y \in Y$ is feasible production plan. Figure 1 illustrates a production possibility set.

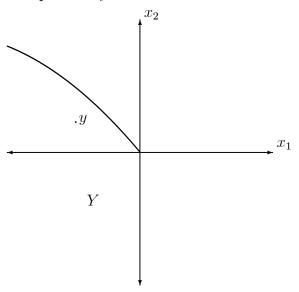


Figure 1: A Production Possibility Set

Throughout our analysis, we will make the innocent technical assumptions that Y is non-empty (so as to have something to study!) and closed (to make it more likely that optimal production plans exist). Consider the more interesting and substantive economic properties production sets might have:

- Free Disposal. The production set Y satisfies free disposal if $y \in Y$ implies that $y' \in Y$ for any $y' \leq y$.
- Shut Down. The production set Y has the shut-down property if $0 \in Y$; that is, the firm has the option of using no resources and producing nothing.
- Nonincreasing Returns to Scale. The production set Y has nonincreasing returns to scale (loosely, "decreasing returns to scale") if $y \in Y$ implies that $\alpha y \in Y$ for all $0 \le \alpha \le 1$.

- Nondecreasing Returns to Scale. The production set Y has nondecreasing returns to scale (loosely, "increasing returns to scale") if $y \in Y$ implies that $\alpha y \in Y$ for ally $\alpha \geq 1$.
- Constant Returns to Scale. The production set Y has constant returns to scale if $y \in Y$ implies that $\alpha y \in Y$ for all $\alpha \geq 0$.
- Convexity. The production set Y is convex if... Y is convex. This condition incorporates a kind of "nonincreasing returns to specialization," meaning that if two "extreme" plans are feasible, their combination will be as well. In addition, if $0 \in Y$, then convexity implies nonincreasing returns to scale.

Another way to represent production possibility sets is using a transformation function $T: \mathbb{R}^n \to \mathbb{R}$, where $T(y) \leq 0$ implies that y is feasible, and T(y) > 0 implies that y is infeasible. This is represented in Figure 2. You can think of the transformation function simply as a convenient way to represent a set. The set of boundary points $\{y \in \mathbb{R}^n : T(y) = 0\}$ is called the transformation frontier.

When the transformation function is differentiable, we can define the marginal rate of transformation between goods k and l as:

$$MRT_{k,l}(y) = -\frac{\partial T(y)/\partial y_l}{\partial T(y)/\partial y_k}.$$

The marginal rate of transformation measures the extra amount of good k that can be obtained per unit reduction of good k. As Figure 2 shows, it is equal to the slope of the boundary of the production set at point y.

Thinking in terms of production sets leads to a very general model where each good k can be either an input or an output — that is, a firm may both produce widgets, and also use widgets to make gadgets, with y_k being the net amount of widgets produced. Often, it is convenient to separate inputs and outputs, letting $q = (q_1, ..., q_L)$ denote the vector of the firm's outputs, and $z = (z_1, ..., z_M)$ the vector of inputs (where L + M = N).

¹Several interpretations can be offered of the function T(y). As just one example among many, one might interpret it to define the amount of technical progress required to make the combination y a feasible one. With that interpretation, using the currently available technology, one can produce any element of the set $\{y|T(y) \le 0\}$.

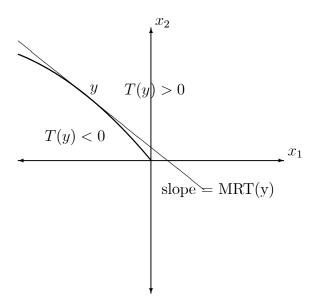


Figure 2: Marginal Rate of Transformation

If the firm has only a single output, we can write output as a function of the inputs used, q = f(z). In this case, we refer to $f(\cdot)$ as the firm's production function. We can also define the marginal rate of technological substitution to be:

$$MRTS_{k,l}(y) = -\frac{\partial f(z)/\partial z_l}{\partial f(z)/\partial z_k}$$

The marginal rate of technological substitution tells us how much of input k must be used in place of one unit of input l to maintain the same level of output. It is illustrated in Figure 3.

3 Profit Maximization

We write the profit maximization problem for the firm as:

$$\max_{y} p \cdot y$$

s.t. $y \in Y$

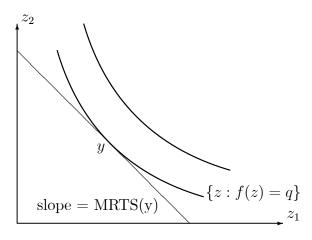


Figure 3: Isoquants and MRTS.

where we assume that $p \gg 0$.² The constraint that $y \in Y$ can be alternatively written as $T(y) \leq 0$.

