CHAPTER

5

Production

5.A Introduction

In this chapter, we move to the supply side of the economy, studying the process by which the goods and services consumed by individuals are produced. We view the supply side as composed of a number of productive units, or, as we shall call them, "firms." Firms may be corporations or other legally recognized businesses. But they must also represent the productive possibilities of individuals or households. Moreover, the set of all firms may include some potential productive units that are never actually organized. Thus, the theory will be able to accommodate both active production processes and potential but inactive ones.

Many aspects enter a full description of a firm: Who owns it? Who manages it? How is it managed? How is it organized? What can it do? Of all these questions, we concentrate on the last one. Our justification is not that the other questions are not interesting (indeed, they are), but that we want to arrive as quickly as possible at a minimal conceptual apparatus that allows us to analyze market behavior. Thus, our model of production possibilities is going to be very parsimonious: The firm is viewed merely as a "black box", able to transform inputs into outputs.

In Section 5.B, we begin by introducing the firm's *production set*, a set that represents the production *activities*, or *production plans*, that are technologically feasible for the firm. We then enumerate and discuss some commonly assumed properties of production sets, introducing concepts such as *returns to scale*, *free disposal*, and *free entry*.

After studying the firm's technological possibilities in Section 5.B, we introduce its objective, the goal of *profit maximization*, in Section 5.C. We then formulate and study the firm's profit maximization problem and two associated objects, the firm's *profit function* and its *supply correspondence*. These are, respectively, the value function and the optimizing vectors of the firm's profit maximization problem. Related to the firm's goal of profit maximization is the task of achieving cost-minimizing production. We also study the firm's cost minimization problem and two objects associated with it: The firm's *cost function* and its *conditional factor demand correspondence*. As with the utility maximization and expenditure minimization problems in the theory of demand, there is a rich duality theory associated with the profit maximization and cost minimization problems.

Section 5.D analyzes in detail the geometry associated with cost and production relationships for the special but theoretically important case of a technology that produces a single output.

Aggregation theory is studied in Section 5.E. We show that aggregation on the supply side is simpler and more powerful than the corresponding theory for demand covered in Chapter 4.

Section 5.F constitutes an excursion into welfare economics. We define the concept of *efficient production* and study its relation to profit maximization. With some minor qualifications, we see that profit-maximizing production plans are efficient and that when suitable convexity properties hold, the converse is also true: An efficient plan is profit maximizing for an appropriately chosen vector of prices. This constitutes our first look at the important ideas of the *fundamental theorems of welfare economics.*

In Section 5.G, we point out that profit maximization does not have the same primitive status as preference maximization. Rigorously, it should be derived from the latter. We discuss this point and related issues.

In Appendix A, we study in more detail a particular, important case of production technologies: Those describable by means of linear constraints. It is known as the *linear activity model.*

5.B Production Sets

As in the previous chapters, we consider an economy with L commodities. A *production vector* (also known as an *input–output*, or *netput*, vector, or as a *production plan*) is a vector $y = (y_1, \dots, y_L) \in \mathbb{R}^L$ that describes the (net) outputs of the L commodities from a production process. We adopt the convention that positive numbers denote outputs and negative numbers denote inputs. Some elements of a production vector may be zero; this just means that the process has no net output of that commodity.

Example 5.B.1: Suppose that $L = 5$. Then $y = (-5, 2, -6, 3, 0)$ means that 2 and 3 units of goods 2 and 4, respectively, are produced, while 5 and 6 units of goods 1 and 3, respectively, are used. Good 5 is neither produced nor used as an input in this production vector. ■

To analyze the behavior of the firm, we need to start by identifying those production vectors that are technologically possible. The set of all production vectors that constitute feasible plans for the firm is known as the *production set* and is denoted by $Y \subset \mathbb{R}^L$. Any $y \in Y$ is possible; any $y \notin Y$ is not. The production set is taken as a primitive datum of the theory.

The set of feasible production plans is limited first and foremost by technological constraints. However, in any particular model, legal restrictions or prior contractual commitments may also contribute to the determination of the production set.

It is sometimes convenient to describe the production set Y using a function $F(\cdot)$, called the *transformation function.* The transformation function $F(\cdot)$ has the property that $Y = \{y \in \mathbb{R}^L\colon F(y) \leq 0\}$ and $F(y) = 0$ if and only if y is an element of the boundary of Y. The set of boundary points of Y, $\{y \in \mathbb{R}^L\colon F(y) = 0\}$, is known as the *transformation frontier.* Figure 5.B.1 presents a two-good example.

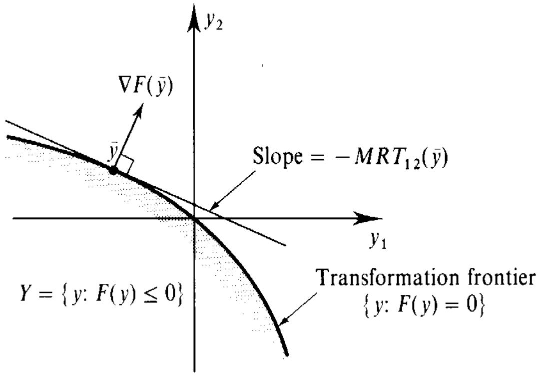


Figure 5.B.1
The production set and transformation frontier.

If $F(\cdot)$ is differentiable, and if the production vector $\bar{y}$ satisfies $F(\bar{y}) = 0$, then for any commodities ℓ and k, the ratio

$$MRT_{\ell k}(\bar{y}) = \frac{\partial F(\bar{y})/\partial y_\ell}{\partial F(\bar{y})/\partial y_k}$$

is called the *marginal rate of transformation (MRT) of good ℓ for good k at $\bar{y}$*.[1] The marginal rate of transformation is a measure of how much the (net) output of good k can increase if the firm decreases the (net) output of good ℓ by one marginal unit. Indeed, from $F(\bar{y}) = 0$, we get

$$\frac{\partial F(\bar{y})}{\partial y_k} dy_k + \frac{\partial F(\bar{y})}{\partial y_\ell} dy_\ell = 0,$$

and therefore the slope of the transformation frontier at $\bar{y}$ in Figure 5.B.1 is precisely $-MRT_{12}(\bar{y})$.

Technologies with Distinct Inputs and Outputs

In many actual production processes, the set of goods that can be outputs is distinct from the set that can be inputs. In this case, it is sometimes convenient to notationally distinguish the firm's inputs and outputs. We could, for example, let $q = (q_1, \ldots, q_M) \geq 0$ denote the production levels of the firm's M outputs and $z = (z_1, \ldots, z_{L-M}) \geq 0$ denote the amounts of the firm's $L - M$ inputs, with the convention that the amount of input z_ℓ used is now measured as a *nonnegative* number (as a matter of notation, we count all goods not actually used in the process as inputs).

One of the most frequently encountered production models is that in which there is a single output. A single-output technology is commonly described by means of a *production function* $f(z)$ that gives the maximum amount q of output that can be produced using input amounts $(z_1, \ldots, z_{L-1}) \geq 0$. For example, if the output is good L, then (assuming that output can be disposed of at no cost) the production function $f(\cdot)$ gives rise to the production set:

$$Y = \{(-z_1, \ldots, -z_{L-1}, q): q - f(z_1, \ldots, z_{L-1}) \leq 0 \quad \text{and} \quad (z_1, \ldots, z_{L-1}) \geq 0\}.$$

Holding the level of output fixed, we can define the *marginal rate of technical*

1. As in Chapter 3, in computing ratios such as this, we always assume that $\partial F(\bar{y})/\partial y_k \neq 0$.

substitution (MRTS) of input ℓ *for input k at* $\bar{z}$ as

$$MRTS_{\ell k}(\bar{z}) = \frac{\partial f(\bar{z})/\partial z_\ell}{\partial f(\bar{z})/\partial z_k}.$$

The number $MRTS_{\ell k}(\bar{z})$ measures the additional amount of input k that must be used to keep output at level $\bar{q} = f(\bar{z})$ when the amount of input ℓ is decreased marginally. It is the production theory analog to the consumer's marginal rate of substitution. In consumer theory, we look at the trade-off between commodities that keeps utility constant, here, we examine the trade-off between inputs that keeps the amount of output constant. Note that $MRTS_{\ell k}$ is simply a renaming of the marginal rate of transformation of input ℓ for input k in the special case of a single-output, many-input technology.

Example 5.B.2: *The Cobb–Douglas Production Function* The *Cobb–Douglas* production function with two inputs is given by $f(z_1, z_2) = z_1^\alpha z_2^\beta$, where $\alpha \geq 0$ and $\beta \geq 0$. The marginal rate of technical substitution between the two inputs at $z = (z_1, z_2)$ is $MRTS_{12}(z) = \alpha z_2/\beta z_1$. ■

Properties of Production Sets

We now introduce and discuss a fairly exhaustive list of commonly assumed properties of production sets. The appropriateness of each of these assumptions depends on the particular circumstances (indeed, some of them are mutually exclusive).[2]

(i) *Y is nonempty.* This assumption simply says that the firm has something it can plan to do. Otherwise, there is no need to study the behavior of the firm in question.

(ii) *Y is closed.* The set Y includes its boundary. Thus, the limit of a sequence of technologically feasible input–output vectors is also feasible; in symbols, $y^n \to y$ and $y^n \in Y$ imply $y \in Y$. This condition should be thought of as primarily technical.[3]

(iii) *No free lunch.* Suppose that $y \in Y$ and $y \geq 0$, so that the vector y does not use any inputs. The no-free-lunch property is satisfied if this production vector cannot produce output either. That is, whenever $y \in Y$ and $y \geq 0$, then $y = 0$; it is not possible to produce something from nothing. Geometrically, $Y \cap \mathbb{R}^L_+ \subset \{0\}$. For $L = 2$, Figure 5.B.2(a) depicts a set that violates the no-free-lunch property, the set in Figure 5.B.2(b) satisfies it.

(iv) *Possibility of inaction* This property says that $0 \in Y$: Complete shutdown is possible. Both sets in Figure 5.B.2, for example, satisfy this property. The point in time at which production possibilities are being analyzed is often important for the validity of this assumption. If we are contemplating a firm that could access a set of technological possibilities but that has not yet been organized, then inaction is clearly

2. For further discussion of these properties, see Koopmans (1957) and Chapter 3 of Debreu (1959).

3. Nonetheless, we show in Exercise 5.B.4 that there is an important case of economic interest when it raises difficulties.

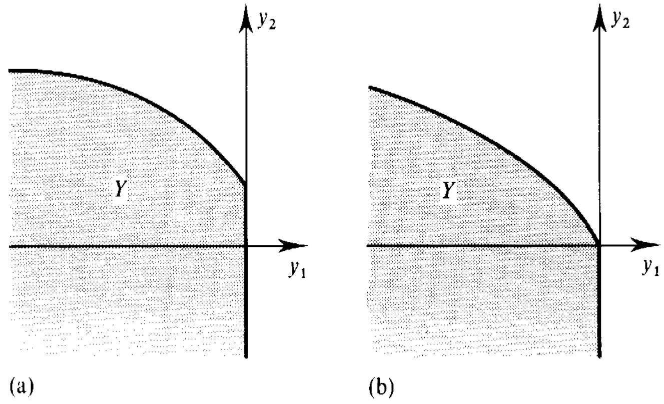


Figure 5.B.2
The no free lunch property.
(a) Violates no free lunch.
(b) Satisfies no free lunch.

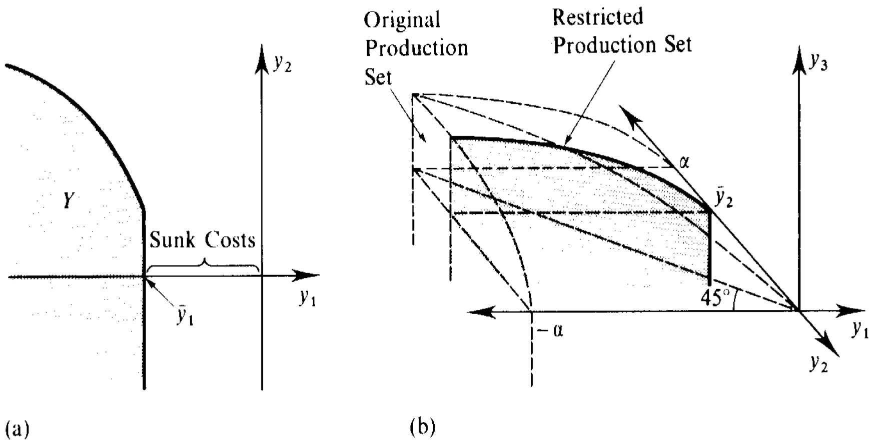
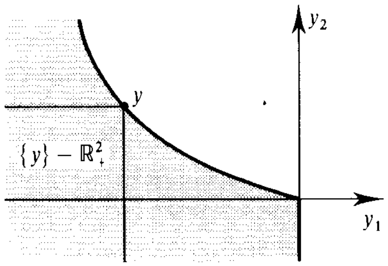


Figure 5.B.3
Two production sets with sunk costs.
(a) A minimal level of expenditure committed.
(b) One kind of input fixed.

possible. But if some production decisions have already been made, or if irrevocable contracts for the delivery of some inputs have been signed, inaction is not possible. In that case, we say that some costs are *sunk*. Figure 5.B.3 depicts two examples. The production set in Figure 5.B.3(a) represents the *interim* production possibilities arising when the firm is already committed to use at least $-\bar{y}_1$ units of good 1 (perhaps because it has already signed a contract for the purchase of this amount); that is, the set is a *restricted production set* that reflects the firm's remaining choices from some original production set Y like the ones in Figure 5.B.2. In Figure 5.B.3(b), we have a second example of sunk costs. For a case with one output (good 3) and two inputs (goods 1 and 2), the figure illustrates the restricted production set arising when the level of the second input has been irrevocably set at $\bar{y}_2 < 0$ [here, in contrast with Figure 5.B.3(a), increases in the use of the input are impossible].

(v) *Free disposal.* The property of free disposal holds if the absorption of any additional amounts of inputs without any reduction in output is always possible. That is, if $y \in Y$ and $y' \leq y$ (so that y' produces at most the same amount of outputs using at least the same amount of inputs), then $y' \in Y$. More succinctly, $Y - \mathbb{R}^L_+ \subset Y$ (see Figure 5.B.4). The interpretation is that the extra amount of inputs (or outputs) can be disposed of or eliminated at no cost.



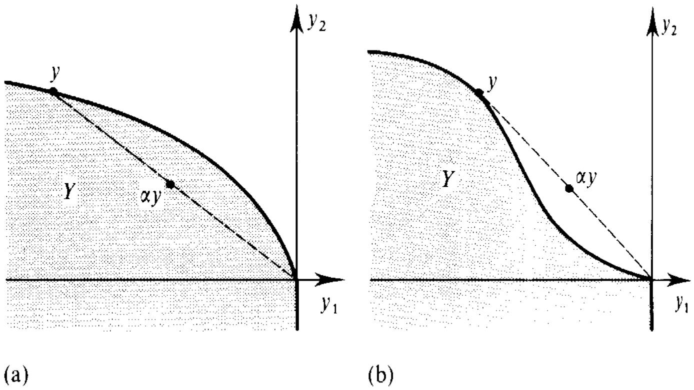


Figure 5.B.4 (left) The free disposal property.

Figure 5.B.5 (right) The nonincreasing returns to scale property. (a) Nonincreasing returns satisfied. (b) Nonincreasing returns violated.

(vi) *Irreversibility*. Suppose that $y \in Y$ and $y \neq 0$. Then irreversiblity says that $-y \notin Y$. In words, it is impossible to reverse a technologically possible production vector to transform an amount of output into the same amount of input that was used to generate it. If, for example, the description of a commodity includes the time of its availability, then irreversibility follows from the requirement that inputs be used before outputs emerge.

Exercise 5.B.1: Draw two production sets: one that violates irreversibility and one that satisfies this property.

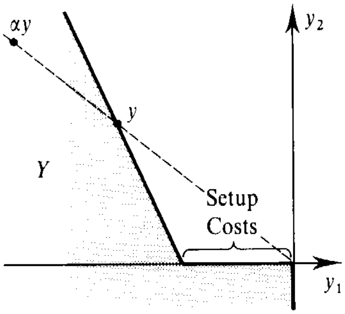
(vii) *Nonincreasing returns to scale*. The production technology Y exhibits nonincreasing returns to scale if for any $y \in Y$, we have $\alpha y \in Y$ for all scalars $\alpha \in [0, 1]$. In words, any feasible input–output vector can be scaled down (see Figure 5.B.5). Note that nonincreasing returns to scale imply that inaction is possible [property (iv)].

