Production

6

In the last three chapters, we focused on the *demand side* of the market—the preferences and behavior of consumers. Now we turn to the *supply side* and examine the behavior of producers. We will see how firms can produce efficiently and how their costs of production change with changes in both input prices and the level of output. We will also see that there are strong similarities between the optimizing decisions made by firms and those made by consumers. In other words, understanding consumer behavior will help us understand producer behavior.

In this chapter and the next we discuss the **theory of the firm**, which describes how a firm makes cost-minimizing production decisions and how the firm's resulting cost varies with its output. Our knowledge of production and cost will help us understand the characteristics of market supply. It will also prove useful for dealing with problems that arise regularly in business. To see this, just consider some of the problems often faced by a company like General Motors. How much assembly-line machinery and how much labor should it use in its new automobile plants? If it wants to increase production, should it hire more workers, construct new plants, or both? Does it make more sense for one automobile plant to produce different models, or should each model be manufactured in a separate plant? What should GM expect its costs to be during the coming year? How are these costs likely to change over time and be affected by the level of production? These questions apply not only to business firms but also to other producers of goods and services, such as governments and nonprofit agencies.

CHAPTER OUTLINE

LIST OF EXAMPLES

The Production Decisions of a Firm

In Chapters 3 and 4, we studied consumer behavior by breaking it down into three steps. First, we explained how to describe consumer preferences. Second, we accounted for the fact that consumers face budget constraints. Third, we saw how, given their preferences and budget constraints, consumers can choose combinations of goods to maximize their satisfaction. The production decisions of firms are analogous to the purchasing decisions of consumers, and can likewise be understood in three steps:

1. **Production Technology:** We need a practical way of describing how *inputs* (such as labor, capital, and raw materials) can be transformed into *outputs* (such as cars and televisions). Just as a consumer can reach a level of satisfaction from buying different combinations of goods, the firm can produce a particular level of output by using different combinations of inputs. For example, an electronics firm

might produce 10,000 televisions per month by using a substantial amount of labor (e.g., workers assembling the televisions by hand) and very little capital, or by building a highly automated capital-intensive factory and using very little labor.

2. **Cost Constraints:** Firms must take into account the *prices* of labor, capital, and other inputs. Just as a consumer is constrained by a limited budget, the firm will be concerned about its cost of production. For example, the firm that produces 10,000 televisions per month will want to do so in a way that minimizes its total production cost, which is determined in part by the prices of the inputs it uses.
3. **Input Choices:** Given its production technology and the prices of labor, capital, and other inputs, the firm must choose *how much of each input* to use in producing its output. Just as a consumer takes account of the prices of different goods when deciding how much of each good to buy, the firm must take into account the prices of different inputs when deciding how much of each input to use. If our electronics firm operates in a country with low wage rates, it may decide to produce televisions by using a large amount of labor, thereby using very little capital.

• **theory of the firm** Explanation of how a firm makes cost-minimizing production decisions and how its cost varies with its output.

These three steps are the building blocks of the theory of the firm, and we will discuss them in detail in this chapter and the next. We will also address other important aspects of firm behavior. For example, assuming that the firm is always using a cost-minimizing combination of inputs, we will see how its total cost of production varies with the quantity it produces and how it can choose that quantity to maximize its profit.

We begin this chapter by showing how the firm's production technology can be represented in the form of a *production function*—a compact description of how inputs are turned into output. We then use the production function to show how the firm's output changes when just one of its inputs (labor) is varied, holding the other inputs fixed. Next, we turn to the more general case in which the firm can vary all of its inputs, and we show how the firm chooses a cost-minimizing combination of inputs to produce its output. We will be particularly concerned with the *scale* of the firm's operation. Are there, for example, any technological advantages that make the firm more productive as its scale increases?

6.1 THE TECHNOLOGY OF PRODUCTION

• **factors of production** Inputs into the production process (e.g., labor, capital, and materials).

In the production process, firms turn *inputs* into *outputs* (or products). Inputs, which are also called **factors of production**, include anything that the firm must use as part of the production process. In a bakery, for example, inputs include the labor of its workers; raw materials, such as flour and sugar; and the capital invested in its ovens, mixers, and other equipment needed to produce such outputs as bread, cakes, and pastries.