There are two functions of special interest in studying the problem. The first is called the "optimal production correspondence" and is denoted by y(p). The correspondence $y: \mathbb{R}^n_+ \rightrightarrows Y$ maps a vector of prices into the set of profit-maximizing production plans. The second, called the "profit function," identifies the maximal value of the problem and is denoted by $\pi(p)$. That is, the profit function $\pi: \mathbb{R}^n_+ \to \mathbb{R}$ is defined to be:

$$\pi(p) = \max_{y \in Y} p \cdot y.$$

We now record some useful properties of the profit function and the optimal production correspondence. Recall that the vector notation p > p' is defined by the conjunction: $p \ge p'$ and $p \ne p'$.

Proposition 1 (Properties of π) The profit function π has the following properties:

²We have not yet made sufficient assumptions to ensure that a maximum exists, so it would be more proper to write $\sup_{y \in Y} p \cdot y$. The focus of our investigation here, however, will be on the properties of the maximum when it exists, rather than on conditions for a maximum to exist, so we will continue to use the "max" notation.

- 1. $\pi(\cdot)$ is homogeneous of degree one, i.e. for all $\lambda > 0$, $\pi(\lambda p) = \lambda \pi(p)$.
- 2. $\pi(\cdot)$ is convex in p.
- 3. If Y is closed and convex, then $Y = \{ y \in \mathbb{R}^n : p \cdot y \leq \pi(p) \text{ for all } p \in \mathbb{R}^N \}.$
- 4. If Y is closed and convex and has the free disposal property, then $Y = \{y \in \mathbb{R}^n : p \cdot y \leq \pi(p) \text{ for all } p \in \mathbb{R}^N_+\}.$
- **Proof.** (1) Note that $\pi(\lambda p) = \max_{y \in Y} \lambda p \cdot y = \lambda \max_{y \in Y} p \cdot y = \lambda \pi(p)$.
- (2) Fix p, p' and define $p^t = tp + (1 t)p'$ for $t \in [0, 1]$. And let $y^t \in y(p^t)$. Then

$$t\pi(p) + (1-t)\pi(p') \ge tp \cdot y^t + (1-t)p' \cdot y^t = p^t \cdot y^t = \pi(p^t).$$

- (3) Let $\widehat{Y} = \{x \in \mathbb{R}^n : p \cdot x \leq \pi(p) \text{ for all } p \in \mathbb{R}^n\}$. We need to show that $Y \subset \widehat{Y}$ and that $\widehat{Y} \subset Y$. The first inclusion follows from the definition of π . For the reverse inclusion, suppose that Y is closed and convex and $x \notin Y$. Then, by the separating hyperplane theorem, there exists $p \in \mathbb{R}^n$ such that $p \cdot x > \max_{y \in Y} p \cdot y = \pi(p)$. It follows that $x \notin \widehat{Y}$. The finding $x \notin Y \Longrightarrow x \notin \widehat{Y}$ establishes that $\widehat{Y} \subset Y$
- (4) We argue exactly as in (3), but with one additional step. Let $\widetilde{Y} = \{y \in \mathbb{R}^n : p \cdot y \leq \pi(p) \text{ for all } p \in \mathbb{R}^N_+\}$. We need to show that $Y \subset \widetilde{Y}$ and that $\widetilde{Y} \subset Y$. The first inclusion follows from the definition of π . For the reverse inclusion, suppose that Y is closed and convex and $x \notin Y$. Then, by the separating hyperplane theorem, there exists $p \in \mathbb{R}^n$ such that $p \cdot x > \max_{y \in Y} p \cdot y = \pi(p)$. By free disposal, if any component of p were negative, then $\sup_{y \in Y} p \cdot y = +\infty$. So, no component is negative, that is, $p \in \mathbb{R}^N_+$. It follows that $x \notin \widetilde{Y}$. The finding $x \notin Y \Longrightarrow x \notin \widetilde{Y}$ establishes that $\widetilde{Y} \subset Y$.

Proposition 2 (Properties of y) The optimal production correspondence y has the following properties:

- 1. $y(\cdot)$ is homogeneous of degree zero, i.e. for all $\lambda > 0$, $y(\lambda p) = y(p)$.
- 2. If Y is convex, then for all p, the set y(p) is convex. If Y is strictly convex, $p \neq 0$ and $y(p) \neq \emptyset$, then y(p) is a singleton.