(viii) *Nondecreasing returns to scale*. In contrast with the previous case, the production process exhibits nondecreasing returns to scale if for any $y \in Y$, we have $\alpha y \in Y$ for any scale $\alpha \geq 1$. In words, any feasible input–output vector can be scaled up. Figure 5.B.6(a) presents a typical example; in the figure, units of output (good 2) can be produced at a constant cost of input (good 1) except that in order to produce at all, a fixed setup cost is required. It does not matter for the existence of nondecreasing returns if this fixed cost is sunk [as in Figure 5.B.6(b)] or not [as in Figure 5.B.6(a), where inaction is possible].

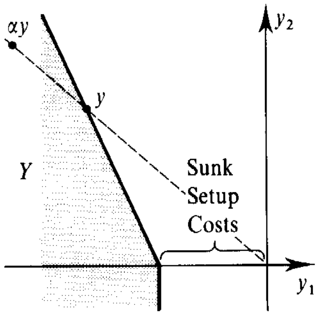
(ix) *Constant returns to scale*. This property is the conjunction of properties (vii) and (viii). The production set Y exhibits constant returns to scale if $y \in Y$ implies $\alpha y \in Y$ for any scalar $\alpha \geq 0$. Geometrically, Y is a *cone* (see Figure 5.B.7).

For single-output technologies, properties of the production set translate readily into properties of the production function $f(\cdot)$. Consider Exercise 5.B.2 and Example 5.B.3.

Exercise 5.B.2: Suppose that $f(\cdot)$ is the production function associated with a single-output technology, and let Y be the production set of this technology. Show that Y satisfies constant returns to scale if and only if $f(\cdot)$ is homogeneous of degree one.



(a)



(b)

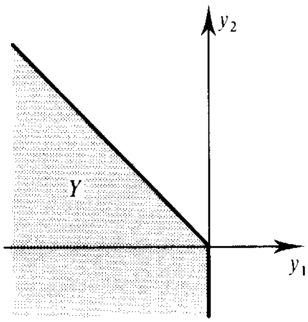


Figure 5.B.6 (left)
The nondecreasing returns to scale property.

Figure 5.B.7 (right)
A technology satisfying the constant returns to scale property.

Example 5.B.3: *Returns to Scale with the Cobb–Douglas Production Function*: For the Cobb–Douglas production function introduced in Example 5.B.2, $f(2z_1, 2z_2) = 2^{\alpha+\beta}z_1^{\alpha}z_2^{\beta} = 2^{\alpha+\beta}f(z_1, z_2)$. Thus, when $\alpha + \beta = 1$, we have constant returns to scale; when $\alpha + \beta < 1$, we have decreasing returns to scale; and when $\alpha + \beta > 1$, we have increasing returns to scale. ■

(x) *Additivity* (*or free entry*). Suppose that $y \in Y$ and $y' \in Y$. The additivity property requires that $y + y' \in Y$. More succinctly, $Y + Y \subset Y$. This implies, for example, that $ky \in Y$ for any positive *integer* k. In Figure 5.B.8, we see an example where Y is additive. Note that in this example, output is available only in integer amounts (perhaps because of indivisibilities). The economic interpretation of the additivity condition is that if y and y' are both possible, then one can set up two plants that do not interfere with each other and carry out production plans y and y' independently. The result is then the production vector $y + y'$.

Additivity is also related to the idea of entry. If $y \in Y$ is being produced by a firm and another firm enters and produces $y' \in Y$, then the net result is the vector $y + y'$. Hence, the *aggregate production set* (the production set describing feasible production plans for the economy as a whole) must satisfy additivity whenever unrestricted entry, or (as it is called in the literature) *free entry*, is possible.

(xi) *Convexity*. This is one of the fundamental assumptions of microeconomics. It postulates that the production set Y is convex. That is, if $y, y' \in Y$ and $\alpha \in [0, 1]$, then $\alpha y + (1 - \alpha)y' \in Y$. For example, Y is convex in Figure 5.B.5(a) but is not convex in Figure 5.B.5(b).

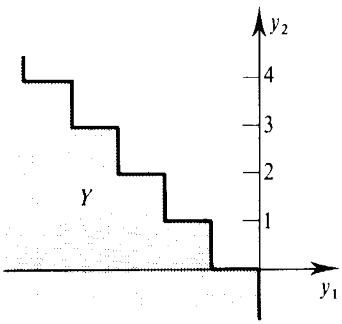


Figure 5.B.8
A production set satisfying the additivity property.

The convexity assumption can be interpreted as incorporating two ideas about production possibilities. The first is nonincreasing returns. In particular, if inaction is possible (i.e., if $0 \in Y$), then convexity implies that Y has nonincreasing returns to scale. To see this, note that for any $\alpha \in [0, 1]$, we can write $\alpha y = \alpha y + (1 - \alpha)0$. Hence, if $y \in Y$ and $0 \in Y$, convexity implies that $\alpha y \in Y$. Second, convexity captures the idea that "unbalanced" input combinations are not more productive than balanced ones (or, symmetrically, that "unbalanced" output combinations are not least costly to produce than balanced ones). In particular, if production plans y and y' produce exactly the same amount of output but use different input combinations, then a production vector that uses a level of each input that is the average of the levels used in these two plans can do at least as well as either y or y'.

Exercise 5.B.3 illustrates these two ideas for the case of a single-output technology.

Exercise 5.B.3: Show that for a single-output technology, Y is convex if and only if the production function $f(z)$ is concave.

(xii) *Y is a convex cone.* This is the conjunction of the convexity (xi) and constant returns to scale (ix) properties. Formally, Y is a convex cone if for any production vector $y, y' \in Y$ and constants $\alpha \geq 0$ and $\beta \geq 0$, we have $\alpha y + \beta y' \in Y$. The production set depicted in Figure 5.B.7 is a convex cone.

An important fact is given in Proposition 5.B.1.

Proposition 5.B.1: The production set Y is additive and satisfies the nonincreasing returns condition if and only if it is a convex cone.

Proof: The definition of a convex cone directly implies the nonincreasing returns and additivity properties. Conversely, we want to show that if nonincreasing returns and additivity hold, then for any $y, y' \in Y$ and any $\alpha > 0$, and $\beta > 0$, we have $\alpha y + \beta y' \in Y$. To this effect, let k be any integer such that $k > \text{Max}\,\{\alpha, \beta\}$. By additivity, $ky \in Y$ and $ky' \in Y$. Since $(\alpha/k) < 1$ and $\alpha y = (\alpha/k)ky$, the nonincreasing returns condition implies that $\alpha y \in Y$. Similarly, $\beta y' \in Y$. Finally, again by additivity, $\alpha y + \beta y' \in Y$. ∎

Proposition 5.B.1 provides a justification for the convexity assumption in production. Informally, we could say that if feasible input–output combinations can always be scaled down, and if the simultaneous operation of several technologies without mutual interference is always possible, then, in particular, convexity obtains. (See Appendix A of Chapter 11 for several examples in which there is mutual interference and, as a consequence, convexity does not arise.)

It is important not to lose sight of the fact that the production set describes technology, not limits on resources. It can be argued that if all inputs (including, say, entrepreneurial inputs) are explicitly accounted for, then it should always be possible to replicate production. After all, we are not saying that doubling output is actually feasible, only that in principle it would be possible if *all* inputs (however esoteric, be they marketed or not) were doubled. In this view, which originated with Marshall and has been much emphasized by McKenzie (1959), decreasing returns must reflect the scarcity of an underlying, unlisted input of production. For this reason, some economists believe that among models with convex technologies the constant returns model is the most fundamental. Proposition 5.B.2 makes this idea precise.

Proposition 5.B.2: For any convex production set $Y \subset \mathbb{R}^L$ with $0 \in Y$, there is a constant returns, convex production set $Y' \subset \mathbb{R}^{L+1}$ such that $Y = \{y \in \mathbb{R}^L : (y, -1) \in Y'\}$.

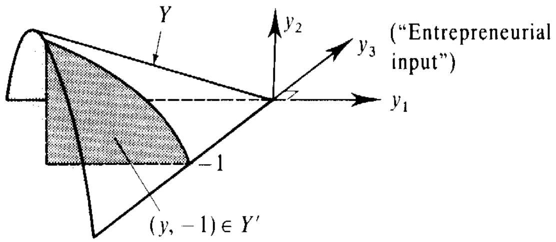


Figure 5.B.9
A constant returns production set with an "entrepreneurial factor."

Proof: Simply let $Y' = \{y' \in \mathbb{R}^{L+1}: y' = \alpha(y, -1) \text{ for some } y \in Y \text{ and } \alpha \geq 0\}$. (See Figure 5.B.9.) ■

The additional input included in the extended production set (good $L + 1$) can be called the "entrepreneurial factor." (The justification for this can be seen in Exercise 5.C.12; in a competitive environment, the return to this entrepreneurial factor is precisely the firm's profit.) In essence, the implication of Proposition 5.B.2 is that in a competitive, convex setting, there may be little loss of conceptual generality in limiting ourselves to constant returns technologies.

5.C Profit Maximization and Cost Minimization

In this section, we begin our study of the market behavior of the firm. In parallel to our study of consumer demand, we assume that there is a vector of prices quoted for the L goods, denoted by $p = (p_1, \ldots, p_L) \gg 0$, and that these prices are independent of the production plans of the firm (the *price-taking assumption*).

We assume throughout this chapter that the firm's objective is to maximize its profit. (It is quite legitimate to ask why this should be so, and we will offer a brief discussion of the issue in Section 5.G.) Moreover, we always assume that the firm's production set Y satisfies the properties of *nonemptiness*, *closedness*, and *free disposal* (see Section 5.B).

The Profit Maximization Problem

Given a price vector $p \gg 0$ and a production vector $y \in \mathbb{R}^L$, the profit generated by implementing y is $p \cdot y = \sum_{\ell=1}^{L} p_\ell y_\ell$. By the sign convention, this is precisely the total revenue minus the total cost. Given the technological constraints represented by its production set Y, the firm's *profit maximization problem* (*PMP*) is then

$$\begin{aligned} \underset{y}{\text{Max}} \quad & p \cdot y \\ \text{s.t. } & y \in Y. \end{aligned} \qquad \text{(PMP)}$$

Using a transformation function to describe Y, $F(\cdot)$, we can equivalently state the PMP as

$$\begin{aligned} \underset{y}{\text{Max}} \quad & p \cdot y \\ \text{s.t. } & F(y) \leq 0. \end{aligned}$$

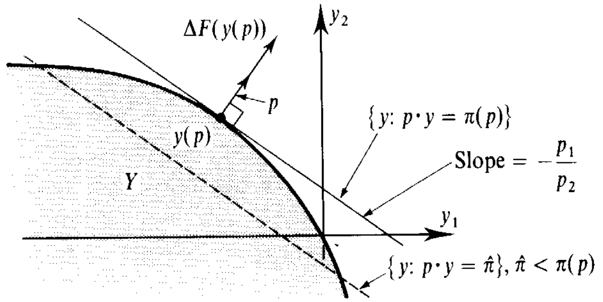


Figure 5.C.1
The profit maximization problem.

Given a production set Y, the firm's *profit function* $\pi(p)$ associates to every p the amount $\pi(p) = \text{Max}\,\{p\cdot y\colon y \in Y\}$, the value of the solution to the PMP. Correspondingly, we define the firm's *supply correspondence* at p, denoted $y(p)$, as the set of profit-maximizing vectors $y(p) = \{y \in Y\colon p\cdot y = \pi(p)\}$.[4] Figure 5.C.1 depicts the supply to the PMP for a strictly convex production set Y. The optimizing vector $y(p)$ lies at the point in Y associated with the highest level of profit. In the figure, $y(p)$ therefore lies on the *iso-profit line* (a line in $\mathbb{R}^2$ along which all points generate equal profits) that intersects the production set farthest to the northeast and is, therefore, tangent to the boundary of Y at $y(p)$.

In general, $y(p)$ may be a set rather than a single vector. Also, it is possible that no profit-maximizing production plan exists. For example, the price system may be such that there is no bound on how high profits may be. In this case, we say that $\pi(p) = +\infty$.[5] To take a concrete example, suppose that $L = 2$ and that a firm with a constant returns technology produces one unit of good 2 for every unit of good 1 used as an input. Then $\pi(p) = 0$ whenever $p_2 \le p_1$. But if $p_2 > p_1$, then the firm's profit is $(p_2 - p_1)y_2$, where y_2 is the production of good 2. Clearly, by choosing y_2 appropriately, we can make profits arbitrarily large. Hence, $\pi(p) = +\infty$ if $p_2 > p_1$.

Exercise 5.C.1: Prove that, in general, if the production set Y exhibits nondecreasing returns to scale, then either $\pi(p) \le 0$ or $\pi(p) = +\infty$.

If the transformation function $F(\cdot)$ is differentiable, then first-order conditions can be used to characterize the solution to the PMP. If $y^* \in y(p)$, then, for some $\lambda \ge 0$, y^* must satisfy the first-order conditions

$$p_\ell = \lambda \frac{\partial F(y^*)}{\partial y_\ell} \quad \text{for } \ell = 1, \ldots, L$$

or, equivalently, in matrix notation,

$$p = \lambda \nabla F(y^*). \tag{5.C.1}$$

4. We use the term *supply correspondence* to keep the parallel with the *demand* terminology of the consumption side. Recall however that $y(p)$ is more properly thought of as the firm's *net* supply to the market. In particular, the negative entries of a supply vector should be interpreted as demand for inputs.

5. Rigorously, to allow for the possibility that $\pi(p) = +\infty$ (as well as for other cases where no profit-maximizing production plan exists), the profit function should be defined by $\pi(p) = \text{Sup}\,\{p\cdot y\colon y \in Y\}$. We will be somewhat loose, however, and continue to use *Max* while allowing for this possiblity.

In words, the *price vector p and the gradient* $\nabla F(y^*)$ *are proportional* (Figure 5.C.1 depicts this fact). Condition (5.C.1) also yields the following ratio equality: $p_\ell/p_k = MRT_{\ell k}(y^*)$ for all ℓ, k. For $L = 2$, this says that the slope of the transformation frontier at the profit-maximizing production plan must be equal to the negative of the price ratio, as shown in Figure 5.C.1. Were this not so, a small change in the firm's production plan could be found that increases the firm's profits.

When Y corresponds to a single-output technology with differentiable production function $f(z)$, we can view the firm's decision as simply a choice over its input levels z. In this special case, we shall let the scalar $p > 0$ denote the price of the firm's output and the vector $w \gg 0$ denote its input prices.[6] The input vector z^* maximizes profit given (p, w) if it solves

$$\underset{z \geq 0}{\text{Max}}\ pf(z) - w \cdot z.$$

If z^* is optimal, then the following first-order conditions must be satisfied for $\ell = 1, \ldots, L-1$:

$$p\frac{\partial f(z^*)}{\partial z_\ell} \leq w_\ell, \quad \text{with equality if } z_\ell^* > 0,$$

or, in matrix notation,

$$p\,\nabla f(z^*) \leq w \qquad \text{and} \qquad [p\,\nabla f(z^*) - w] \cdot z^* = 0.^7 \qquad (5.C.2)$$

Thus, the marginal product of every input ℓ actually used (i.e., with $z_\ell^* > 0$) must equal its price in terms of output, w_ℓ/p. Note also that for any two inputs ℓ and k with $(z_\ell^*, z_k^*) \gg 0$, condition (5.C.2) implies that $MRTS_{\ell k} = w_\ell/w_k$; that is, the marginal rate of technical substitution between the two inputs is equal to their price ratio, the economic rate of substitution between them. This ratio condition is merely a special case of the more general condition derived in (5.C.1).

If the production set Y is convex, then the first-order conditions in (5.C.1) and (5.C.2) are not only necessary but also sufficient for the determination of a solution to the PMP.

Proposition 5.C.1, which lists the properties of the profit function and supply correspondence, can be established using methods similar to those we employed in Chapter 3 when studying consumer demand. Observe, for example, that mathematically the concept of the profit function should be familiar from the discussion of duality in Chapter 3. In fact, $\pi(p) = -\mu_{-Y}(p)$, where $\mu_{-Y}(p) = \text{Min}\,\{p \cdot (-y) \colon y \in Y\}$ is the support function of the set $-Y$. Thus, the list of important properties in Proposition 5.C.1 can be seen to follow from the general properties of support functions discussed in Section 3.F.

6. Up to now, we have always used the symbol p for an overall vector of prices; here we use it only for the output price and we denote the vector of input prices by w. This notation is fairly standard. As a rule of thumb, unless we are in a context of explicit classification of commodities as inputs or outputs (as in the single-output case), we will continue to use p to denote an overall vector of prices $p = (p_1, \ldots, p_L)$.