As you can see, we can divide inputs into the broad categories of *labor*, *materials*, and *capital*, each of which might include more narrow subdivisions. Labor inputs include skilled workers (carpenters, engineers) and unskilled workers (agricultural workers), as well as the entrepreneurial efforts of the firm's managers. Materials include steel, plastics, electricity, water, and any other goods that the firm buys and transforms into final products. Capital includes land, buildings, machinery and other equipment, as well as inventories.

The Production Function

Firms can turn inputs into outputs in a variety of ways, using various combinations of labor, materials, and capital. We can describe the relationship between the inputs into the production process and the resulting output by a *production function*. A **production function** indicates the highest output q that a firm can produce for every specified combination of inputs.[1] Although in practice firms use a wide variety of inputs, we will keep our analysis simple by focusing on only two, labor L and capital K. We can then write the production function as

• **production function** Function showing the highest output that a firm can produce for every specified combination of inputs.

$$q = F(K, L) \tag{6.1}$$

This equation relates the quantity of output to the quantities of the two inputs, capital and labor. For example, the production function might describe the number of personal computers that can be produced each year with a 10,000-square-foot plant and a specific amount of assembly-line labor. Or it might describe the crop that a farmer can obtain using specific amounts of machinery and workers.

It is important to keep in mind that inputs and outputs are *flows*. For example, our PC manufacturer uses a certain amount of labor *each year* to produce some number of computers over that year. Although it might own its plant and machinery, we can think of the firm as paying a cost for the use of that plant and machinery over the year. To simplify things, we will frequently ignore the reference to time and refer only to amounts of labor, capital, and output. Unless otherwise indicated, however, we mean the amount of labor and capital used each year and the amount of output produced each year.

Because the production function allows inputs to be combined in varying proportions, output can be produced in many ways. For the production function in equation (6.1), this could mean using more capital and less labor, or vice versa. For example, wine can be produced in a labor-intensive way using many workers, or in a capital-intensive way using machines and only a few workers.

Note that equation (6.1) applies to a *given technology*—that is, to a given state of knowledge about the various methods that might be used to transform inputs into outputs. As the technology becomes more advanced and the production function changes, a firm can obtain more output for a given set of inputs. For example, a new, faster assembly line may allow a hardware manufacturer to produce more high-speed computers in a given period of time.

Production functions describe what is *technically feasible* when the firm operates *efficiently*—that is, when the firm uses each combination of inputs as effectively as possible. The presumption that production is always technically efficient need not always hold, but it is reasonable to expect that profit-seeking firms will not waste resources.

The Short Run versus the Long Run

It takes time for a firm to adjust its inputs to produce its product with differing amounts of labor and capital. A new factory must be planned and built, and machinery and other capital equipment must be ordered and delivered. Such activities can easily take a year or more to complete. As a result, if we

[1] In this chapter and those that follow, we will use the variable q for the output of the firm, and Q for the output of the industry.

are looking at production decisions over a short period of time, such as a month or two, the firm is unlikely to be able to substitute very much capital for labor.

Because firms must consider whether or not inputs can be varied, and if they can, over what period of time, it is important to distinguish between the short and long run when analyzing production. The **short run** refers to a period of time in which the quantities of one or more factors of production cannot be changed. In other words, in the short run there is at least one factor that cannot be varied; such a factor is called a **fixed input**. The **long run** is the amount of time needed to make *all* inputs variable.

• **short run** Period of time in which quantities of one or more production factors cannot be changed.

• **fixed input** Production factor that cannot be varied.

• **long run** Amount of time needed to make all production inputs variable.

As you might expect, the kinds of decisions that firms can make are very different in the short run than *those made* in the long run. In the short run, firms vary the intensity with which they utilize a given plant and machinery; in the long run, they vary the size of the plant. All fixed inputs in the short run represent the outcomes of previous long-run decisions based on estimates of what a firm could profitably produce and sell.

There is no specific time period, such as one year, that separates the short run from the long run. Rather, one must distinguish them on a case-by-case basis. For example, the long run can be as brief as a day or two for a child's lemonade stand or as long as five or ten years for a petrochemical producer or an automobile manufacturer.

We will see that in the long run firms can vary the amounts of all their inputs to minimize the cost of production. Before treating this general case, however, we begin with an analysis of the short run, in which only one input to the production process can be varied. We assume that capital is the fixed input, and labor is variable.