3. The Law of Supply. For any $p, p', y \in y(p)$ and $y' \in y(p')$,

$$(p'-p)(y'-y) \ge 0.$$

Proof. (1) Note that $\pi(\lambda p) = \max_{y \in Y} \lambda p \cdot y = \lambda \max_{y \in Y} p \cdot y = \lambda \pi(p)$, so for $\lambda > 0$, $y(p) = \{y \in Y | p \cdot y = \pi(p)\} = \{y \in Y | \lambda p \cdot y = \pi(\lambda p)\} = y(\lambda p)$.

(2) Observe that $y(p) = Y \cap \{y \in \mathbb{R}^n_+ | p \cdot y = \pi(p)\}$. If Y is convex, then y(p) is the intersection of two convex sets and hence is itself convex.

Suppose Y is strictly convex but y(p) is not a singleton. Then for any $y \neq y' \in y(p)$, we have $y'' = \frac{1}{2}y + \frac{1}{2}y' \in interior(Y)$ and, since y(p) is convex, $y'' \in y(p)$. That's impossible, because a non-trivial linear function (one with $p \neq 0$) has no local maximum.

(3) Given any $p, p', y \in y(p)$ and $y' \in y(p')$, profit maximization at price vectors p and p' imply that $p \cdot y \geq p \cdot y'$ and $p' \cdot y' \geq p' \cdot y$, respectively. So, $p \cdot (y - y') \geq 0 \geq p' \cdot (y - y')$, from which conclusion follows. Q.E.D.

Next, we obtain two comparative statics results that relate the profit function to optimal production choices.

Proposition 3 (Local Properties) Assume that Y is closed and satisfies free disposal. Then,

1. Hotelling's Lemma (traditional form): If y is singleton-valued in a neighborhood of p, then $\pi(\cdot)$ is differentiable at p and:

$$\frac{\partial \pi(p)}{\partial p_i} = y_i(p).$$

2. Hotelling's Lemma (producer surplus): If $y(\cdot)$ is non-empty valued and $\widehat{y}(p) \in y(p)$ for all p, then for all p' and p''_i ,

$$\pi(p'_j, p'_{-j}) - \pi(p''_j, p'_{-j}) = \int_{p''_j}^{p'_j} \widehat{y}(s, p'_{-j}) ds.$$

3. If $y(\cdot)$ is singleton-valued and continuously differentiable, the matrix $D_p y(p) = D_p^2 \pi(p)$ is symmetric and positive semi-definite, with $[D_p y(p)]p = 0$.

Proof. For (1), notice that the condition that y is single-valued in a neighborhood of p implies the conditions of the envelope theorem, after which it follows immediately that $\partial \pi/\partial p_i = y_i(p)$. Conclusion (2) is immediate from the envelope theorem. The first part of (3), namely that $D_p^2\pi(p)$ is symmetric and positive semi-definite, follows from the convexity of π . For the second part of (3), observe that since y(p) solves $\max_{y\in Y} p\cdot y$, it follows that p solves $\max_{p'\in\mathbb{R}^n} p\cdot y(p')$. The first-order optimality condition for that latter problem is $[D_p y(p)]p = 0$. Q.E.D.

The traditional form of Hotelling's lemma allows us to recover the firm's choices from the profit function. The alternative form recovers the profit function from the choices. The latter form uses weaker assumptions and allows us to recover the profit function even if the choice function is not continuous. The symmetry of the matrix $D_p y(p)$ is a subtle empirical implication of optimization theory that was missed by economists working in a verbal tradition. Historically, this conclusion was argued to be important evidence that a mathematical approach to economic theory could lead to new insights that would be missed by a merely verbal approach.

4 Cost Minimization with a Single Output

Suppose that the firm produces a single output whose quantity is denoted by q. Using the production function notation, the firm's cost minimization problem can be written as:

$$\min_{z \in \mathbb{R}^n_+} w \cdot z$$

s.t. $f(z) \ge q$

In our analysis of this problem, we will always assume that all input prices are strictly positive: $w \gg 0$. As usual in our study of optimization problems, two functions are of central interest. One is the solution as z(q, w). We refer to z(q, w) as the *conditional factor demand* to indicate that it is conditional on a fixed level of output q. The second is the optimal value function. The optimal value function for this problem is:

$$c(q,w) = \min_{\{z: f(z) \geq q\}} w \cdot z,$$

It is called the cost function gives the minimum cost at which output q can be produced.