7. The concern over boundary conditions arises here, but not in condition (5.C.1), because the assumption of distinct inputs and outputs requires that $z \geq 0$, whereas the formulation leading to (5.C.1) allows the net output of every good to be either positive or negative. Nonetheless, when using the first-order conditions (5.C.2), we will typically assume that $z^* \gg 0$.

Proposition 5.C.1: Suppose that $\pi(\cdot)$ is the profit function of the production set Y and that $y(\cdot)$ is the associated supply correspondence. Assume also that Y is closed and satisfies the free disposal property. Then

(i) $\pi(\cdot)$ is homogeneous of degree one.

(ii) $\pi(\cdot)$ is convex.

(iii) If Y is convex, then $Y = \{y \in \mathbb{R}^L: p \cdot y \leq \pi(p) \text{ for all } p \gg 0\}$.

(iv) $y(\cdot)$ is homogeneous of degree zero.

(v) If Y is convex, then $y(p)$ is a convex set for all p. Moreover, if Y is strictly convex, then $y(p)$ is single-valued (if nonempty).

(vi) (*Hotelling's lemma*) If $y(\bar{p})$ consists of a single point, then $\pi(\cdot)$ is differentiable at $\bar{p}$ and $\nabla\pi(\bar{p}) = y(\bar{p})$.

(vii) If $y(\cdot)$ is a function differentiable at $\bar{p}$, then $Dy(\bar{p}) = D^2\pi(\bar{p})$ is a symmetric and positive semidefinite matrix with $Dy(\bar{p})\bar{p} = 0$.

Properties (ii), (iii), (vi), and (vii) are the nontrivial ones.

Exercise 5.C.2: Prove that $\pi(\cdot)$ is a convex function [Property (ii) of Proposition 5.C.1]. [*Hint*: Suppose that $y \in y(\alpha p + (1-\alpha)p')$. Then

$$\pi(\alpha p + (1-\alpha)p') = \alpha p \cdot y + (1-\alpha)p' \cdot y \leq \alpha\pi(p) + (1-\alpha)\pi(p').]$$

Property (iii) tells us that if Y is closed, convex, and satisfies free disposal, then $\pi(p)$ provides an alternative ("dual") description of the technology. As for the indirect utility function's (or expenditure function's) representation of preferences (discussed in Chapter 3), it is a less primitive description than Y itself because it depends on the notions of prices and of price-taking behavior. But thanks to property (vi), it has the great virtue in applications of often allowing for an immediate computation of supply.

Property (vi) relates supply behavior to the derivatives of the profit function. It is a direct consequence of the duality theorem (Proposition 3.F.1). As in Proposition 3.G.1, the fact that $\nabla\pi(\bar{p}) = y(\bar{p})$ can also be established by the related arguments of the envelope theorem and of first-order conditions.

The positive semidefiniteness of the matrix $Dy(p)$ in property (vii), which in view of property (vi) is a consequence of the convexity of $\pi(\cdot)$, is the general mathematical expression of the *law of supply: Quantities respond in the same direction as price changes*. By the sign convention, this means that *if the price of an output increases* (all other prices remaining the same), then *the supply of the output increases*; and *if the price of an input increases*, then *the demand for the input decreases*.

Note that the law of supply holds for *any* price change. Because, in contrast with demand theory, there is no budget constraint, there is no compensation requirement of any sort. In essence, we have no wealth effects here, only substitution effects.

In nondifferentiable terms, the law of supply can be expressed as

$$(p - p') \cdot (y - y') \geq 0 \qquad (5.C.3)$$

for all p, p', $y \in y(p)$, and $y' \in y(p')$. In this form, it can also be established by a straightforward revealed preference argument. In particular,

$$(p - p') \cdot (y - y') = (p \cdot y - p \cdot y') + (p' \cdot y' - p' \cdot y) \geq 0,$$

where the inequality follows from the fact that $y \in y(p)$ and $y' \in y(p')$ (i.e., from the fact that y is profit maximizing given prices p and y' is profit maximizing for prices p').

Property (vii) of Proposition 5.C.1 implies that the matrix $Dy(p)$, the *supply substitution matrix*, has properties that parallel (although with the reverse sign) those for the substitution matrix of demand theory. Thus, own-substitution effects are nonnegative as noted above $[\partial y_\ell(p)/\partial p_\ell \geq 0$ for all $\ell]$, and substitution effects are symmetric $[\partial y_\ell(p)/\partial p_k = \partial y_k(p)/\partial p_\ell$ for all $\ell, k]$. The fact that $Dy(p)p = 0$ follows from the homogeneity of $y(\cdot)$ [property (iv)] in a manner similar to the parallel property of the demand substitution matrix discussed in Chapter 3.

Cost Minimization

An important implication of the firm choosing a profit-maximizing production plan is that there is no way to produce the same amounts of outputs at a lower total input cost. Thus, cost minimization is a necessary condition for profit maximization. This observation motivates us to an independent study of the firm's *cost minimization problem*. The problem is of interest for several reasons. First, it leads us to a number of results and constructions that are technically very useful. Second, as we shall see in Chapter 12, when a firm is not a price taker in its output market, we can no longer use the profit function for analysis. Nevertheless, as long as the firm is a price taker in its input market, the results flowing from the cost minimization problem continue to be valid. Third, when the production set exhibits nondecreasing returns to scale, the value function and optimizing vectors of the cost minimization problem, which keep the levels of outputs fixed, are better behaved than the profit function and supply correspondence of the PMP (e.g., recall from Exercise 5.C.1 that the profit function can take only the values 0 and $+\infty$).

To be concrete, we focus our analysis on the single-output case. As usual, we let z be a nonnegative vector of inputs, $f(z)$ the production function, q the amounts of output, and $w \gg 0$ the vector of input prices. The *cost minimization problem* (CMP) can then be stated as follows (we assume free disposal of output):

$$\begin{aligned} \underset{z \geq 0}{\text{Min}} \quad & w \cdot z \\ \text{s.t.} \quad & f(z) \geq q. \end{aligned} \qquad \text{(CMP)}$$

The optimized value of the CMP is given by the *cost function* $c(w, q)$. The corresponding optimizing set of input (or factor) choices, denoted by $z(w, q)$, is known as the *conditional factor demand correspondence* (or *function* if it is always single-valued). The term *conditional* arises because these factor demands are conditional on the requirement that the output level q be produced.

The solution to the CMP is depicted in Figure 5.C.2(a) for a case with two inputs. The shaded region represents the set of input vectors z that can produce at least the amount q of output. It is the projection (into the positive orthant of the input space) of the part of the production set Y than generates output of at least q, as shown in Figure 5.C.2(b). In Figure 5.C.2(a), the solution $z(w, q)$ lies on the iso-cost line (a line in $\mathbb{R}^2$ on which all input combinations generate equal cost) that intersects the set $\{z \in \mathbb{R}^L_+ : f(z) \geq q\}$ closest to the origin.

If z^* is optimal in the CMP, and if the production function $f(\cdot)$ is differentiable,

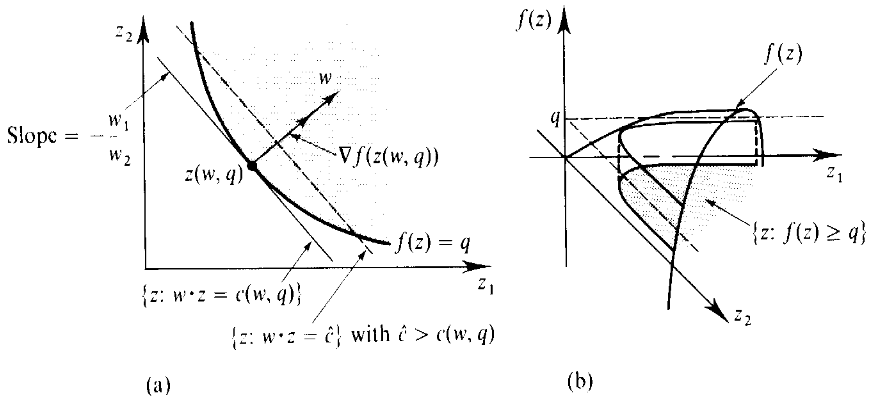


Figure 5.C.2
The cost minimization problem.
(a) Two inputs.
(b) The isoquant as a section of the production set.

then for some $\lambda \geq 0$, the following first-order conditions must hold for every input $\ell = 1, \ldots, L-1$:

$$w_\ell \geq \lambda \frac{\partial f(z^*)}{\partial z_\ell}, \quad \text{with equality if } z_\ell^* > 0,$$

or, in matrix notation,

$$w \geq \lambda \nabla f(z^*) \qquad \text{and} \qquad [w - \lambda \nabla f(z^*)] \cdot z^* = 0. \tag{5.C.4}$$

As with the PMP, if the production set Y is convex [i.e., if $f(\cdot)$ is concave], then condition (5.C.4) is not only necessary but also sufficient for z^* to be an optimum in the CMP.[8]

Condition (5.C.4), like condition (5.C.2) of the PMP, implies that for any two inputs ℓ and k with $(z_\ell, z_k) \gg 0$, we have $MRTS_{\ell k} = w_\ell/w_k$. This correspondence is to be expected because, as we have noted, profit maximization implies that input choices are cost minimizing for the chosen output level q. For $L = 2$, condition (5.C.4) entails that the slope at z^* of the *isoquant* associated with production level q is exactly equal to the negative of the ratio of the input prices $-w_1/w_2$. Figure 5.C.2(a) depicts this fact as well.

As usual, the Lagrange multiplier λ can be interpreted as the marginal value of relaxing the constraint $f(z^*) \geq q$. Thus, λ equals $\partial c(w, q)/\partial q$, the *marginal cost of production*.

Note the close formal analogy with consumption theory here. Replace $f(\cdot)$ by $u(\cdot)$, q by u, and z by x (i.e., interpret the production function as a utility function), and the CMP becomes the expenditure minimization problem (EMP) discussed in Section 3.E. Therefore, in Proposition 5.C.2, properties (i) to (vii) of the cost function and conditional factor demand correspondence follow from the analysis in Sections 3.E to 3.G by this reinterpretation. [You are asked to prove properties (viii) and (ix) in Exercise 5.C.3.]

Proposition 5.C.2: Suppose that $c(w, q)$ is the cost function of a single-output technology Y with production function $f(\cdot)$ and that $z(w, q)$ is the associated

8. Note, however, that the first-order conditions are sufficient for a solution to the CMP as long as the set $\{z\colon f(z) \geq q\}$ is convex. Thus, the key condition for the sufficiency of the first-order conditions of the CMP is the *quasiconcavity* of $f(\cdot)$. This is an important fact because the quasiconcavity of $f(\cdot)$ is compatible with increasing returns to scale (see Example 5.C.1).

conditional factor demand correspondence. Assume also that Y is closed and satisfies the free disposal property. Then

(i) $c(\cdot)$ is homogeneous of degree one in w and nondecreasing in q.

(ii) $c(\cdot)$ is a concave function of w.

(iii) If the sets $\{z \geq 0 : f(z) \geq q\}$ are convex for every q, then $Y = \{(-z, q) : w \cdot z \geq c(w, q) \text{ for all } w \gg 0\}$.

(iv) $z(\cdot)$ is homogeneous of degree zero in w.

(v) If the set $\{z \geq 0 : f(z) \geq q\}$ is convex, then $z(w, q)$ is a convex set. Moreover, if $\{z \geq 0 : f(z) \geq q\}$ is a strictly convex set, then $z(w, q)$ is single-valued.

(vi) (*Shepard's lemma*) If $z(\bar{w}, q)$ consists of a single point, then $c(\cdot)$ is differentiable with respect to w at $\bar{w}$ and $\nabla_w c(\bar{w}, q) = z(\bar{w}, q)$.

(vii) If $z(\cdot)$ is differentiable at $\bar{w}$, then $D_w z(\bar{w}, q) = D_w^2 c(\bar{w}, q)$ is a symmetric and negative semidefinite matrix with $D_w z(\bar{w}, q)\bar{w} = 0$.

(viii) If $f(\cdot)$ is homogeneous of degree one (i.e., exhibits constant returns to scale), then $c(\cdot)$ and $z(\cdot)$ are homogeneous of degree one in q.

(ix) If $f(\cdot)$ is concave, then $c(\cdot)$ is a convex function of q (in particular, marginal costs are nondecreasing in q).

In Exercise 5.C.4 we are asked to show that properties (i) to (vii) of Proposition 5.C.2 also hold for technologies with multiple outputs.

The cost function can be particularly useful when the production set is of the constant returns type. In this case, $y(\cdot)$ is not single-valued at any price vector allowing for nonzero production, making Hotelling's lemma [Proposition 5.C.1(vi)] inapplicable at these prices. Yet, the conditional input demand $z(w, q)$ may nevertheless be single-valued, allowing us to use Shepard's lemma. Keep in mind, however, that the cost function does not contain more information than the profit function. In fact, we know from property (iii) of Propositions 5.C.1 and 5.C.2 that under convexity restrictions there is a one-to-one correspondence between profit and cost functions; that is, from either function, the production set can be recovered, and the other function can then be derived.

Using the cost function, we can restate the firm's problem of determining its profit-maximizing production level as

$$\underset{q \geq 0}{\text{Max}}\ pq - c(w, q). \tag{5.C.5}$$

The necessary first-order condition for q^* to be profit maximizing is then

$$p - \frac{\partial c(w, q^*)}{\partial q} \leq 0, \quad \text{with equality if } q^* > 0. \tag{5.C.6}$$

In words, at an interior optimum (i.e., if $q^* > 0$), *price equals marginal cost.*[9] If $c(w, q)$ is convex in q, then the first-order condition (5.C.6) is also sufficient for q^* to be the firm's optimal output level. (We study the relationship between the firm's supply behavior and the properties of its technology and cost function in detail in Section 5.D.)

9. This can also be seen by noting that the first-order condition (5.C.4) of the CMP coincides with first-order condition (5.C.2) of the PMP if and only if $\lambda = p$. Recall that λ, the multiplier on the constraint in the CMP, is equal to $\partial c(w, q)/\partial q$.

We could go on for many pages analyzing profit and cost functions. Some examples and further properties are contained in the exercises. See McFadden (1978) for an extensive treatment of this topic.

Example 5.C.1: *Profit and Cost Functions for the Cobb–Douglas Production Function.* Here we derive the profit and cost functions for the Cobb–Douglas production function of Example 5.B.2, $f(z_1, z_2) = z_1^{\alpha} z_2^{\beta}$. Recall from Example 5.B.3 that $\alpha + \beta = 1$ corresponds to the case of constant returns to scale, $\alpha + \beta < 1$ corresponds to decreasing returns, and $\alpha + \beta > 1$ corresponds to increasing returns.

The conditional factor demand equations and cost function have exactly the same form, and are derived in exactly the same way, as the expenditure function in Section 3.E (see Example 3.E.1; the only difference in the computations is that we now do not impose $\alpha + \beta = 1$):

$$z_1(w_1, w_2, q) = q^{1/(\alpha+\beta)}(\alpha w_2/\beta w_1)^{\beta/(\alpha+\beta)},$$
$$z_2(w_1, w_2, q) = q^{1/(\alpha+\beta)}(\beta w_1/\alpha w_2)^{\alpha/(\alpha+\beta)},$$

and

$$c(w_1, w_2, q) = q^{1/(\alpha+\beta)}[(a/\beta)^{\beta/(\alpha+\beta)} + (\alpha/\beta)^{-\alpha/(\alpha+\beta)}]\, w_1^{\alpha/(\alpha+\beta)} w_2^{\beta/(\alpha+\beta)}.$$

This cost function has the form $c(w_1, w_2, q) = q^{1/(\alpha+\beta)}\theta\phi(w_1, w_2)$, where

$$\theta = [(\alpha/\beta)^{\beta/(\alpha+\beta)} + (\alpha/\beta)^{-\alpha/(\alpha+\beta)}]$$

is a constant and $\phi(w_1, w_2) = w_1^{\alpha/(\alpha+\beta)} w_2^{\beta/(\alpha+\beta)}$ is a function that does not depend on the output level q. When we have constant returns, $\theta\phi(w_1, w_2)$ is the per-unit cost of production.

One way to derive the firm's supply function and profit function is to use this cost function and solve problem (5.C.5). Applying (5.C.6), the first-order condition for this problem is

$$p \leq \theta\phi(w_1, w_2)\left(\frac{1}{\alpha + \beta}\right) q^{(1/(\alpha+\beta))-1}, \quad \text{with equality if } q > 0 \qquad (5.C.7)$$

The first-order condition (5.C.7) is sufficient for a maximum when $\alpha + \beta \leq 1$ because the firm's cost function is then convex in q.