6.2 PRODUCTION WITH ONE VARIABLE INPUT (LABOR)

When deciding how much of a particular input to buy, a firm has to compare the benefit that will result with the cost. Sometimes it is useful to look at the benefit and the cost on an *incremental* basis by focusing on the additional output that results from an incremental addition to an input. In other situations, it is useful to make the comparison on an *average* basis by considering the result of substantially increasing an input. We will look at benefits and costs in both ways.

When capital is fixed but labor is variable, the only way the firm can produce more output is by increasing its labor input. Imagine, for example, that you are managing a clothing factory. Although you have a fixed amount of equipment, you can hire more or less labor to sew and to run the machines. You must decide how much labor to hire and how much clothing to produce. To make the decision, you will need to know how the amount of output q increases (if at all) as the input of labor L increases.

Table 6.1 gives this information. The first three columns show the amount of output that can be produced in one month with different amounts of labor and capital fixed at 10 units. The first column shows the amount of labor, the second the fixed amount of capital, and the third total output. When labor input is zero, output is also zero. Output then increases as labor is increased up to an input of 8 units. Beyond that point, total output declines: Although initially each unit of labor can take greater and greater advantage of the existing machinery and plant, after a certain point, additional labor is no longer useful and indeed can

TABLE 6.1 Production with One Variable Input

Amount of Labor (L)	Amount of Capital (K)	Total Output (q)	Average Product (q/L)	Marginal Product ($\Delta q/\Delta L$)
0	10	0	—	—
1	10	10	10	10
2	10	30	15	20
3	10	60	20	30
4	10	80	20	20
5	10	95	19	15
6	10	108	18	13
7	10	112	16	4
8	10	112	14	0
9	10	108	12	–4
10	10	100	10	–8

be counterproductive. Five people can run an assembly line better than two, but ten people may get in one another's way.

Average and Marginal Products

The contribution that labor makes to the production process can be described on both an *average* and a *marginal* (i.e., incremental) basis. The fourth column in Table 6.1 shows the **average product** of labor (AP_L), which is the output per unit of labor input. The average product is calculated by dividing the total output q by the total input of labor L. The average product of labor measures the productivity of the firm's workforce in terms of how much output each worker produces on average. In our example, the average product increases initially but falls when the labor input becomes greater than four.

• **average product** Output per unit of a particular input.

The fifth column of Table 6.1 shows the **marginal product** of labor (MP_L). This is the *additional* output produced as the labor input is increased by 1 unit. For example, with capital fixed at 10 units, when the labor input increases from 2 to 3, total output increases from 30 to 60, creating an additional output of 30 (i.e., 60 – 30) units. The marginal product of labor can be written as $\Delta q/\Delta L$—in other words, the change in output Δq resulting from a 1-unit increase in labor input ΔL.

• **marginal product** Additional output produced as an input is increased by one unit.

Remember that the marginal product of labor depends on the amount of capital used. If the capital input increased from 10 to 20, the marginal product of labor most likely would increase. Why? Because additional workers are likely to be more productive if they have more capital to use. Like the average product, the marginal product first increases then falls—in this case, after the third unit of labor.

To summarize:

Average product of labor = Output/labor input = q/L
Marginal product of labor = Change in output/change in labor input
$= \Delta q/\Delta L$

The Slopes of the Product Curve

Figure 6.1 plots the information contained in Table 6.1. (We have connected all the points in the figure with solid lines.) Figure 6.1(a) shows that as labor is increased, output increases until it reaches the maximum output of 112; thereafter, it falls. The portion of the total output curve that is declining is drawn with a dashed line to denote that producing with more than eight workers is not