If f(z) is differentiable and concave, we can use the Kuhn-Tucker method to solve for the conditional factor demands. The Lagrangian for the cost problem is:

$$\max_{\lambda \ge 0, \mu \ge 0} \min_{z} w \cdot z - \lambda \left[f(z) - q \right] - \sum_{i=1}^{m} \mu_{i} z_{i}.$$

The first-order conditions from the Lagrangian problem are:

$$\lambda \frac{\partial f(z)}{\partial z_i} \le w_i$$
 with equality if $z_i > 0$.

and of course the solution must satisfy the production constraint. $f(z) \geq q$.

Later, we will compare these first-order conditions to the ones arising from the profit maximization problem. In the meantime, we record a few properties of the cost function. (For a full recital, including properties of the conditional factor demands, see MWG, Proposition 5.C.2.)

Proposition 4 (Properties of c) The cost function c has the following properties:

- 1. $c(\cdot)$ is homogeneous of degree one in w, and increasing in q.
- 2. $c(\cdot)$ is a concave function of w.
- 3. If $f(\cdot)$ is concave, then $c(\cdot)$ is a convex function of q (i.e. marginal costs are increasing in q).
- 4. Shepard's Lemma: If $z(\cdot)$ is single-valued, then $c(\cdot)$ is differentiable with respect to w and

$$\frac{\partial c(q, w)}{\partial w_i} = z_i(q, w)$$

5. If $z(\cdot)$ is a differentiable function, then the matrix $D_w z(q, w) = D_w^2 c(q, w)$ is symmetric and negative semi-definite, and $D_w z(w, q) w = 0$.

Proof. Left as an exercise.

Q.E.D.

Returning to our characterization of the firm's problem, suppose the firm solves the cost minimization problem for every q, giving it a cost function c(q, w). The profit maximization problem can then be seen as:

$$\max_{q} pq - c(q, w).$$

This problem gives the famous first-order conditions:

$$p = \frac{\partial c(q, w)}{\partial q},$$

equating price and marginal cost. So profit maximization implies that the correct shadow price is the market price for output p.

Finally, we state a result about how conditional factor demands will change with prices.

Proposition 5 Conditional factor demands are decreasing in own price: for all $i, q, z_i(q, w)$ is nonincreasing in w_i .

Proof. The proof mimics the argument for the law of supply; try it yourself. Q.E.D.

5 Profit Maximization with a Single Output

With a vector $z = (z_1, ..., z_M)$ of inputs and a single output f(z), the profit maximization problem can be simplified to:

$$\max_{z \in \overline{\mathbb{R}}_{+}^{m}} p \cdot f(z) - w \cdot z,$$

where p reflects the price of output and the vector $w \gg 0$ reflects the input prices. Denote the maximizing solution by z(p, w). There may be multiple solutions or no solution for some price vectors (p, w), but we will focus our analysis on the cases where the solution is a singleton, in which case z(p, w) denotes the factor demands at prices (p, w). If f(z) is differentiable and concave, the factor demands (when they exist!) can be found by solving the first-order conditions: for all i,

$$p \frac{\partial f(z)}{\partial z_i} \leq w_i$$
 with equality if $z_i > 0$.

It is interesting to compare this solution to that of the cost minimization problem. There, the first order conditions were that

$$\lambda \frac{\partial f(z)}{\partial z_i} \leq w_i$$
 with equality if $z_i > 0$

Thus, with f concave, one can think of profit maximization as the special case of cost minimization in which the shadow price of output is the market price p. There is more to this account. From the envelope therem, we have:

$$\lambda = \frac{\partial c(q, w)}{\partial q}.$$

Thus, at the solution to the cost minimization problem, the shadow value of output λ is exactly the marginal cost of production.

When f is concave, the approach based on first-order conditions is useful for working examples and obtaining formula that can be used to compute solutions numerically. The convexity assumption fails in several interesting cases, such as ones where there are fixed costs of production or where the production sets exhibits increasing returns. It turns out that comparative statics conclusions are largely independent of convexity assumptions, so we approach the problem of comparative statics using methods that do not rely on convexity. The first result is an easy one that follows from the law of supply.

Proposition 6 Restrict attention to the domain of price vectors (p, w) upon which z(p, w) is singleton-valued. Then, for all $i, z_i(p, w)$ is non-increasing in w_i .