When $\alpha + \beta < 1$, (5.C.7) can be solved for a unique optimal output level:

$$q(w_1, w_2, p) = (\alpha + \beta)[p/\theta\phi(w_1, w_2)]^{(\alpha+\beta)/(1-\alpha-\beta)}.$$

The factor demands can then be obtained through substitution,

$$z_\ell(w_1, w_2, p) = z_\ell(w_1, w_2, q(w_1, w_2, p)) \quad \text{for } \ell = 1, 2,$$

as can the profit function,

$$\pi(w_1, w_2, p) = pq(w_1, w_2, p) - w \cdot z(w_1, w_2, q(w_1, w_2, p)).$$

When $\alpha + \beta = 1$, the right-hand side of the first-order condition (5.C.7) becomes $\theta\phi(w_1, w_2)$, the unit cost of production (which is independent of q). If $\theta\phi(w_1, w_2)$ is greater than p, then $q = 0$ is optimal; if it is smaller than p, then no solution exists (again, unbounded profits can be obtained by increasing q); and when $\theta\phi(w_1, w_2) = p$, any non-negative output level is a solution to the PMP and generates zero profits.

Finally, when $\alpha + \beta > 1$ (so that we have increasing returns to scale), a quantity q satisfying the first-order condition (5.C.7) does not yield a profit-maximizing production. [Actually, in this case, the cost function is strictly concave in q, so that any solution to the first-order condition (5.C.7) yields a local *minimum* of profits, subject to output being always produced at minimum cost]. Indeed, since $p > 0$, a doubling of the output level starting from any q doubles the firm's revenue but increases input costs only by a factor of $2^{1/(\alpha+\beta)} > 2$. With enough doublings, the firm's profits can therefore be made arbitrarily large. Hence, with increasing returns to scale, there is no solution to the PMP. ∎

5.D The Geometry of Cost and Supply in the Single-Output Case

In this section, we continue our analysis of the relationships among a firm's technology, its cost function, and its supply behavior for the special but commonly used case in which there is a single output. A significant advantage of considering the single-output case is that it lends itself to extensive graphical illustration.

Throughout, we denote the amount of output by q and hold the vector of factor prices constant at $\bar{w} \gg 0$. For notational convenience, we write the firm's cost function as $C(q) = c(\bar{w}, q)$. For $q > 0$, we can denote the firm's average cost by $AC(q) = C(q)/q$ and assuming that the derivative exists, we denote its *marginal cost* by $C'(q) = dC(q)/dq$.

Recall from expression (5.C.6) that for a given output price p, all profit-maximizing output levels $q \in q(p)$ must satisfy the first-order condition [assuming that $C'(q)$ exists]:

$$p \leq C'(q) \quad \text{with equality if } q > 0. \tag{5.D.1}$$

If the production set Y is convex, $C(\cdot)$ is a convex function [see property (ix) of Proposition 5.C.2], and therefore marginal cost is nondecreasing. In this case, as we noted in Section 5.C, satisfaction of this first-order condition is also sufficient to establish that q is a profit-maximizing output level at price p.

Two examples of convex production sets are given in Figures 5.D.1 and 5.D.2. In the figures, we assume that there is only one input, and we normalize its price to equal 1 (you can think of this input as the total expense of factor use).[10] Figure 5.D.1 depicts the production set (a), cost function (b), and average and marginal cost functions (c) for a case with decreasing returns to scale. Observe that the cost function is obtained from the production set by a 90-degree rotation. The determination of average cost and marginal cost from the cost function is shown in Figure 5.D.1(b) (for an output level $\hat{q}$). Figure 5.D.2 depicts the same objects for a case with constant returns to scale.

In Figures 5.D.1(c) and 5.D.2(c), we use a heavier trace to indicate the firm's profit-maximizing supply locus, the graph of $q(\cdot)$. (*Note*: *In this and subsequent figures, the supply locus is always indicated by a heavier trace.*) Because the technologies in these two examples are convex, the supply locus in each case coincides exactly with the (q, p) combinations that satisfy the first-order condition (5.D.1).

If the technology is not convex, perhaps because of the presence of some underlying indivisibility, then satisfaction of the first-order necessary condition

10. Thus, the single input can be thought of as a Hicksian composite commodity in a sense analogous to that in Exercise 3.G.5.

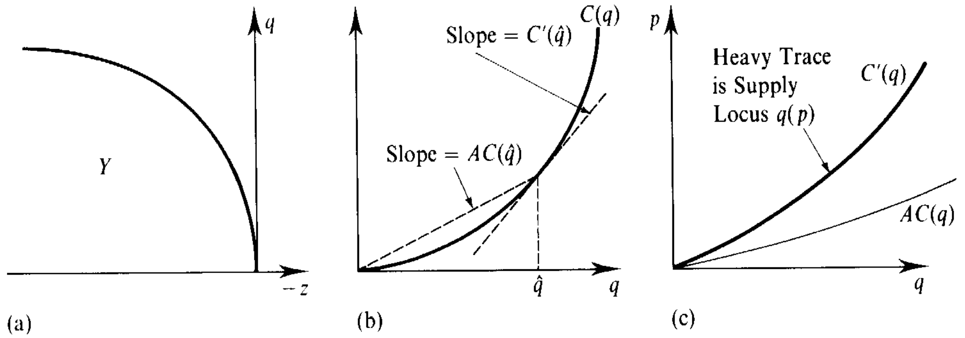


Figure 5.D.1
A strictly convex technology (strictly decreasing returns to scale).
(a) Production set.
(b) Cost function.
(c) Average cost, marginal cost, and supply.

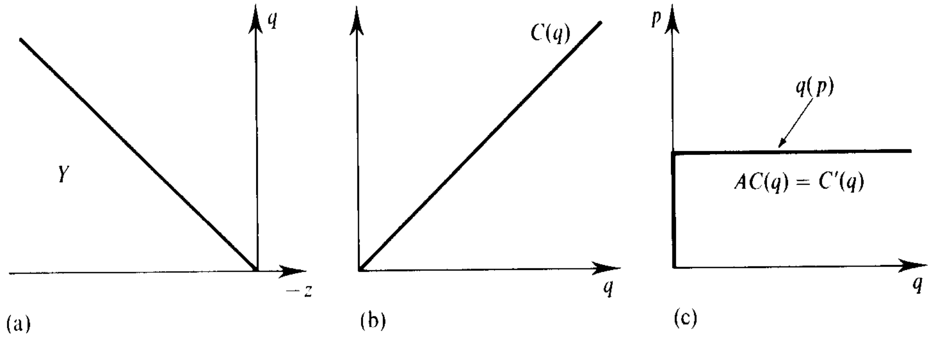


Figure 5.D.2
A constant returns to scale technology.
(a) Production set.
(b) Cost function.
(c) Average cost, marginal cost, and supply.

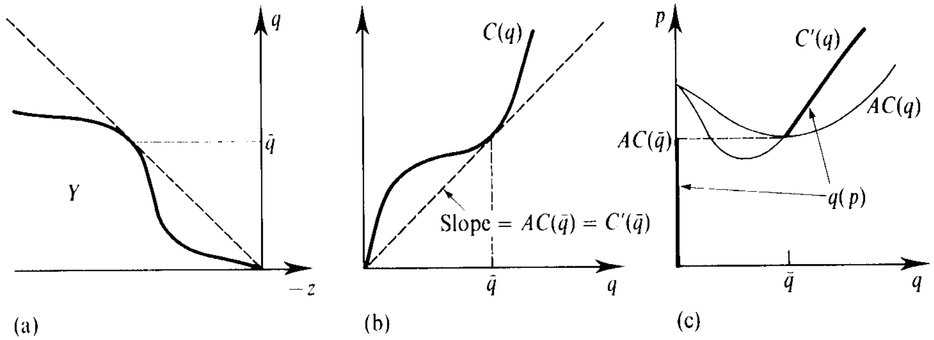


Figure 5.D.3
A nonconvex technology.
(a) Production set.
(b) Cost function.
(c) Average cost, marginal cost, and supply.

(5.D.1) no longer implies that q is profit maximizing. The supply locus will then be only a subset of the set of (q, p) combinations that satisfy (5.D.1).

Figure 5.D.3 depicts a situation with a nonconvex technology. In the figure, we have an initial segment of increasing returns over which the average cost decreases and then a region of decreasing returns over which the average cost increases. The level (or levels) of production corresponding to the minimum average cost is called the *efficient scale*, which, if unique, we denote by $\bar{q}$. Looking at the cost functions in Figure 5.D.3(a) and (b), we see that at $\bar{q}$ we have $AC(\bar{q}) = C'(\bar{q})$. In Exercise 5.D.1, you are asked to establish this fact as a general result.

Exercise 5.D.1: Show that $AC(\bar{q}) = C'(\bar{q})$ at any $\bar{q}$ satisfying $AC(\bar{q}) \leq AC(q)$ for all q. Does this result depend on the differentiability of $C(\cdot)$ everywhere?

The supply locus for this nonconvex example is depicted by the heavy trace in

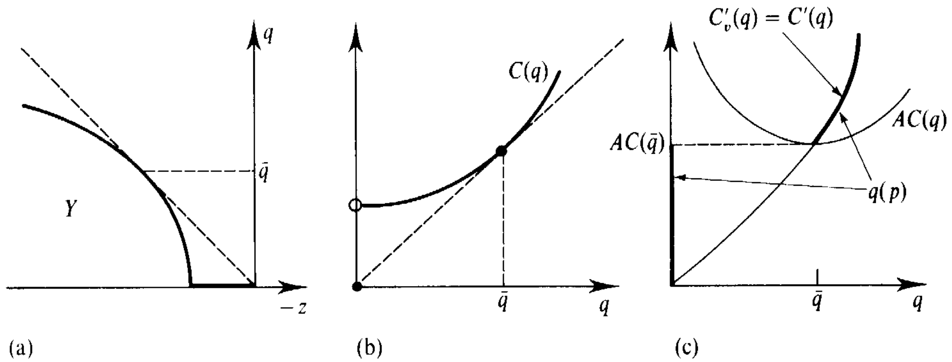


Figure 5.D.4
Strictly convex variable costs with a nonsunk setup cost. (a) Production set. (b) Cost function. (c) Average cost, marginal cost, and supply.

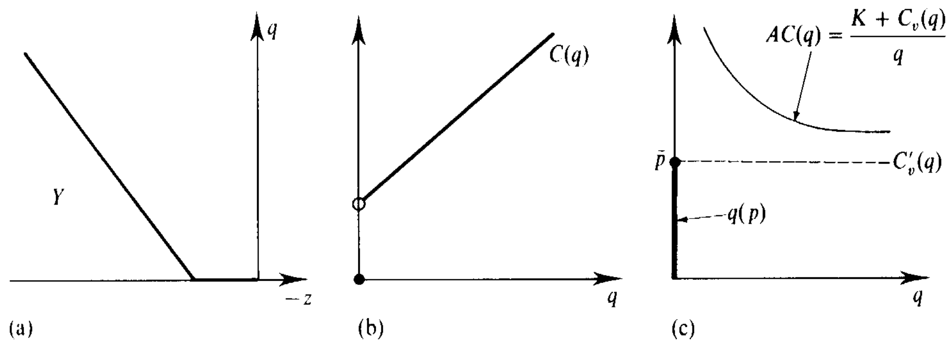


Figure 5.D.5
Constant returns variable costs with a nonsunk setup cost. (a) Production set. (b) Cost function. (c) Average cost, marginal cost, and supply.

Figure 5.D.3(c). When $p > AC(\bar{q})$, the firm maximizes its profit by producing at the unique level of q satisfying $p = C'(q) > AC(q)$. [Note that the firm earns strictly positive profits doing so, exceeding the zero profits earned by setting $q = 0$, which in turn exceed the strictly negative profits earned by choosing any $q > 0$ with $p = C'(q) < AC(q)$.] On the other hand, when $p < AC(\bar{q})$, any $q > 0$ earns strictly negative profits, and so the firm's optimal supply is $q = 0$ [note that $q = 0$ satisfies the necessary first-order condition (5.D.1) because $p < C'(0)$]. When $p = AC(\bar{q})$, the profit-maximizing set of output levels is $\{0, \bar{q}\}$. The supply locus is therefore as shown in Figure 5.D.3(c).

An important source of nonconvexities is fixed setup costs. These may or may not be sunk. Figures 5.D.4 and 5.D.5 (which parallel 5.D.1 and 5.D.2) depict two cases with nonsunk fixed setup costs (so inaction is possible). In these figures, we consider a case in which the firm incurs a fixed cost K if and only if it produces a positive amount of output and otherwise has convex costs. In particular, total cost is of the form $C(0) = 0$, and $C(q) = C_v(q) + K$ for $q > 0$, where $K > 0$ and $C_v(q)$, the *variable cost function*, is convex [and has $C_v(0) = 0$]. Figure 5.D.4 depicts the case in which $C_v(\cdot)$ is strictly convex, whereas $C_v(\cdot)$ is linear in Figure 5.D.5. The supply loci are indicated in the figures. In both illustrations, the firm will produce a positive amount of output only if its profit is sufficient to cover not only its variable costs but also the fixed cost K. You should read the supply locus in Figure 5.D.5(c) as saying that for $p > \bar{p}$, the supply is "infinite," and that $q = 0$ is optimal for $p \leq \bar{p}$.

In Figure 5.D.6, we alter the case studied in Figure 5.D.4 by making the fixed costs sunk, so that $C(0) > 0$. In particular, we now have $C(q) = C_v(q) + K$ for all $q \geq 0$; therefore, the firm must pay K whether or not it produces a positive quantity.

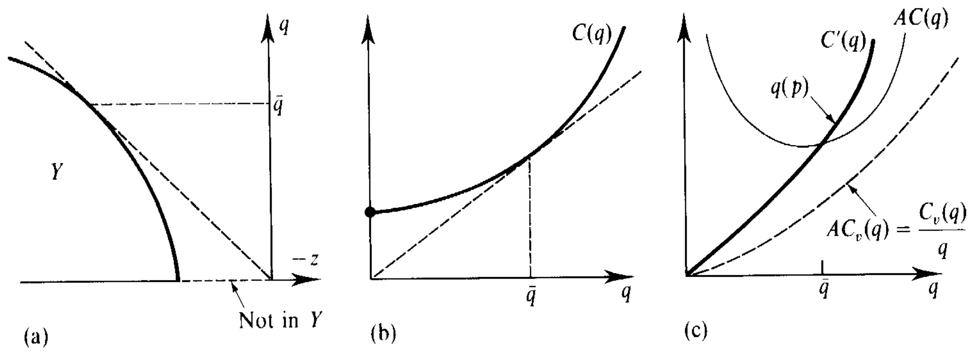


Figure 5.D.6 Strictly convex variable costs with sunk costs. (a) Production set. (b) Cost function. (c) Average cost, marginal cost, and supply.

Although inaction is not possible here, the firm's cost function is convex, and so we are back to the case in which the first-order condition (5.D.1) is sufficient. Because the firm must pay K regardless of whether it produces a positive output level, it will not shut down simply because profits are negative. Note that because $C_v(\cdot)$ is convex and $C_v(0) = 0$, $p = C_v'(q)$ implies that $pq > C_v(q)$; hence, the firm covers its variable costs when it sets output to satisfy its first-order condition. The firm's supply locus is therefore that depicted in Figure 5.D.6(c). Note that its supply behavior is exactly the same as if it did not have to pay the sunk cost K at all [compare with Figure 5.D.1(c)].

Exercise 5.D.2: Depict the supply locus for a case with partially sunk costs, that is, where $C(q) = K + C_v(q)$ if $q > 0$ and $0 < C(0) < K$.