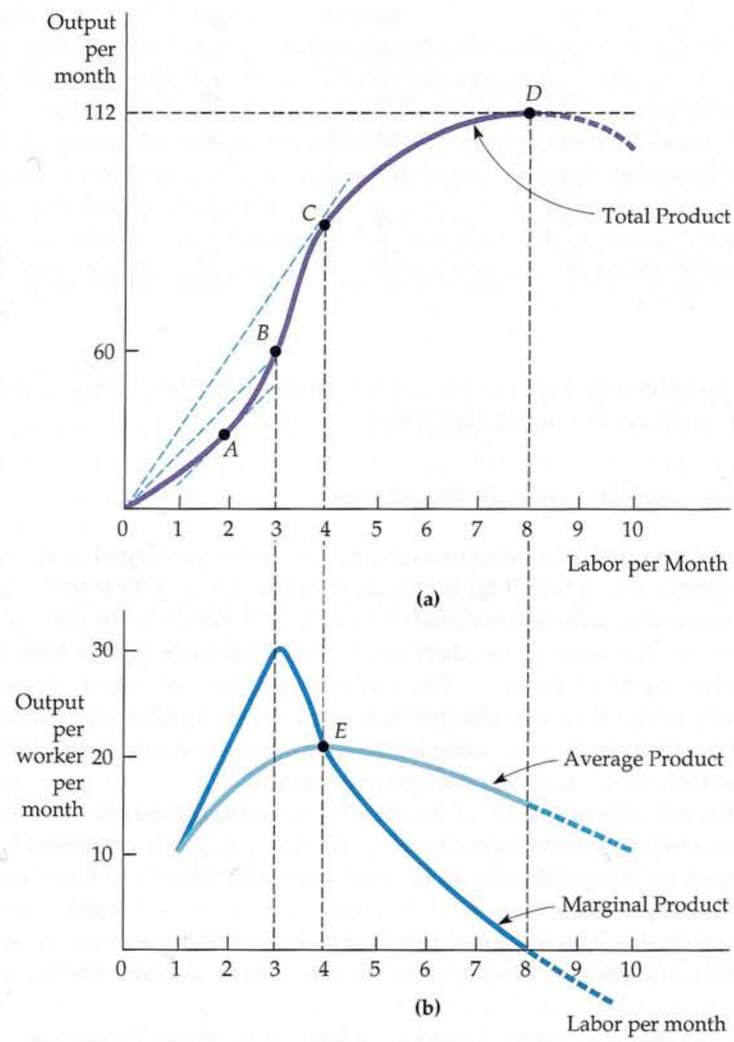


FIGURE 6.1 Production with One Variable Input

The total product curve in **(a)** shows the output produced for different amounts of labor input. The average and marginal products in **(b)** can be obtained (using the data in Table 6.1) from the total product curve. At point *A* in **(a)**, the marginal product is 20 because the tangent to the total product curve has a slope of 20. At point *B* in **(a)** the average product of labor is 20, which is the slope of the line from the origin to *B*. The average product of labor at point *C* in **(a)** is given by the slope of the line 0*C*. To the left of point *E* in **(b)**, the marginal product is above the average product and the average is increasing; to the right of *E*, the marginal product is below the average product and the average is decreasing. As a result, *E* represents the point at which the average and marginal products are equal, when the average product reaches its maximum.

economically rational; it can never be profitable to use additional amounts of a costly input to produce *less* output.

Figure 6.1(b) shows the average and marginal product curves. (The units on the vertical axis have changed from output per month to output per worker per month.) Note that the marginal product is positive as long as output is increasing, but becomes negative when output is decreasing.

It is no coincidence that the marginal product curve crosses the horizontal axis of the graph at the point of maximum total product. This happens because adding a worker in a manner that slows production and decreases total output implies a negative marginal product for that worker.

The average product and marginal product curves are closely related. *When the marginal product is greater than the average product, the average product is increasing*. This is the case for labor inputs up to 4 in Figure 6.1(b). If the output of an additional worker is greater than the average output of each existing worker (i.e., the marginal product is greater than the average product), then adding the worker causes average output to rise. In Table 6.1, two workers produce 30 units of output, for an average product of 15 units per worker. Adding a third worker increases output by 30 units (to 60), which raises the average product from 15 to 20.

Similarly, *when the marginal product is less than the average product, the average product is decreasing*. This is the case when the labor input is greater than 4 in Figure 6.1(b). In Table 6.1, six workers produce 108 units of output, for an average product of 18. Adding a seventh worker contributes a marginal product of only 4 units (less than the average product), reducing the average product to 16.

We have seen that the marginal product is above the average product when the average product is increasing and below the average product when the average product is decreasing. It follows, therefore, that the marginal product must equal the average product when the average product reaches its maximum. This happens at point *E* in Figure 6.1(b).