Proof. This is an immediate consequence of the Law of Supply. Q.E.D.

6 Complements and Substitutes

Informally, two inputs are called *substitutes* when an increase in the price of one leads to an increase in input demand for the second and *complements* when it leads to a decrease in input demand for the second. Several things conspire to complicate this seemingly simple definition. First, it is perfectly possible that the change in demand in response to a price increase is not uniform; for example, the demand $z_j(p, w)$ for input j may increase as the price of input i increases from w_i to w'_i and may then decrease as the input price increases further to w''_i . Second, the response to a price increase can depend on which optimization problem we are using to determine demand. Input demands might correspond to the solution of a cost minimization problem in which output q is held constant, or from the long-run profit maximization, with all inputs and output free to vary, or a short-run profit maximization problem, with some inputs fixed.

The first result reported below is for the profit maximization problem with a single kind of output and all inputs free to vary.

Proposition 7 Restrict attention to the domain of price vectors $(p, w) \in \mathbb{R}^{n+1}_+$ upon which z(p, w) is singleton-valued. If f(z) is increasing and supermodular, then z(p, w) is isotone ("weakly increasing") in p and antitone ("weakly decreasing") in w.

Proof. Since $f(\cdot)$ is increasing and supermodular, the firm's objective function $p \cdot f(z) - w \cdot z$ is supermodular in (z, p). Also, the choice set \mathbb{R}^n_+ is a sublattice. So by Topkis' Monotonicity Theorem, z(p, w) must be increasing in p. Similarly, the firm's objective is also supermodular in in $(z, -w_i)$. So z(p, w) is antitone in w_i . Q.E.D.

The preceding proposition is easily extended to "short run" profit maximization. For consider the problem in which some set of inputs S is held fixed at the levels x_S . Define $z(p, w, x_S)$ to be the solution to the firm's profit maximization problem given the extra constraint $z_S = x_S$. This additional constraint defines a sublattice, so the original proof still applies.

Supermodularity is stronger than the long-run price-theory notion of complementary, because it implies the price theory concept not only for the long-run problem but also for all possible short-run problems. It is stronger in another way, as well: it characterizes the behavior of f even around choices z that would never be justified by any price vector. The next theorem asserts that when f is strictly concave, so each choice is the unique optimum for some set of prices, then supermodularity is identical to this notion of long-run and short-run price theory complementarity.

Proposition 8 Suppose that f is increasing and strictly concave. If for all S and x_S , $z(p, w, x_S)$ is antitone in w, then f is supermodular.

Proof. Left as an exercise. (Hint: Suppose that all but 2 inputs are fixed. How does f vary in the remaining two inputs?) Q.E.D.

By Topkis's Monotonicity Theorem, if f is supermodular, then $z(p, w, x_S)$ is antitone in w, for all S, even without the assumptions that f is increasing and concave. So, the import of the proposition is that for the case of an increasing concave production function f, inputs are complements in the strong sense defined by the proposition if and only if f is supermodular.

The substitutes case is similar for the two-input case, but subtler for the general cases. Here are the relevant propositions.

Proposition 9 Restrict attention to the domain of price vectors (p, w) upon which z(p, w) is singleton-valued and suppose there are just two inputs. If -f(z) is supermodular, then $z_1(p, w)$ is everywhere non-decreasing in p_2 and $z_2(p, w)$ is everywhere non-decreasing in p_1 .

Proof. With two inputs, define $\widehat{f}(z_1, -z_2) = f(z)$. If -f is supermodular, then the function \widehat{f} is also supermodular. The conclusion then follows directly by applying Topkis's Monotonicity Theorem to the problem: $\max_{x_1 \geq 0, x_2 \leq 0} p\widehat{f}(x) - w_1x_1 + w_2x_2$. Q.E.D.

Note well that this characterization applies only to the two-input case. For the case of multiple inputs, one can obtain characterizations of substitutes using the indirect profit function or the cost function, as follows.

Proposition 10 Suppose that the cost function c(q, w) is continuously differentiable. Then, the following two are equivalent: (1) for all $i \neq j$ and all w_{-j} and q, $z_i(q, w)$ is non-increasing in w_j , and (2) for all q, c(q, w) is submodular in w.

Proof. By Shepard's lemma, $z_i(q, w) = -\frac{\partial}{\partial w_i}c(q, w)$, and the right-hand-side is always non-increasing in w_j if and only if c is submodular in w. Q.E.D.