As we noted in Section 5.B, one source of sunk costs, at least in the short run, is input choices irrevocably set by prior decisions. Suppose, for example, that we have two inputs and a production function $f(z_1, z_2)$. Recall that we keep the prices of the two inputs fixed at $(\bar{w}_1, \bar{w}_2)$. In Figure 5.D.7(a), the cost function excluding any prior input commitments is depicted by $C(\cdot)$. We call it the *long-run cost function*. If one input, say z_2, is fixed at level $\bar{z}_2$ in the short-run, then the *short-run cost function* of the firm becomes $C(q|\bar{z}_2) = \bar{w}_1 z_1 + \bar{w}_2 \bar{z}_2$, where z_1 is chosen so that $f(z_1, \bar{z}_2) = q$. Several such short-run cost functions corresponding to different levels of z_2 are illustrated in Figure 5.D.7(a). Because restrictions on the firm's input decisions can only increase its costs of production, $C(q|\bar{z}_2)$ lies above $C(q)$ at all q except the q for

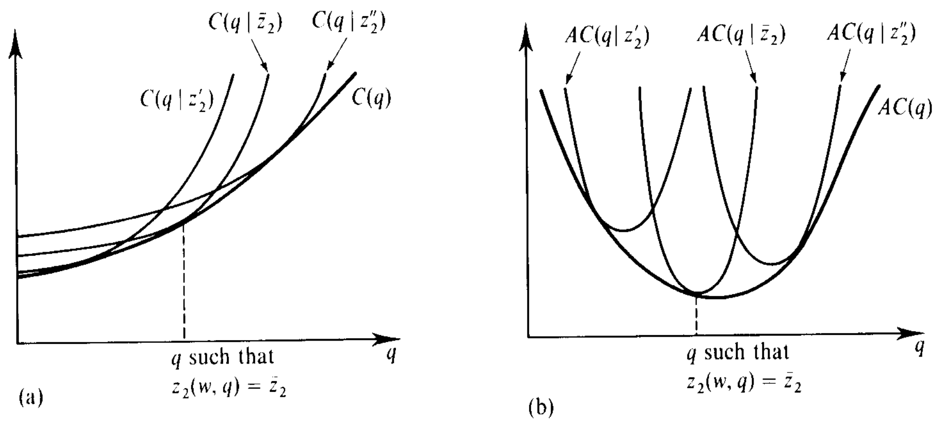


Figure 5.D.7 Costs when an input level is fixed in the short run but is free to vary in the long run. (a) Long-run and short-run cost functions. (b) Long-run and short-run average cost.

which $\bar{z}_2$ is the optimal long-run input level [i.e., the q such that $z_2(\bar{w}, q) = \bar{z}_2$]. Thus, $C(q|z_2(\bar{w},q)) = C(q)$ for all q. It follows from this and from the fact that $C(q'|z_2(\bar{w}, q)) \geq C(q')$ for all q', that $C'(q) = C'(q|z_2(\bar{w}, q))$ for all q; that is, if the level of z_2 is at its long-run value, then the short-run marginal cost equals the long-run marginal cost. Geometrically, $C(\cdot)$ is the lower envelope of the family of short-run functions $C(q|z_2)$ generated by letting z_2 take all possible values.

Observe finally that given the long-run and short-run cost functions, the long-run and short-run average cost functions and long-run and short-run supply functions of the firm can be derived in the manner discussed earlier in the section. The average-cost version of Figure 5.D.7(a) is given in Figure 5.D.7(b). (Exercise 5.D.3 asks you to investigate the short-run and long-run supply behavior of the firm in more detail.)

5.E Aggregation

In this section, we study the theory of aggregate (net) supply. As we saw in Section 5.C, the absence of a budget constraint implies that individual supply is not subject to wealth effects. As prices change, there are only substitution effects along the production frontier. In contrast with the theory of aggregate demand, this fact makes for an aggregation theory that is simple and powerful.[11]

Suppose there are J production units (firms or, perhaps, plants) in the economy, each specified by a production set $Y_1, \ldots, Y_J$. We assume that each Y_j is nonempty, closed, and satisfies the free disposal property. Denote the profit function and supply correspondences of Y_j by $\pi_j(p)$ and $y_j(p)$, respectively. The *aggregate supply correspondence* is the sum of the individual supply correspondences:

$$y(p) = \sum_{j=1}^{J} y_j(p) = \{y \in \mathbb{R}^L: y = \textstyle\sum_j y_j \text{ for some } y_j \in y_j(p), j = 1, \ldots, J\}.$$

Assume, for a moment, that every $y_j(\cdot)$ is a single-valued, differentiable function at a price vector p. From Proposition 5.C.1, we know that every $Dy_j(p)$ is a symmetric, positive semidefinite matrix. Because these two properties are preserved under addition, we can conclude that the matrix $Dy(p)$ is *symmetric and positive semidefinite.*

As in the theory of individual production, the positive semidefiniteness of $Dy(p)$ implies the *law of supply* in the aggregate: If a price increases, then so does the corresponding *aggregate* supply. As with the law of supply at the firm level, this property of aggregate supply holds for *all* price changes. We can also prove this aggregate law of supply directly because we know from (5.C.3) that $(p - p')\cdot[y_j(p) - y_j(p')] \geq 0$ for every j; therefore, adding over j, we get

$$(p - p')\cdot[y(p) - y(p')] \geq 0.$$

The symmetry of $Dy(p)$ suggests that underlying $y(p)$ there is a "representative producer." As we now show, this is true in a particularly strong manner.

Given $Y_1, \ldots, Y_J$, we can define the *aggregate production set* by

$$Y = Y_1 + \cdots + Y_J = \{y \in \mathbb{R}^L: y = \textstyle\sum_j y_j \text{ for some } y_j \in Y_j, j = 1, \ldots, J\}.$$

11. A classical and very readable account for the material in this section and in Section 5.F is Koopmans (1957).

The aggregate production set Y describes the production vectors that are feasible in the aggregate if all the production sets are used together. Let $\pi^*(p)$ and $y^*(p)$ be the profit function and the supply correspondence of the aggregate production set Y. They are the profit function and supply correspondence that would arise if a single price-taking firm were to operate, under the same management so to speak, all the individual production sets.

Proposition 5.E.1 establishes a strong aggregation result for the supply side: *The aggregate profit obtained by each production unit maximizing profit separately taking prices as given is the same as that which would be obtained if they were to coordinate their actions* (i.e., *their* y_js) *in a joint profit maximizing decision.*

Proposition 5.E.1: For all $p \gg 0$, we have

(i) $\pi^*(p) = \sum_j \pi_j(p)$

(ii) $y^*(p) = \sum_j y_j(p)$ $(= \{\sum_j y_j : y_j \in y_j(p) \text{ for every } j\})$.

Proof: (i) For the first equality, note that if we take any collection of production plans $y_j \in Y_j$, $j = 1, \dots, J$, then $\sum_j y_j \in Y$. Because $\pi^*(\cdot)$ is the profit function associated with Y, we therefore have $\pi^*(p) \geq p \cdot (\sum_j y_j) = \sum_j p \cdot y_j$. Hence, it follows that $\pi^*(p) \geq \sum_j \pi_j(p)$. In the other direction, consider any $y \in Y$. By the definition of the set Y, there are $y_j \in Y_j, j = 1, \dots, J$, such that $\sum_j y_j = y$. So $p \cdot y = p \cdot (\sum_j y_j) = \sum_j p \cdot y_j \leq \sum_j \pi_j(p)$ for all $y \in Y$. Thus, $\pi^*(p) \leq \sum_j \pi_j(p)$. Together, these two inequalities imply that $\pi^*(p) = \sum_j \pi(p)$.

(ii) For the second equality, we must show that $\sum_j y_j(p) \subset y^*(p)$ and that $y^*(p) \subset \sum_j y_j(p)$. For the former relation, consider any set of individual production plans $y_j \in y_j(p), j = 1, \dots, J$. Then $p \cdot (\sum_j y_j) = \sum_j p \cdot y_j = \sum_j \pi_j(p) = \pi^*(p)$, where the last equality follows from part (i) of the proposition. Hence, $\sum_j y_j \in y^*(p)$, and therefore, $\sum_j y_j(p) \subset y^*(p)$. In the other direction, take any $y \in y^*(p)$. Then $y = \sum_j y_j$ for some $y_j \in Y_j, j = 1, \dots, J$. Since $p \cdot (\sum_j y_j) = \pi^*(p) = \sum_j \pi_j(p)$ and, for every j, we have $p \cdot y_j \leq \pi_j(p)$, it must be that $p \cdot y_j = \pi_j(p)$ for every j. Thus, $y_j \in y_j(p)$ for all j, and so $y \in \sum_j y_j(p)$. Thus, we have shown that $y^*(p) \subset \sum_j y_j(p)$. ∎

The content of Proposition 5.E.1 is illustrated in Figure 5.E.1. The proposition can be interpreted as a decentralization result: To find the solution of the aggregate profit maximization problem for given prices p, it is enough to add the solutions of the corresponding individual problems.

Simple as this result may seem, it nevertheless has many important implications. Consider, for example, the single-output case. The result tells us that if firms are maximizing profit facing output price p and factor prices w, then their supply behavior maximizes aggregate profits. But this must mean that if $q = \sum_j q_j$ is the aggregate output produced by the firms, then the total cost of production is exactly equal to $c(w, q)$, the value of the *aggregate cost function* (the cost function corresponding to the aggregate production set Y). *Thus, the allocation of the production of output level q among the firms is cost minimizing.* In addition, this allows us to relate the firms' aggregate supply function for output $q(p)$ to the aggregate cost function in the same manner as done in Section 5.D for an individual firm. (This fact will prove useful when we study partial equilibrium models of competitive markets in Chapter 10.)

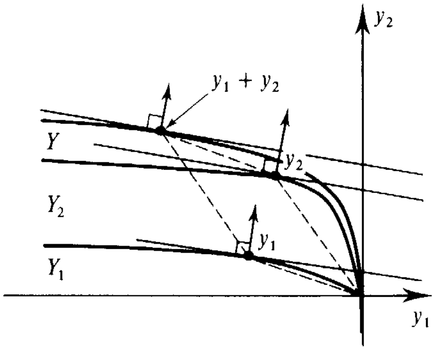


Figure 5.E.1
Joint profit maximization as a result of individual profit maximization.

In summary: If firms maximize profits taking prices as given, then the production side of the economy aggregates beautifully.

As in the consumption case (see Appendix A of Chapter 4), aggregation can also have helpful regularizing effects in the production context. An interesting and important fact is that the existence of many firms or plants with technologies that are not too dissimilar can make the *average* production set almost convex, even if the individual production sets are not so. This is illustrated in Figure 5.E.2, where there are J firms with identical production sets equal to

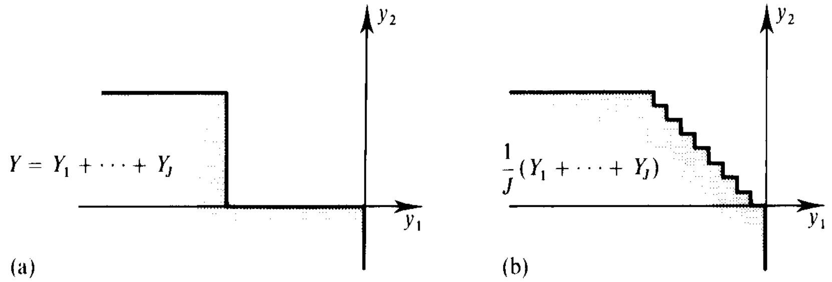


Figure 5.E.2
An example of the convexifying effects of aggregation.
(a) The individual production set.
(b) The average production set.

that displayed in 5.E.2(a). Defining the average production set as $(1/J)(Y_1 + \cdots + Y_J) = \{y: y = (1/J)(y_1 + \cdots + y_J)$ for some $y_j \in Y_j, j = 1, \ldots, J\}$, we see that for large J, this set is nearly convex, as depicted in Figure 5.E.2(b).[12]

5.F Efficient Production

Because much of welfare economics focuses on efficiency (see, for example, Chapters 10 and 16), it is useful to have algebraic and geometric characterizations of productions plans that can unambiguously be regarded as nonwasteful. This motivates Definition 5.F.1.

12. Note that this production set is bounded above. This is important because it insures that the individual nonconvexity is of finite size. If the individual production set was like that shown in, say, Figure 5.B.4, where neither the set nor the nonconvexity is bounded, then the average set would display a large nonconvexity (for any J). In Figure 5.B.5, we have a case of an unbounded production set but with a bounded nonconvexity; as for Figure 5.E.2, the average set will in this case be almost convex.

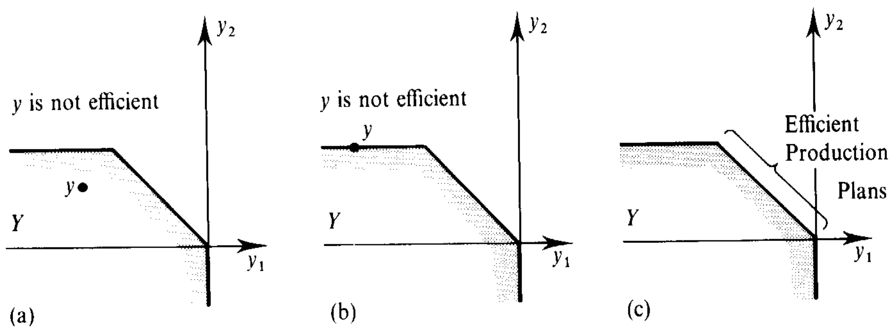


Figure 5.F.1
An efficient production plan must be on the boundary of Y, but not all points on the boundary of Y are efficient.
(a) An inefficient production plan in the interior of Y.
(b) An inefficient production plan at the boundary of Y.
(c) The set of efficient production plans.

Definition 5.F.1: A production vector $y \in Y$ is *efficient* if there is no $y' \in Y$ such that $y' \geq y$ and $y' \neq y$.

In words, a production vector is efficient if there is no other feasible production vector that generates as much output as y using no additional inputs, and that actually produces more of some output or uses less of some input.

As we see in Figure 5.F.1, every efficient y must be on the boundary of Y, but the converse is not necessarily the case: There may be boundary points of Y that are not efficient.

We now show that the concept of efficiency is intimately related to that of supportability by profit maximization. This constitutes our first look at a topic that we explore in much more depth in Chapter 10 and especially in Chapter 16

Proposition 5.F.1 provides an elementary but important result. It is a version of the *first fundamental theorem of welfare economics.*

Proposition 5.F.1: If $y \in Y$ is profit maximizing for some $p \gg 0$, then y is efficient.

Proof: Suppose otherwise: That there is a $y' \in Y$ such that $y' \neq y$ and $y' \geq y$. Because $p \gg 0$, this implies that $p \cdot y' > p \cdot y$, contradicting the assumption that y is profit maximizing. ■

It is worth emphasizing that Proposition 5.F.1 is valid even if the production set is nonconvex. This is illustrated in Figure 5.F.2.

When combined with the aggregation results discussed in Section 5.E, Proposition 5.F.1 tells us that *if a collection of firms each independently maximizes profits with respect to the same fixed price vector* $p \gg 0$, *then the aggregate production is*

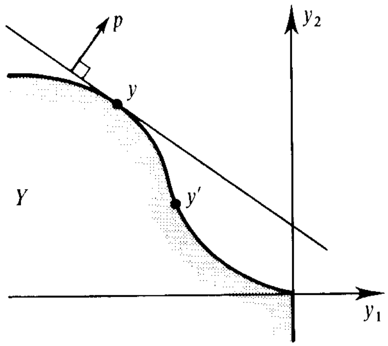


Figure 5.F.2
A profit-maximizing production plan (for $p \gg 0$) is efficient.

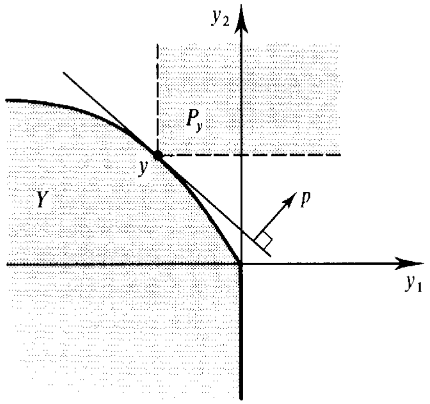


Figure 5.F.3
The use of the separating hyperplane theorem to prove Proposition 5.F.2: If Y is convex, every efficient $y \in Y$ is profit maximizing for some $p \geq 0$.

socially efficient. That is, there is no other production plan for the economy as a whole that could produce more output using no additional inputs. This is in line with our conclusion in Section 5.E that, in the single-output case, the aggregate output level is produced at the lowest-possible cost when all firms maximize profits facing the same prices.

The need for strictly positive prices in Proposition 5.F.1 is unpleasant, but it cannot be dispensed with, as Exercise 5.F.1 asks you to demonstrate.

Exercise 5.F.1: Give an example of a $y \in Y$ that is profit maximizing for some $p \geq 0$ with $p \neq 0$ but that is also inefficient (i.e. not efficient).

A converse of Proposition 5.F.1 would assert that any efficient production vector is profit maximizing for *some* price system. However, a glance at the efficient production y' in Figure 5.F.2 shows that this cannot be true in general. Nevertheless, this converse does hold with the added assumption of convexity. Proposition 5.F.2, which is less elementary than Proposition 5.F.1, is a version of the so-called *second fundamental theorem of welfare economics*.

Proposition 5.F.2: Suppose that Y is convex. Then every efficient production $y \in Y$ is a profit-maximizing production for some nonzero price vector $p \geq 0$.[13]

Proof: This proof is an application of the separating hyperplane theorem for convex sets (see Section M.G of the Mathematical Appendix). Suppose that $y \in Y$ is efficient, and define the set $P_y = \{y' \in \mathbb{R}^L: y' \gg y\}$. The set P_y is depicted in Figure 5.F.3. It is convex, and because y is efficient, we have $Y \cap P_y = \varnothing$. We can therefore invoke the separating hyperplane theorem to establish that there is *some* $p \neq 0$ such that $p \cdot y' \geq p \cdot y''$ for every $y' \in P_y$ and $y'' \in Y$ (see Figure 5.F.3). Note, in particular, that this implies $p \cdot y' \geq p \cdot y$ for every $y' \gg y$. Therefore, we must have $p \geq 0$ because if $p_\ell < 0$ for some ℓ, then we would have $p \cdot y' < p \cdot y$ for some $y' \gg y$ with $y'_\ell - y_\ell$ sufficiently large.

Now take any $y'' \in Y$. Then $p \cdot y' \geq p \cdot y''$ for every $y' \in P_y$. Because y' can be chosen to be arbitrarily close to y, we conclude that $p \cdot y \geq p \cdot y''$ for any $y'' \in Y$; that is, y is profit maximizing for p. ■

13. As the proof makes clear, the result also applies to *weakly efficient* productions, that is, to productions such as y in Figure 5.F.1(b) where there is no $y' \in Y$ such that $y' \gg y$.

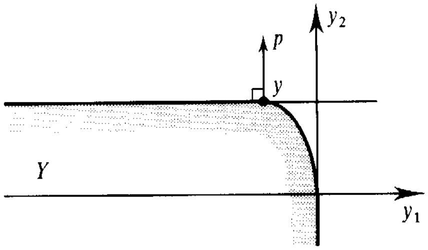


Figure 5.F.4
Proposition 5.C.2 cannot be extended to require $p \gg 0$.

The second part of Proposition 5.F.2 cannot be strengthened to read "$p \gg 0$." In Figure 5.F.4, for example, the production vector y is efficient, but it cannot be supported by any strictly positive price vector.

As an illustration of Proposition 5.F.2, consider a single-output, concave production function $f(z)$. Fix an input vector $\bar{z}$, and suppose that $f(\cdot)$ is differentiable at $\bar{z}$ and $\nabla f(\bar{z}) \gg 0$. Then the production plan that uses input vector $\bar{z}$ to produce output level $f(\bar{z})$ is efficient. Letting the price of output be 1, condition (5.C.2) tells us that the input price vector that makes this efficient production profit maximizing is precisely $w = \nabla f(\bar{z})$, the vector of marginal productivities.

5.G Remarks on the Objectives of the Firm

Although it is logical to take the assumption of preference maximization as a primitive concept for the theory of the consumer, the same cannot be said for the assumption of profit maximization by the firm. Why this objective rather than, say, the maximization of sales revenues or the size of the firm's labor force? The objectives of the firm assumed in our economic analysis should emerge from the objectives of those individuals who control it. Firms in the type of economies we consider are owned by individuals who, wearing another hat, are also consumers. A firm owned by a single individual has well-defined objectives: those of the owner. In this case, the only issue is whether this objective coincides with profit maximization. Whenever there is more than one owner, however, we have an added level of complexity. Indeed, we must either reconcile any conflicting objectives the owners may have or show that no conflict exists.

Fortunately, it is possible to resolve these issues and give a sound theoretical grounding to the objective of profit maximization. We shall now show that under reasonable assumptions this is the goal that all owners would agree upon.

Suppose that a firm with production set Y is owned by consumers. Ownership here simply means that each consumer $i = 1, \ldots, I$ is entitled to a share $\theta_i \geq 0$ of profits, where $\sum_i \theta_i = 1$ (some of the θ_i's may equal zero). Thus, if the production decision is $y \in Y$, then a consumer i with utility function $u_i(\cdot)$ achieves the utility level

$$\begin{aligned} \operatorname*{Max}_{x_i \geq 0} \quad & u_i(x_i) \\ \text{s.t.} \quad & p \cdot x_i \leq w_i + \theta_i p \cdot y, \end{aligned}$$

where w_i is consumer i's nonprofit wealth. Hence at fixed prices, higher profit increases consumer–owner i's overall wealth and expands her budget set, a desirable outcome. It follows that at any fixed price vector p, the consumer–owners *unanimously* prefer that the firm implement a production plan $y' \in Y$ instead of $y \in Y$ whenever $p \cdot y' > p \cdot y$. Hence, we conclude that if we maintain the assumption of price-taking behavior, all owners would agree, whatever their utility functions, to instruct the manager of the firm to maximize profits.[14]

It is worth emphasizing three of the implicit assumptions in the previous reasoning: (i) prices are fixed and do not depend on the actions of the firm, (ii) profits are not uncertain, and (iii) managers can be controlled by owners. We comment on these assumptions very informally.

(i) If prices may depend on the production of the firm, the objective of the owners may depend on their tastes as consumers. Suppose, for example, that each consumer has no wealth from sources other than the firm ($w_i = 0$), that $L = 2$, and that the firm produces good 1 from good 2 with production function $f(\cdot)$. Also, normalize the price of good 2 to be 1, and suppose that the price of good 1, in terms of good 2, is $p(q)$ if output is q. If, for example, the preferences of the owners are such that they care only about the consumption of good 2, then they will unanimously want to solve $\text{Max}_{z \geq 0}\ p(f(z))f(z) - z$. This maximizes the amount of good 2 that they get to consume. On the other hand, if they want to consume only good 1, then they will wish to solve $\text{Max}_{z \geq 0}\ f(z) - [z/p(f(z))]$ because if they earn $p(f(z))f(z) - z$ units of good 2, then end up with $[p(f(z))f(z) - z]/p(f(z))$ units of good 1. But these two problems have different solutions. (Check the first-order conditions.) Moreover, as this suggests, if the owners differ in their tastes as consumers, then they will not agree about what they want the firm to do (Exercise 5.G.1 elaborates on this point.)

(ii) If the output of the firm is random, then it is crucial to distinguish whether the output is sold before or after the uncertainty is resolved. If the output is sold after the uncertainty is resolved (as in the case of agricultural products sold in spot markets after harvesting), then the argument for a unanimous desire for profit maximization breaks down. Because profit, and therefore derived wealth, are now uncertain, the risk attitudes and expectations of owners will influence their preferences with regard to production plans. For example, strong risk averters will prefer relatively less risky production plans than moderate risk averters.

On the other hand, if the output is sold before uncertainty is resolved (as in the case of agricultural products sold in futures markets before harvesting), then the risk is fully carried by the buyer. The profit of the firm is not uncertain, and the argument for unanimity in favor of profit maximization still holds. In effect, the firm can be thought of as producing a commodity that is sold before uncertainty is resolved in a market of the usual kind. (Further analysis of this issue would take us too far afield. We come back to it in Section 19.G after covering the foundations of decision theory under uncertainty in Chapter 6.)

(iii) It is plain that shareholders cannot usually exercise control directly. They need managers, who, naturally enough, have their own objectives. Especially if ownership is very diffuse, it is an important theoretical challenge to understand how and to what extent managers are, or can be, controlled by owners. Some relevant considerations are factors such as the degree of observability of managerial actions

14. In actuality, there are public firms and quasipublic organizations such as universities that do not have *owners* in the sense that private firms have shareholders. Their objectives may be different, and the current discussion does not apply to them.

and the stake of individual owners. [These issues will be touched on in Section 14.C (agency contracts as a mechanism of internal control) and in Section 19.G (stock markets as a mechanism of external control).]

APPENDIX A: THE LINEAR ACTIVITY MODEL

The saliency of the model of production with convexity and constant returns to scale technologies recommends that we examine it in some further detail.

Given a constant returns to scale technology Y, the *ray* generated (or spanned) by a vector $\bar{y} \in Y$ is the set $\{y \in Y: y = \alpha\bar{y}$ for some scalar $\alpha \geq 0\}$. We can think of a ray as representing a production *activity* that can be run at any *scale of operation.* That is, the production plan $\bar{y}$ can be scaled up or down by any factor $\alpha \geq 0$, generating, in this way, other possible production plans.

We focus here on a particular case of constant returns to scale technologies that lends itself to explicit computation and is therefore very important in applications. We assume that we are given as a primitive of our theory a list of *finitely many activities* (say M), each of which can be run at any scale of operation and any number of which can be run simultaneously. Denote the M activities, to be called the *elementary activities*, by $a_1 \in \mathbb{R}^L, \dots, a_M \in \mathbb{R}^L$. Then, the production set is

$$Y = \{y \in \mathbb{R}^L: y = \sum_{m=1}^{M} \alpha_m a_m \text{ for some scalars } (\alpha_1, \dots, \alpha_M) \geq 0\}.$$

The scalar α_m is called the *level of elementary activity m*; it measures the scale of operation of the mth activity. Geometrically, Y is a *polyhedral cone*, a set generated as the convex hull of a finite number of rays.

An activity of the form $(0, \dots, 0, -1, 0, \dots, 0)$, where -1 is in the ℓth place, is known as the *disposal activity* for good ℓ. Henceforth, we shall always assume that, in addition to the M listed elementary activities, the L disposal activities are also available. Figure 5.AA.1 illustrates a production set arising in the case where $L = 2$ and $M = 2$.

Given a price vector $p \in \mathbb{R}^L_+$, a profit-maximizing plan exists in Y if and only if $p \cdot a_m \leq 0$ for every m. To see this, note that if $p \cdot a_m < 0$, then the profit-maximizing level of activity m is $\alpha_m = 0$. If $p \cdot a_m = 0$, then any level of activity m generates zero profits. Finally, if $p \cdot a_m > 0$ for some m, then by making α_m arbitrarily large, we could generate arbitrarily large profits. Note that the presence of the disposal activities implies that we must have $p \in \mathbb{R}^L_+$ for a profit-maximizing plan to exist. If $p_\ell < 0$, then the ℓth disposal activity would generate strictly positive (hence, arbitrarily large) profits.

For any price vector p generating zero profits, let $A(p)$ denote the set of activities that generate exactly zero profits: $A(p) = \{a_m: p \cdot a_m = 0\}$. If $a_m \notin A(p)$, then $p \cdot a_m < 0$, and so activity m is not used at prices p. The profit-maximizing supply set $y(p)$ is therefore the convex cone generated by the activities in $A(p)$; that is, $y(p) = \{\sum_{a_m \in A(p)} \alpha_m a_m: \alpha_m \geq 0\}$. The set $y(p)$ is also illustrated in Figure 5.AA.1. In the figure, at price vector p, activity a_1 makes exactly zero profits, and activity a_2

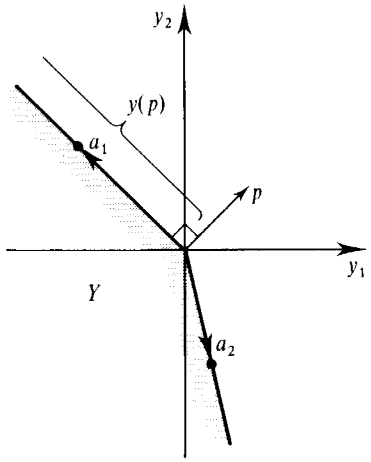


Figure 5.AA.1
A production set generated by two activities.

incurs a loss (if operated at all). Therefore, $A(p) = \{a_1\}$ and $y(p) = \{y: y = \alpha_1 a_1$ for any scalar $\alpha_1 \geq 0\}$, the ray spanned by activity a_1.

A significant result that we shall not prove is that for the linear activity model the converse of the efficiency Proposition 5.F.1 holds exactly; that is, we can strengthen Proposition 5.F.2 to say: *Every efficient* $y \in Y$ *is a profit-maximizing production for some* $p \gg 0$.

An important special case of the linear activity model is *Leontief's input–output model*. It is characterized by two additional features:

(i) There is one commodity, say the Lth, which is not produced by any activity. For this reason, we will call it the *primary factor*. In most applications of the Leontief model, the primary factor is labor.

(ii) Every elementary activity has at most a single positive entry. This is called the assumption of *no joint production*. Thus, it is as if every good except the primary factor is produced from a certain type of constant returns production function using the other goods and the primary factor as inputs.

The Leontief Input–Output Model with No Substitution Possibilities

The simplest Leontief model is one in which each producible good is produced by only one activity. In this case, it is natural to label the activity that produces good $\ell = 1, \ldots, L-1$ as $a_\ell = (a_{1\ell}, \ldots, a_{L\ell}) \in \mathbb{R}^L$. So the number of elementary activities M is equal to $L - 1$. As an example, in Figure 5.AA.2, for a case where $L = 3$, we represent the unit production isoquant [the set $\{(z_2, z_3): f(z_2, z_3) = 1\}$] for the implied production function of good 1. In the figure, the disposal activities for goods 2 and 3 are used to get rid of any excess of inputs. Because inputs must be used in fixed proportions (disposal aside), this special case is called a *Leontief model with no substitution possibilities*.

If we normalize the activity vectors so that $a_{\ell\ell} = 1$ for all $\ell = 1, \ldots, L-1$, then the vector $\alpha = (\alpha_1, \ldots, \alpha_{L-1}) \in \mathbb{R}^{L-1}$ of activity levels equals the vector of *gross* production of goods 1 through $L - 1$. To determine the levels of *net* production, it is convenient to denote by A the $(L-1) \times (L-1)$ matrix in which the ℓth column is

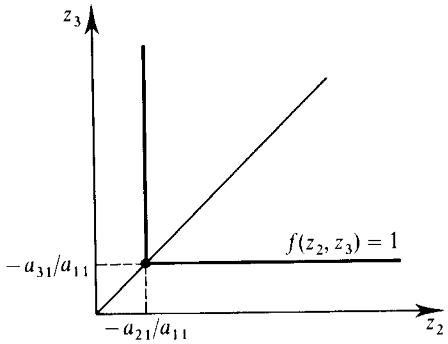


Figure 5.AA.2
Unit isoquant of production function for good 1 in the Leontief model with no substitution.

the negative of the activity vector a_ℓ except that its last entry has been deleted and entry $a_{\ell\ell}$ has been replaced by a zero (recall that entries $a_{k\ell}$ with $k \neq \ell$ are nonpositive):

$$A = \begin{bmatrix} 0 & -a_{12} & \cdots & -a_{1,L-1} \\ -a_{21} & 0 & \cdots & -a_{2,L-1} \\ \vdots & & \ddots & \\ -a_{L-1,1} & -a_{L-1,2} & \cdots & 0 \end{bmatrix}.$$

The matrix A is known as the *Leontief input–output matrix.* Its $k\ell$th entry, $-a_{k\ell} \geq 0$, measures how much of good k is needed to produce one unit of good ℓ. We also denote by $b \in \mathbb{R}^{L-1}$ the vector of primary factor requirements, $b = (-a_{L1}, \ldots, -a_{L,L-1})$. The vector $(I - A)\alpha$ then gives the *net* production levels of the $L - 1$ outputs when the activities are run at levels $\alpha = (\alpha_1, \ldots, \alpha_{L-1})$. To see this, recall that the activities are normalized so that the gross production levels of the $L - 1$ produced goods are exactly $\alpha = (\alpha_1, \ldots, \alpha_{L-1})$. On the other hand, $A\alpha$ gives the amounts of each of these goods that are used as inputs for other produced goods. The difference, $(I - A)\alpha$, is therefore the net production of goods $1, \ldots, L - 1$. In addition, the scalar $b\cdot\alpha$ gives the total use of the primary factor. In summary, with this notation, we can write the set of technologically feasible production vectors (assuming free disposal) as

$$Y = \left\{ y\colon y \leq \begin{bmatrix} I - A \\ -b \end{bmatrix} \alpha \text{ for some } \alpha \in \mathbb{R}^L_+ \right\}.$$

If $(I - A)\bar{\alpha} \gg 0$ for some $\bar{\alpha} \geq 0$, the input–output matrix A is said to be *productive.* That is, the input–output matrix A is productive if there is *some* production plan that can produce positive net amounts of the $L - 1$ outputs, provided only that there is a sufficient amount of primary input available.

A remarkable fact of Leontief input–output theory is the all-or-nothing property stated in Proposition 5.AA.1.

Proposition 5.AA.1: If A is productive, then for any nonnegative amounts of the $L - 1$ producible commodities $c \in \mathbb{R}^{L-1}_+$, there is a vector of activity levels $\alpha \geq 0$ such that $(I - A)\alpha = c$. That is, if A is productive, then it is possible to produce *any* nonnegative net amount of outputs (perhaps for purposes of final consumption), provided only that there is enough primary factor available.

Proof: We will show that if A is productive, then the inverse of the matrix $(I - A)$ exists and is nonnegative. This will give the result because we can then achieve net output levels $c \in \mathbb{R}_+^{L-1}$ by setting the (nonnegative) activity levels $\alpha = (I - A)^{-1}c$.