Recall that we have used the notation z(q, w) to refer to the solution of the cost minimization problem, with output fixed, and z(p, w) to refer to the solution of the profit maximization problem, with output free to vary. Either of these optimization problems can be used to define the term "substitutes," so we have avoided that term in stating the proposition above. A parallel proposition for the alternative optimization problem is stated below. It is proved similarly, using Hotelling's lemma in place of Shepard's lemma.

Proposition 11 Suppose that the profit function $\pi(p, w)$ is continuously differentiable. Then, the following two are equivalent: (1) for all $i \neq j$ and all w_{-j} and p, $z_i(p, w)$ is non-increasing in w_j , and (2) for all p, $\pi(p, w)$ is supermodular in w.

7 The Short-Run and Long-Run

While not treating time explicitly, the neoclassical theory of the firm typically distinguishes between the *long-run*, a length of time over which the firm has the opportunity to adjust all factors of production, and the *short-run*, during which time some factors may be difficult or impossible to adjust.

In his Foundations of Economic Analysis (1947), Samuelson suggested that a firm would react more to input price changes in the long-run than in the short-run, because it has more inputs that it can adjust. This view still persists in

some economics texts.³ Samuelson called this effect the *LeChatelier principle* and argued that it also illuminates how war-time rationing makes demand for non-rationed goods less elastic. Assuming that the optimal production choice y(p) is differentiable, he proved that the principle holds for sufficiently small price changes in a neighborhood of the long-run price. The relation between long and short run effects can be quite important, because data about the short-run effects of policies are frequently used to forecast their long-run effects, and such forecasts can influence policymakers.

We begin our analysis with an example to prove that the Samuelson-LeChatelier principle does not apply to large price changes. Consider the production set $Y = \{(0,0,0), (1,-2,0), (1,-1,-1)\}$ in which two inputs, goods 2 and 3, are used to produce good 1. The set includes the possibilities that the firm can produce a unit of output either by using two units of good 2 or by using one unit of each kind of input. The third and last possibility is that the firm can use no inputs and produce nothing.

Suppose long-run prices are given initially by p=(2,.7,.8). At the corresponding initial long-run optimum, the firm achieves its maximum profit of 0.6 by choosing the point $x^{LR}(p)=(1,-2,0)\in Y$. The superscript designates this as a long-run choice—one that treats all points in the set Y as feasible. Suppose that the use of good 3 is fixed in the short run and that the price of good 2 rises to 1.1, so the new price vector is p'=(2,1.1,.8). Since the firm cannot immediately change its use of good 3, it must choose between its current plan (1,-2,0) incurring a loss (profit=-0.2) or switching to (0,0,0) (profit=0). The latter choice maximizes its profit, so $x^{SR}(p',p)=(0,0,0)$, where the notation indicates that this is the firm's profit-maximizing short-run choice when current prices are p' but fixed inputs were chosen when prices were p. The firm's long run choice at price vector p' is $x^{LR}(p')=(1,-1,-1)$. In this example, when the price of the first input rises, the demand for good 2 changes in the short-run from $x_2^{LR}(p)=-2$ to $x_2^{SR}(p',p)=0$, but then recovers in the long-run to $x_2^{LR}(p')=-1$. So, the

³For example, Varian (1992) writes: "It seems plausible that the firm will respond more to a price change in the long run since, by definition, it has more factors to adjust in the long run than in the short run. This intuitive proposition can be proved rigorously."

short-run change is *larger* than the long-run change, contrary to the Samuelsonian conclusion.

Although the three-point production set may seem unusual, the example can be modified to make Y convex and smooth. The first step is to replace Y by its convex hull \widehat{Y} (the triangle with vertexes at the three points in the original set Y). The choices from \widehat{Y} are the same as those from Y, so that gives us a convex model of the same choices. One can further expand \widehat{Y} by adding free disposal without changing the preceding calculations. For a similar example with a strictly convex production set having a smoothly curved boundary, one can replace \widehat{Y} by the set $\widehat{Y}_{\epsilon} \subset \{y \in \mathbb{R}^3_+ | (\exists x \in \widehat{Y}) | y - x | \leq \varepsilon\}$, where $\varepsilon > 0$. For ε small, the firm's profit-maximizing choices from \widehat{Y}_{ϵ} differ little from its profit-maximizing choices from Y. In particular, the long-run response to the price change will remain smaller than the short-run response.

There is an interesting set of economic models in which it is always true that long-run responses to price changes are larger than short run responses. Intuitively, these are models in which a "positive feedbacks" argument applies, as follows.

Suppose that output is not fixed and that there are two inputs, capital and labor, which are substitutes. Suppose that capital is fixed in the short-run. By the law of demand, if the wage increases, the firm will use less labor both in the short-run and in the long-run. Since the two inputs are substitutes, the increased wage implies an increased use of capital in the long-run. Since $f_{kl} \leq 0$ for substitutes, the additional capital used in the long-run will reduce the marginal product of labor, so in the long-run the firm will use still less labor. In summary, the long-run effect is larger than the short-run effect because, in the short-run the firm responds only to a higher wage, but in the long-run, it responds both to a higher wage and to an increased capital stock that reduces marginal product of labor. Graphically, the additional effect in this example can be represented by a positive feedback loop, as follows.

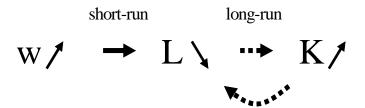


Figure 4: Positive feedbacks when capital and labor are substitutes. If the inputs are complements, the same figure applies with capital decreasing.

Next, suppose that the two inputs, capital and labor, are complements. Again, by the law of demand, if the wage increases, the firm will use less labor input, both in the short-run and in the long-run. Since the inputs are complements, the increased wage implies a reduced use of capital in the long-run. Since $f_{kl} \geq 0$ for complements, the reduced capital used in the long-run will reduce the marginal product of labor, so in the long-run the firm will use still less labor. Again, we have a positive feedback loop.

The general positive feedback argument for two inputs (due to Milgrom and Roberts (1996)) goes as follows. Let X and Y be sublattices (for example, let $X = Y = \mathbb{R}$). Define:

$$x(y,t) = \arg\max_{x \in X} g(x,y,t)$$

and

$$y(t) = \arg\max_{y \in Y} g(x(y, t), y, t).$$

Proposition 12 Suppose that $g: X \times Y \times \mathbb{R} \to \mathbb{R}$ is supermodular, that $t' \geq t$, and that the maximizers described below are unique for the parameter values t and t'. Then:

$$x(y(t'), t') \ge x(y(t), t') \ge x(y(t), t).$$

and

$$x(y(t'), t') \ge x(y(t'), t) \ge x(y(t), t).$$

Proof. By Topkis' Theorem, the function y(t) is isotone ("weakly increasing"). Then, since $t' \geq t$, $y(t') \geq y(t)$. Similarly, by Topkis's Theorem, the function x(y,t) is isotone. The claims in the theorems follow immediately from that and the inequalities $t' \geq t$ and $y(t') \geq y(t)$.

Q.E.D.

Now let's apply the result. Let x be labor input, and y capital input, and let $t = -w_x$, where w_x is the price of labor. The firm's objective is to maximize:

$$g(x, y, t) = pf(x, y) - w_x x - w_y y.$$

If capital and labor are "complements" in the sense that $f_{xy} \geq 0$, then the firm's objective is supermodular in $(x, y, -w_x)$, because it verifies all the pairwise supermodularity conditions. Similarly, if capital and labor are "substitutes" in the sense that $f_{xy} \leq 0$, then the firm's objective is supermodular in $(x, -y, -w_x)$. We then have the following result.

Corollary 1 (LeChatelier Principle) Suppose production is given by f(k,l), where either $f_{kl}(k,l) \geq 0$ for all (k,l) or $f_{kl}(k,l) \leq 0$ for all (k,l). Then if the wage w_l increases (decreases), the firm's labor demand will decrease (increase), and the decrease (increase) will be larger in the long-run than in the short-run.

Returning to our example above, consider the following three price vectors. The first two are p and p' as defined above and the third is p'' = (2, 1.1, 1.1). In the example, the inputs act like substitutes when prices change from p to p' (the long-run input demand for good 3 rises with this increase in the price of good 2) but they act like complements when prices change from p' to p'' (the long-run input demand for good 2 falls with this increase in the price of good 3). It is this non-uniformity that enables the example to contradict the conclusion of the LeChatelier principle.

8 Recovering the Production Set

We now turn to a set of questions concerning what can be learned from a set of observations of the form $\{(p_1, y_1), ..., (p_n, y_n)\}$. The questions are:

- 1. Is the set of observations consistent with profit maximization at fixed prices from some production set?
- 2. What can we infer about the underlying production set?
- 3. If the data set is sufficiently large, can we recover the entire production set? The production function?