To prove the claim, we begin by establishing a matrix-algebra fact. We show that if A is productive, then the matrix $\sum_{n=0}^{N} A^n$, where A^n is the nth power of A, approaches a limit as $N \to \infty$. Because A has only nonnegative entries, every entry of $\sum_{n=0}^{N} A^n$ is *nondecreasing* with N. Therefore, to establish that $\sum_{n=0}^{N} A^n$ has a limit, it suffices to show that there is an upper bound for its entries. Since A is productive, there is an $\bar{\alpha}$ and $\bar{c} \gg 0$ such that $\bar{c} = (I - A)\bar{\alpha}$. If we premultiply both sides of this equality by $\sum_{n=0}^{N} A^n$, we get $(\sum_{n=0}^{N} A^n)\bar{c} = (I - A^{N+1})\bar{\alpha}$ (recall that $A^0 = I$). But $(I - A^{N+1})\bar{\alpha} \leq \bar{\alpha}$ because all elements of the matrix A^{N+1} are nonnegative. Therefore, $(\sum_{n=0}^{N} A^n)\bar{c} \leq \bar{\alpha}$. With $\bar{c} \gg 0$, this implies that no entry of $\sum_{n=0}^{N} A^n$ can exceed $[\text{Max}\{\bar{\alpha}_1, \ldots, \bar{\alpha}_{L-1}\}/\text{Min}\{\bar{c}_1, \ldots, \bar{c}_{L-1}\}]$, and so we have established the desired upper bound. We conclude, therefore, that $\sum_{n=0}^{\infty} A^n$ exists.

The fact that $\sum_{n=0}^{\infty} A^n$ exists must imply that $\lim_{N \to \infty} A^N = 0$. Thus, since $(\sum_{n=0}^{N} A^n)(I - A) = (I - A^{N+1})$ and $\lim_{N \to \infty} (I - A^{N+1}) = I$, it must be that $\sum_{n=0}^{\infty} A^n = (I - A)^{-1}$. (If A is a single number, this is precisely the high-school formula for adding up the terms of a geometric series.) The conclusion is that $(I - A)^{-1}$ exists and that all its entries are nonnegative. This establishes the result. ■

The focus on $\sum_{n=0}^{N} A^n$ in the proof of Proposition 5.AA.1 makes economic sense. Suppose we want to produce the vector of final consumptions $c \in \mathbb{R}_+^{L-1}$. How much total production will be needed? To produce final outputs $c = A^0 c$, we need to use as inputs the amounts $A(A^0 c) = Ac$ of produced goods. In turn, to produce these amounts requires that $A(Ac) = A^2 c$ of additional produced goods be used, and so on ad infinitum. The *total* amounts of goods required to be produced is therefore the limit of $(\sum_{n=0}^{N} A^n)c$ as $N \to \infty$. Thus, we can conclude that the vector $c \geq 0$ will be producible if and only if $\sum_{n=0}^{\infty} A^n$ is well defined (i.e., all its entries are finite).

Example 5.AA.1: Suppose that $L = 3$, and let $a_1 = (1, -1, -2)$ and $a_2 = (-\beta, 1, -4)$ for some constant $\beta \geq 0$. Activity levels $\alpha = (\alpha_1, \alpha_2)$ generate a positive net output of good 2 if $\alpha_2 > \alpha_1$; they generate a positive net output of good 1 if $\alpha_1 - \beta\alpha_2 > 0$. The input–output matrix A and the matrix $(I - A)^{-1}$ are

$$A = \begin{bmatrix} 0 & \beta \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad (I - A)^{-1} = \frac{1}{1 - \beta} \begin{bmatrix} 1 & \beta \\ 1 & 1 \end{bmatrix}.$$

Hence, matrix A is productive if and only if $\beta < 1$. Figure 5.AA.3(a) depicts a case where A is productive. The shaded region represents the vectors of net outputs that can be generated using the two activity vectors; note how the two activity vectors can span all of $\mathbb{R}_+^2$. In contrast, in Figure 5.AA.3(b), the matrix A is not productive: No strictly positive vector of net outputs can be achieved by running the two activities at nonnegative scales. [Again, the shaded region represents those vectors that can be generated using the two activity vectors, here a set whose only intersection with $\mathbb{R}_+^2$ is the point $(0, 0)$]. Note also that the closer β is to the value 1, the larger the levels of activity required to produce any final vector of consumptions. ■

The Leontief Model with Substitution Possibilities

We now move to the consideration of the general Leontief model in which each good may have more than one activity capable of producing it. We shall see that the

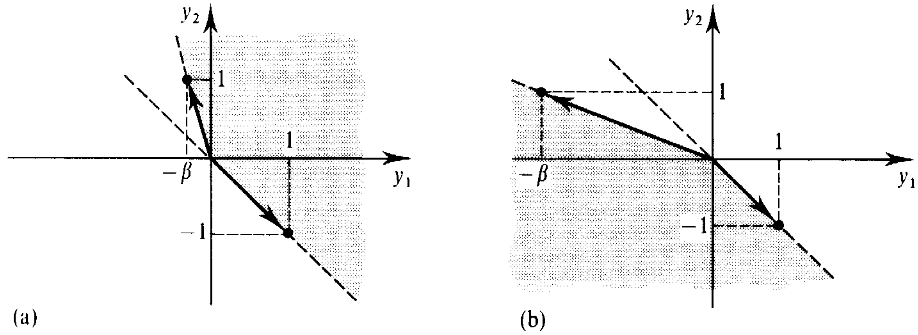


Figure 5.AA.3
Leontief model of Example 5.AA.1.
(a) Productive ($\beta < 1$).
(b) Unproductive ($\beta \geq 1$).

properties of the nonsubstitution model remain very relevant for the more general case where substitution is possible.

The first thing to observe is that the computation of the production function of a good, say good 1, now becomes a linear programming problem (see Section M.M of the Mathematical Appendix). Indeed, suppose that $a_1 \in \mathbb{R}^L, \ldots, a_{M_1} \in \mathbb{R}^L$ is a list of M_1 elementary activities capable of producing good 1 and that we are given initial levels of goods $2, \ldots, L$ equal to $z_2, \ldots, z_L$. Then the maximal possible production of good 1 given these available inputs $f(z_2, \ldots, z_L)$ is the solution to the problem

$$\begin{aligned} \underset{\alpha_1 \geq 0, \ldots, \alpha_{M_1} \geq 0}{\text{Max}} \quad & \alpha_1 a_{11} + \cdots + \alpha_{M_1} a_{1M_1} \\ \text{s.t.} \quad & \textstyle\sum_{m=1}^{M_1} \alpha_m a_{\ell m} \geq -z_\ell \quad \text{for all } \ell = 2, \ldots, L. \end{aligned}$$

We also know from linear programming theory that the $L-1$ dual variables $(\lambda_2, \ldots, \lambda_L)$ of this problem (i.e., the multipliers associated with the $L-1$ constraints) can be interpreted as the marginal productivities of the $L-1$ inputs. More precisely, for any $\ell = 2, \ldots, L$, we have $(\partial f/\partial z_\ell)^+ \leq \lambda_\ell \leq (\partial f/\partial z_\ell)^-$, where $(\partial f/\partial z_\ell)^+$ and $(\partial f/\partial z_\ell)^-$ are, respectively, the left-hand and right-hand ℓth partial derivatives of $f(\cdot)$ at $(z_2, \ldots, z_L)$.

Figure 5.AA.4 illustrates the unit isoquant for the case in which good 1 can be produced using two other goods (goods 2 and 3) as inputs with two possible activities $a_1 = (1, -2, -1)$ and $a_2 = (1, -1, -2)$. If the ratio of inputs is either higher than 2 or lower than $\frac{1}{2}$, one of the disposal activities is used to eliminate any excess inputs.

For any vector $y \in \mathbb{R}^L$, it will be convenient to write $y = (y_{-L}, y_L)$, where $y_{-L} = (y_1, \ldots, y_{L-1})$. We shall assume that our Leontief model is *productive* in the sense that there is a technologically feasible vector $y \in Y$ such that $y_{-L} \gg 0$.

A striking implication of the Leontief structure (constant returns, no joint products, single primary factor) is that we can associate with each good a *single optimal technique* (which could be a mixture of several of the elementary techniques corresponding to that good). What this means is that optimal techniques (one for each output) supporting efficient production vectors can be chosen independently of the particular output vector that is being produced (as long as the net output of every producible good is positive). Thus, although substitution is possible in principle, efficient production requires no substitution of techniques as desired final consumption levels change. This is the content of the celebrated *non-substitution theorem* (due to Samuelson [1951]).

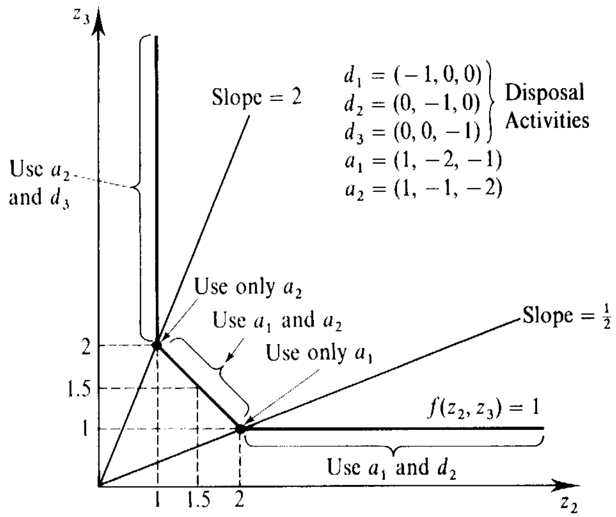


Figure 5.AA.4
Unit isoquant of production function of good 1, in the Leontief model with substitution.

Proposition 5.AA.2: (*The Nonsubstitution Theorem*) Consider a productive Leontief input output model with $L - 1$ producible goods and $M_\ell \geq 1$ elementary activities for the producible good $\ell = 1, \ldots, L - 1$. Then there exist $L - 1$ activities $(a_1, \ldots, a_{L-1})$, with a_ℓ possibly a nonnegative linear combination of the M_ℓ elementary activities for producing good ℓ, such that *all* efficient production vectors with $y_{-L} \gg 0$ can be generated with these $L - 1$ activities.

Proof: Let $y \in Y$ be an efficient production vector with $y_{-L} \gg 0$. As a general matter, the vector y must be generated by a collection of $L - 1$ activities $(a_1, \ldots, a_{L-1})$ (some of these may be "mixtures" of the original activities) run at activity levels $(\alpha_1, \ldots, \alpha_{L-1}) \gg 0$; that is, $y = \sum_{\ell=1}^{L-1} \alpha_\ell a_\ell$. We show that any efficient production plan y' with $y'_{-L} \gg 0$ can be achieved using the activities $(a_1, \ldots, a_{L-1})$.

Since $y \in Y$ is efficient, there exists a $p \gg 0$ such that y is profit maximizing with respect to p (this is from Proposition 5.F.2, as strengthened for the linear activity model). From $p \cdot a_\ell \leq 0$ for all $\ell = 1, \ldots, L - 1$, $\alpha_\ell > 0$, and

$$0 = p \cdot y = p \cdot \left(\sum_{\ell=1}^{L-1} \alpha_\ell a_\ell \right) = \sum_{\ell=1}^{L-1} \alpha_\ell p \cdot a_\ell,$$

it follows that $p \cdot a_\ell = 0$ for all $\ell = 1, \ldots, L - 1$.

Consider now any other efficient production $y' \in Y$ with $y'_{-L} \gg 0$. We want to show that y' can be generated from the activities $(a_1, \ldots, a_{L-1})$. Denote by A the input output matrix associated with $(a_1, \ldots, a_{L-1})$. Because $y_{-L} \gg 0$, it follows by definition that A is productive. Therefore, by Proposition 5.AA.1, we know that there are activity levels $(\alpha''_1, \ldots, \alpha''_{L-1})$ such that the production vector $y'' = \sum_{\ell=1}^{L-1} \alpha''_\ell a_\ell$ has $y''_{-L} = y'_{-L}$. Note that since $p \cdot a_\ell = 0$ for all $\ell = 1, \ldots, L - 1$, we must have $p \cdot y'' = 0$. Thus, y'' is profit maximizing for $p \gg 0$ (recall that the maximum profits for p are zero), and so it follows that y'' is efficient by Proposition 5.F.1. But then we have two production vectors, y' and y'', with $y'_{-L} = y''_{-L}$, and both are efficient. It must therefore be that $y''_L = y'_L$. Hence, we conclude that y' can be produced using only the activities $(a_1, \ldots, a_{L-1})$, which is the desired result. ■

The nonsubstitution theorem depends critically on the presence of only one primary factor. This makes sense. With more than one primary factor, the optimal choice of techniques should depend on the relative prices of these factors. In turn, it is logical to expect that these relative prices will not be independent of the composition of final demand (e.g., if demand moves from land-intensive goods toward labour-intensive goods, we would expect the price of labor relative to the price of land to increase). Nonetheless, it is worth mentioning that the nonsubstitution result remains valid as long as the prices of the primary factors do not change.

For further reading on the material discussed in this appendix see Gale (1960).

REFERENCES

Champsaur, P., and J.-C. Milleron (1983). *Advanced Exercises in Microeconomics.* Cambridge, Mass.: Harvard University Press.

Debreu, G. (1959). *Theory of Value.* New York: Wiley.

Gale, D. (1960). *The Theory of Linear Economic Models.* New York: McGraw-Hill.

Koopmans, T. (1957). *Three Essays on the State of Economic Science.* Essay 1. New York: McGraw-Hill.

McFadden, D. (1978). Cost, revenue and profit functions. In *Production economics: A dual approach to theory and applications*, edited by M. Fuss and D. McFadden. Amsterdam: North-Holland.

McKenzie, L. (1959). On the existence of a general equilibrium for a competitive market. *Econometrica* **27**: 54–71.

Samuelson, P. (1951). Abstract of a theorem concerning substitutability in open Leontief models. In *Activity analysis of production and allocation*, edited by T. Koopmans. New York: Wiley.

EXERCISES

5.B.1^A In text

5.B.2^A In text.

5.B.3^A In text.

5.B.4^B Suppose that Y is a production set, interpreted now as the technology of a single production unit. Denote by Y^+ the additive closure of Y, that is, the smallest production set that is additive and contains Y (in other words, Y^+ is the total production set if technology Y can be replicated an arbitrary number of times). Represent Y^+ for each of the examples of production sets depicted graphically in Section 5.B. In particular, note that for the typical decreasing returns technology of Figure 5.B.5(a), the additive closure Y^+ violates the closedness condition (ii). Discuss and compare with the case corresponding to Figure 5.B.5(b), where Y^+ is closed.

5.B.5^C Show that if Y is closed and convex, and $-\mathbb{R}^L_+ \subset Y$, then free disposal holds.

5.B.6^B There are three goods. Goods 1 and 2 are inputs. The third, with amounts denoted by q, is an output. Output can be produced by two techniques that can be operated simultaneously or separately. The techniques are not necessarily linear. The first (respectively, the second) technique uses only the first (respectively, the second) input. Thus, the first (respectively, the second) technique is completely specified by $\phi_1(q_1)$ [respectively, $\phi_2(q_2)$], the minimal amount of input one (respectively, two) *sufficient* to produce the amount of output q_1 (respectively, q_2). The two functions $\phi_1(\cdot)$ and $\phi_2(\cdot)$ are increasing and $\phi_1(0) = \phi_2(0) = 0$.

(a) Describe the three-dimensional production set associated with these two techniques. Assume free disposal.

(b) Give sufficient conditions on $\phi_1(\cdot), \phi_2(\cdot)$ for the production set to display additivity.

(c) Suppose that the input prices are w_1 and w_2. Write the first-order necessary conditions for profit maximization and interpret. Under which conditions on $\phi_1(\cdot), \phi_2(\cdot)$ will the necessary conditions be sufficient?

(d) Show that if $\phi_1(\cdot)$ and $\phi_2(\cdot)$ are strictly concave, then a cost-minimizing plan cannot involve the simultaneous use of the two techniques. Interpret the meaning of the concavity requirement, and draw isoquants in the two-dimensional space of input uses.

5.C.1^A In text.

5.C.2^A In text.

5.C.3^B Establish properties (viii) and (ix) of Proposition 5.C.2. [*Hint*: Property (viii) is easy; (ix) is more difficult. Try the one-input case first.]

5.C.4^A Establish properties (i) to (vii) of Proposition 5.C.2 for the case in which there are multiple outputs.

5.C.5^A Argue that for property (iii) of Proposition 5.C.2 to hold, it suffices that $f(\cdot)$ be quasiconcave. Show that quasiconcavity of $f(\cdot)$ is compatible with increasing returns.

5.C.6^C Suppose $f(z)$ is a concave production function with $L-1$ inputs $(z_1, \dots, z_{L-1})$. Suppose also that $\partial f(z)/\partial z_\ell \geq 0$ for all ℓ and $z \geq 0$ and that the matrix $D^2 f(z)$ is negative definite at all z. Use the firm's first-order conditions and the implicit function theorem to prove the following statements:

(a) An increase in the output price always increases the profit-maximizing level of output.