For the first question, we can certainly infer from these observations that $\{y_1, ..., y_n\} \subset Y$. So, the firm's choices could be profit-maximizing only if for all $n, m, p_n \cdot (y_n - y_m) \geq 0$. If this inequality failed, then the choice made at prices p_n were less profitable than choosing the feasible alternative y_m .⁴ Conversely, if all those inequalities are satisfied, then if the production set is $Y = \{y_1, ..., y_n\}$, then the choice is profit-maximizing for every price vector p_n . These inequalities, then, characterize a dataset that is consistent with profit maximization.

For the second question, we can certainly infer that $Y^I = \{y_1, ..., y_n\} \subset Y$. We may call Y^I the *inner bound* on the production set. If we assume that the firm is maximizing profits, then the production set can only contain points for which the profits at prices p_n are no more than $p_n \cdot y_n$, that is, $Y \subset \{y | p_n \cdot y \leq p_n \cdot y_n\} = Y^O$, where Y^O is the *outer bound* on Y. Thus, $Y^I \subset Y \subset Y^O$.

If we assume that the production set satisfies free disposal, we can expand that inner bound. The free disposal inner bound is:

$$Y_{FD}^{I} = \{y | (\exists n)y \le y_n\}.$$

The condition of free disposal implies that $Y_{FD}^I \subset Y$. With that extra assumption, the answer to the second question is $Y_{FD}^I \subset Y \subset Y^O$.

⁴Notice that if this inequality fails, that does not establish that the firm is not a profit maximizer. An alternative explanation is that the firm is not a price taker.

The third question, about large data sets is interpreted as supposing that we know the entire decision function y(p). That implies that we know $\pi(p)$ because for each p, taking any $y \in y(p)$, $\pi(p) = p \cdot y$. By a previous proposition, we know that if Y is closed and convex and has the free disposal property, then $Y = \{y \in \mathbb{R}^n : p \cdot y \leq \pi(p) \text{ for all } p > 0\}$. So, in that case, the production set coincides with its outer bound $Y = Y^O$. Moreover, by the supporting hyperplane theorem, every point on the boundary of Y is chosen for some price vector, so in the limit, the inner bound also coincides with the production set: $Y = Y^I$. So, when Y is closed and convex and satisfies free disposal, the inner and outer bounds derived from y(p) coincide with each other- $Y^I = Y^O$ -and it follows that we can recover Y from the function y(p) in that case.

Knowledge of the correspondence y(p), however, is not sufficient when the production set Y when the set is not convex, even if we make the usual (and relatively innocuous) assumptions that Y is closed and satisfies free disposal. The outer bound, Y^O , being an intersection of closed half-spaces, is always convex, so it can't coincide with Y when Y is not convex. Hence, it cannot correspond to the inner bound, either. There is insufficient information in y(p) to decide whether the points in the difference set $Y^O - Y^I$ are in the set Y.

When the production set is convex and production is described by a production function, the possibility of recovering the production set can take a particularly nice form.

Proposition 13 (Duality of Profit and Production Functions) Suppose that f(x) is a production function, that $\pi(p)$ is the associated profit function $\pi(p) = \max_x f(x) - p \cdot x$, where the price of output is normalized to one. If f is concave, then

$$f(x) = \min_{p} \pi(p) + p \cdot x$$

Proof. By definition, for all $x, p, \pi(p) \ge f(x) - p \cdot x$, so $f(x) \le \pi(p) + p \cdot x$ for all p. It follows that $f(x) \le \min_{p} \pi(p) + p \cdot x$.

Assuming that f is concave, we now show the reverse inequality. Sine f is concave, then the production set $Y = \{(y, -x)|y \leq f(x)\}$ is convex. So, by the supporting hyperplane theorem, there exists some hyperplane that supports Y

at the boundary point (f(x), -x). That is, there is some \widehat{p} such that for all x', $f(x') - \widehat{p} \cdot x' \leq f(x) - \widehat{p} \cdot x$. Then, $\pi(\widehat{p}) = f(x) - \widehat{p} \cdot x$, so $f(x) = \pi(\widehat{p}) + \widehat{p} \cdot x$. Hence, $f(x) \geq \min_p \pi(p) + p \cdot x$. Q.E.D.

We have already seen that we cannot recover the production set or production function for the non-convex case, because there is necessarily a gap between the inner and outer estimates of the set. Hence, the preceding results give a complete answer to the question of when the production set can be recovered from sufficiently rich data about the choices of a competitive profit-maximizing firm.