(b) An increase in output price increases the demand for *some* input.

(c) An increase in the price of an input leads to a reduction in the demand for the input.

5.C.7^C A price-taking firm producing a single product according to the technology $q = f(z_1, \dots, z_{L-1})$ faces prices p for its output and $w_1, \dots, w_{L-1}$ for each of its inputs. Assume that $f(\cdot)$ is strictly concave and increasing, and that $\partial^2 f(z)/\partial z_\ell \, \partial z_k < 0$ for all $\ell \neq k$. Show that for all $\ell = 1, \dots, L-1$, the factor demand functions $z_\ell(p, w)$ satisfy $\partial z_\ell(p, w)/\partial p > 0$ and $\partial z_\ell(p, w)/\partial w_k < 0$ for all $k \neq \ell$.

5.C.8^B Alpha Incorporated (AI) produces a single output q from two inputs z_1 and z_2. You are assigned to determine AI's technology. You are given 100 monthly observations. Two of these monthly observations are shown in the following table:

	Input prices		*Input levels*		*Output price*	*Output level*
Month	w_1	w_2	z_1	z_2	p	q
3	3	1	40	50	4	60
95	2	2	55	40	4	60

In light of these two monthly observations, what problem will you encounter in trying to accomplish your task?

5.C.9^A Derive the profit function $\pi(p)$ and supply function (or correspondence) $y(p)$ for the single-output technologies whose production functions $f(z)$ are given by

(a) $f(z) = \sqrt{z_1} + z_2$.

(b) $f(z) = \sqrt{\text{Min}\,\{z_1, z_2\}}$.

(c) $f(z) = (z_1^\rho + z_2^\rho)^{1/\rho}$, for $\rho \leq 1$.

5.C.10[A] Derive the cost function $c(w, q)$ and conditional factor demand functions (or correspondences) $z(w, q)$ for each of the following single-output constant return technologies with production functions given by

(a) $f(z) = z_1 + z_2$ (perfect substitutable inputs)

(b) $f(z) = \text{Min}\,\{z_1, z_2\}$ (Leontief technology)

(c) $f(z) = (z_1^\rho + z_2^\rho)^{1/\rho}$, $\rho \leq 1$ (constant elasticity of substitution technology)

5.C.11[A] Show that $\partial z_\ell(w, q)/\partial q > 0$ if and only if marginal cost at q is increasing in w_ℓ.

5.C.12[A] We saw at the end of Section 5.B that any convex Y can be viewed as the section of a constant returns technology $Y' \subset \mathbb{R}^{L+1}$, where the $L + 1$ coordinate is fixed at the level -1. Show that if $y \in Y$ is profit maximizing at prices p then $(y, -1) \in Y'$ is profit maximizing at $(p, \pi(p))$, that is, profits emerge as the price of the implicit fixed input. The converse is also true: If $(y, -1) \in Y'$ is profit maximizing at prices (p, p_{L+1}), then $y \in Y$ is profit maximizing at p and the profit is p_{L+1}.

5.C.13[B] A price-taking firm produces output q from inputs z_1 and z_2 according to a differentiable concave production function $f(z_1, z_2)$. The price of its output is $p > 0$, and the prices of its inputs are $(w_1, w_2) \gg 0$. However, there are two unusual things about this firm. First, rather than maximizing profit, the firm maximizes revenue (the manager wants her firm to have bigger dollar sales than any other). Second, the firm is cash constrained. In particular, it has only C dollars on hand before production and, as a result, its total expenditures on inputs cannot exceed C.

Suppose one of your econometrician friends tells you that she has used repeated observations of the firm's revenues under various output prices, input prices, and levels of the financial constraint and has determined that the firm's revenue level R can be expressed as the following function of the variables (p, w_1, w_2, C):

$$R(p, w_1, w_2, C) = p[\gamma + \ln C - \alpha \ln w_1 - (1 - \alpha) \ln w_2].$$

(γ and α are scalars whose values she tells you.) What is the firm's use of input z_1 when prices are (p, w_1, w_2) and it has C dollars of cash on hand?

5.D.1[A] In text.

5.D.2[A] In text.

5.D.3[B] Suppose that a firm can produce good L from $L - 1$ factor inputs ($L > 2$). Factor prices are $w \in \mathbb{R}^{L-1}$ and the price of output is p. The firm's differentiable cost function is $c(w, q)$. Assume that this function is strictly convex in q. However, although $c(w, q)$ is the cost function when all factors can be freely adjusted, factor 1 cannot be adjusted in the short run.

Suppose that the firm is initially at a point where it is producing its long-run profit-maximizing output level of good L given prices w and p, $q(w, p)$ [i.e., the level that is optimal under the long-run cost conditions described by $c(w, q)$], and that all inputs are optimally adjusted [i.e., $z_\ell = z_\ell(w, q(w, p))$ for all $\ell = 1, \ldots, L - 1$, where $z_\ell(\cdot,\cdot)$ is the long-run input demand function]. Show that the firm's profit-maximizing output response to a marginal increase in the price of good L is larger in the long run than in the short run. [*Hint*: Define a short-run cost function $c_s(w, q \mid z_1)$ that gives the minimized costs of producing output level q given that input 1 is fixed at level z_1.]

5.D.4[B] Consider a firm that has a distinct set of inputs and outputs. The firm produces M outputs; let $q = (q_1, \dots, q_M)$ denote a vector of its output levels. Holding factor prices fixed, $C(q_1, \dots, q_M)$ is the firm's cost function. We say that $C(\cdot)$ is *subadditive* if for all $(q_1, \dots, q_M)$, there is no way to break up the production of amounts $(q_1, \dots, q_M)$ among several firms, each with cost function $C(\cdot)$, and lower the costs of production. That is, there is no set of, say, J firms and collection of production vectors $\{q_j = (q_{1j}, \dots, q_{Mj})\}_{j=1}^{J}$ such that $\sum_j q_j = q$ and $\sum_j C(q_j) < C(q)$. When $C(\cdot)$ is subadditive, it is usual to say that the industry is a *natural monopoly* because production is cheapest when it is done by only one firm.

(a) Consider the single-output case, $M = 1$. Show that if $C(\cdot)$ exhibits decreasing average costs, then $C(\cdot)$ is subadditive.

(b) Now consider the multiple-output case, $M > 1$. Show by example that the following multiple-output extension of the decreasing average cost assumption is *not* sufficient for $C(\cdot)$ to be subadditive:

$C(\cdot)$ exhibits *decreasing ray average cost* if for any $q \in \mathbb{R}^M_+$,

$$C(q) > C(kq)/k \text{ for all } k > 1.$$

(c) (Harder) Prove that, if $C(\cdot)$ exhibits decreasing ray average cost *and* is quasiconvex, then $C(\cdot)$ is subadditive. [Assume that $C(\cdot)$ is continuous, increasing, and satisfies $C(0) = 0$.]

5.D.5[B] Suppose there are two goods: an input z and an output q. The production function is $q = f(z)$. We assume that $f(\cdot)$ exhibits increasing returns to scale.

(a) Assume that $f(\cdot)$ is differentiable. Do the increasing returns of $f(\cdot)$ imply that the average product is necessarily nondecreasing in input? What about the marginal product?

(b) Suppose there is a representative consumer with the utility function $u(q) - z$ (the negative sign indicates that the input is taken away from the consumer). Suppose that $\bar{q} = f(\bar{z})$ is a production plan that maximizes the representative consumer utility. Argue, either mathematically or economically (disregard boundary solutions), that the equality of marginal utility and marginal cost is a necessary condition for this maximization problem.

(c) Assume the existence of a representative consumer as in (b). "The equality of marginal cost and marginal utility is a sufficient condition for the optimality of a production plan." Right or wrong? Discuss.

5.E.1[A] Assuming that every $\pi_j(\cdot)$ is differentiable and that you already know that $\pi^*(p) = \sum_{j=1}^{J} \pi_j(p)$, give a proof of $y^*(p) = \sum_{j=1}^{J} y_j(p)$ using differentiability techniques.

5.E.2[A] Verify that Proposition 5.E.1 and its interpretation do not depend on any convexity hypothesis on the sets $Y_1, \dots, Y_J$.

5.E.3[B] Assuming that the sets $Y_1, \dots, Y_J$ are convex and satisfy the free disposal property, and that $\sum_{j=1}^{J} Y_j$ is closed, show that the latter set equals $\{y\colon p \cdot y \le \sum_{j=1}^{J} \pi_j(p) \text{ for all } p \gg 0\}$.

5.E.4[B] One output is produced from two inputs. There are many technologies. Every technology can produce up to one unit of output (but no more) with fixed and proportional input requirements z_1 and z_2. So a technology is characterized by $z = (z_1, z_2)$, and we can describe the population of technologies by a density function $g(z_1, z_2)$. Take this density to be uniform on the square $[0, 10] \times [0, 10]$.

(a) Given the input prices $w = (w_1, w_2)$, solve the profit maximization problem of a firm with characteristics z. The output price is 1.

(b) More generally, find the profit function $\pi(w_1, w_2, 1)$ for

$$w_1 \geq \frac{1}{10} \quad \text{and} \quad w_2 \geq \frac{1}{10}.$$

(c) Compute the aggregate input demand function. Ideally, do that directly, and check that the answer is correct by using your finding in (b); this way you also verify (b).

(d) What can you say about the aggregate production function? If you were to assume that the profit function derived in (b) is valid for $w_1 \geq 0$ and $w_2 \geq 0$, what would the underlying aggregate production function be?

5.E.5^A (M. Weitzman) Suppose that there are J single-output plants. Plant j's average cost is $AC_j(q_j) = \alpha + \beta_j q_j$ for $q_j \geq 0$. Note that the coefficient α is the same for all plants but that the coefficient β_j may differ from plant to plant. Consider the problem of determining the cost-minimizing aggregate production plan for producing a total output of q, where $q < (\alpha/\text{Max}_j |\beta_j|)$.

(a) If $\beta_j > 0$ for all j, how should output be allocated among the J plants?

(b) If $\beta_j < 0$ for all j, how should output be allocated among the J plants?

(c) What if $\beta_j > 0$ for some plants and $\beta_j < 0$ for others?

5.F.1^A In text.

5.G.1^B Let $f(z)$ be a single-input, single-output production function. Suppose that owners have quasilinear utilities with the firm's input as the numeraire.

(a) Show that a necessary condition for consumer–owners to unanimously agree to a production plan z is that consumption shares among owners at prices $p(z)$ coincide with ownership shares.

(b) Suppose that ownership shares are identical. Comment on the conflicting instructions to managers and how they depend on the consumer–owners' tastes for output.

(c) With identical preferences and ownership shares, argue that owners will unanimously agree to maximize profits in terms of input. (Recall that we are assuming preferences are quasilinear with respect to input; hence, the numeraire is intrinsically determined.)

5.AA.1^A Compute the cost function $c(w, 1)$ and the input demand $z(w, 1)$ for the production function in Figure 5.AA.4. Verify that whenever $z(w, 1)$ is single-valued, we have $z(w, 1) = \nabla_w c(w, 1)$.

5.AA.2^B Consider a Leontief input–output model with no substitution. Assume that the input matrix A is productive and that the vector of primary factor requirements b is strictly positive.

(a) Show that for any $\alpha \geq 0$, the production plan

$$y = \begin{bmatrix} I - A \\ -b \end{bmatrix} \alpha.$$

is efficient.

(b) Fixing the price of the primary factor to equal 1, show that any production plan with $\alpha \gg 0$ is profit maximizing at a unique vector of prices.

(c) Show that the prices obtained in (b) have the interpretation of amounts of the primary factor directly or indirectly embodied in the production of one unit of the different goods.

(d) (Harder) Suppose that A corresponds to the techniques singled out by the nonsubstitution theorem for a model that, in principle, admits substitution. Show that every component of the price vector obtained from A in (c) is less than or equal to the corresponding component of the price vector obtained from any other selection of techniques.

5.AA.3[B] There are two produced goods and labor. The input–output matrix is

$$A = \begin{bmatrix} 0 & 1 \\ \alpha & 0 \end{bmatrix}.$$

Here $a_{\ell k}$ is the amount of good ℓ required to produce one unit of good k.

(a) Let $\alpha = \frac{1}{2}$, and suppose that the labor coefficients vector is

$$b = \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

where b_1 (respectively, b_2) is the amount of labor required to produce one unit of good 1 (respectively, good 2). Represent graphically the production possibility set (i.e., the locus of possible productions) for the two goods if the total availability of labor is 10.

(b) For the values of α and b in (a), compute equilibrium prices p_1, p_2 (normalize the wage to equal 1) from the profit maximization conditions (assume positive production of the two goods).

(c) For the values of α and b in (a), compute the amount of labor directly or indirectly incorporated into the production of one net (i.e., available for consumption) unit of good 1. How does this amount relate to your answer in **(b)**?

(d) Suppose there is a second technique to produce good 2. To

$$\begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad b_2 = 2$$

we now add

$$\begin{bmatrix} a'_{12} \\ a'_{22} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix}, \qquad b'_2 = \beta.$$

Taking the two techniques into account, represent graphically the locus of amounts of good 1 and of labor necessary to produce one unit of good 2. (Assume free disposal.)

(e) In the context of **(d)**, what does the nonsubstitution theorem say? Determine the value of β at which there is a switch of optimal techniques.

5.AA.4[B] Consider the following linear activity model:

$$\begin{aligned} a_1 &= (1, -1, 0, 0) \\ a_2 &= (0, -1, 1, 0) \\ a_3 &= (0, 0, -1, 1) \\ a_4 &= (2, 0, 0, -1) \end{aligned}$$

(a) For each of the following input–output vectors, check whether they belong or do not belong to the aggregate production set. Justify your answers:

$$\begin{aligned} y_1 &= (6, 0, 0, -2) \\ y_2 &= (5, -3, 0, -1) \\ y_3 &= (6, -3, 0, 0) \\ y_4 &= (0, -4, 0, 4) \\ y_5 &= (0, -3, 4, 0) \end{aligned}$$

(b) The input–output vector $y = (0, -5, 5, 0)$ is efficient. Prove this by finding a $p \gg 0$ for which y is profit-maximizing.

(c) The input–output vector $y = (1, -1, 0, 0)$ is feasible, but it is not efficient. Why?

5.AA.5[B] [This exercise was inspired by an exercise of Champsaur and Milleron (1983).] There are four commodities indexed by $\ell = 1, 2, 3, 4$. The technology of a firm is described by eight elementary activities a_m, $m = 1, \ldots, 8$. With the usual sign convention, the numerical values of these activities are

$$a_1 = (\ -3,\ \ -6,\ \ 4,\ \ 0)$$
$$a_2 = (\ -7,\ \ -9,\ \ 3,\ \ 2)$$
$$a_3 = (\ -1,\ \ -2,\ \ 3,\ -1)$$
$$a_4 = (\ -8,\ -13,\ \ 3,\ \ 1)$$
$$a_5 = (-11,\ -19,\ 12,\ \ 0)$$
$$a_6 = (\ -4,\ \ -3,\ -2,\ \ 5)$$
$$a_7 = (\ -8,\ \ -5,\ \ 0,\ 10)$$
$$a_8 = (\ -2,\ \ -4,\ \ 5,\ \ 2)$$

It is assumed that any activity can be operated at any nonnegative level $\alpha_m \geq 0$ and that all activities can operate simultaneously at any scale (i.e., for any $\alpha_m \geq 0$, $m = 1, \ldots, 8$, the production $\sum_m \alpha_m a_m$ is feasible).

(a) Define the corresponding production set Y, and show that it is convex.

(b) Verify the no-free-lunch property.

(c) Verify that Y does *not* satisfy the free-disposal property. The free-disposal property would be satisfied if we added new elementary activities to our list. How would you choose them (given specific numerical values)?

(d) Show by direct comparison of a_1 with a_5, a_2 with a_4, a_3 with a_8, and a_6 with a_7 that four of the elementary activities are not efficient.

(e) Show that a_1 and a_2 are inefficient be exhibiting two positive linear combinations of a_3 and a_7 that dominate a_1 and a_2, respectively.

(f) Could you venture a complete description of the set of efficient production vectors?

(g) Suppose that the amounts of the four goods available as initial resources to the firm are

$$s_1 = 480, \qquad s_2 = 300, \qquad s_3 = 0, \qquad s_4 = 0.$$

Subject to those limitations on the net use of resources, the firm is interested in maximizing the net production of the third good. How would you set up the problem as a linear program?

(h) By using all the insights you have gained on the set of efficient production vectors, can you solve the optimization problem in **(g)**? [*Hint*: It can be done graphically.]