Why, in practice, should we expect the marginal product curve to rise and then fall? Think of a television assembly plant. Fewer than ten workers might be insufficient to operate the assembly line at all. Ten to fifteen workers might be able to run the assembly line, but not very efficiently. If adding a few more workers allowed the assembly line to operate much more efficiently, the marginal product of those workers would be very high. This added efficiency, however, might start to diminish once there were more than 20 workers. The marginal product of the twenty-second worker, for example, might still be very high (and above the average product), but not as high as the marginal product of the nineteenth or twentieth worker. The marginal product of the twenty-fifth worker might be lower still, perhaps equal to the average product. With 30 workers, adding one more worker would yield more output, but not very much more (so that the marginal product, while positive, would be below the average product). Once there were more than 40 workers, additional workers would simply get in each other's way and actually reduce output (so that the marginal product would be negative).

The Average Product of Labor Curve

The geometric relationship between the total product and the average and marginal product curves is shown in Figure 6.1(a). The average product of labor is the total product divided by the quantity of labor input. At *B*, for example, the average product is equal to the output of 60 divided by the input of 3, or 20 units of output per unit of labor input. This ratio, however, is exactly the slope of the line running from the origin to B in Figure 6.1(a). In general, *the average product of labor is given by the slope of the line drawn from the origin to the corresponding point on the total product curve.*

The Marginal Product of Labor Curve

As we have seen, the marginal product of labor is the change in the total product resulting from an increase of one unit of labor. At A, for example, the marginal product is 20 because the tangent to the total product curve has a slope of 20. In general, *the marginal product of labor at a point is given by the slope of the total product at that point*. We can see in Figure 6.1(b) that the marginal product of labor increases initially, peaks at an input of 3, and then declines as we move up the total product curve to C and D. At D, when total output is maximized, the slope of the tangent to the total product curve is 0, as is the marginal product. Beyond that point, the marginal product becomes negative.

The Relationship between the Average and Marginal Products Note the graphical relationship between average and marginal products in Figure 6.1(a). At B, the marginal product of labor (the slope of the tangent to the total product curve at B—not shown explicitly) is greater than the average product (dashed line $0B$). As a result, the average product of labor increases as we move from B to C. At C, the average and marginal products of labor are equal: While the average product is the slope of the line from the origin, $0C$, the marginal product is the tangent to the total product curve at C (note the equality of the average and marginal products at point E in Figure 6.1(b)). Finally, as we move beyond C toward D, the marginal product falls below the average product; you can check that the slope of the tangent to the total product curve at any point between C and D is lower than the slope of the line from the origin.

The Law of Diminishing Marginal Returns

• **law of diminishing marginal returns** Principle that as the use of an input increases with other inputs fixed, the resulting additions to output will eventually decrease.

A diminishing marginal product of labor (as well as a diminishing marginal product of other inputs) holds for most production processes. The **law of diminishing marginal returns** states that as the use of an input increases in equal increments (with other inputs fixed), a point will eventually be reached at which the resulting additions to output decrease. When the labor input is small (and capital is fixed), extra labor adds considerably to output, often because workers are allowed to devote themselves to specialized tasks. Eventually, however, the law of diminishing marginal returns applies: When there are too many workers, some workers become ineffective and the marginal product of labor falls.

The law of diminishing marginal returns usually applies to the short run when at least one input is fixed. However, it can also apply to the long run. Even though inputs are variable in the long run, a manager may still want to analyze production choices for which one or more inputs are unchanged. Suppose, for example, that only two plant sizes are feasible and that management must decide which to build. In that case, management would want to know when diminishing marginal returns will set in for each of the two options.

Do not confuse the law of diminishing marginal returns with possible changes in the *quality* of labor as labor inputs are increased (as would likely occur, for example, if the most highly qualified laborers are hired first and the least qualified last). In our analysis of production, we have assumed that all labor inputs are of equal quality; diminishing marginal returns results from limitations on the use of other fixed inputs (e.g., machinery), not from declines in worker quality. In addition, do not confuse diminishing marginal returns with *negative* returns. The law of diminishing marginal returns describes a *declining* marginal product but not necessarily a negative one.

The law of diminishing marginal returns applies to a given production technology. Over time, however, inventions and other improvements in technology may allow the entire total product curve in Figure 6.1(a) to shift upward, so that more output can be produced with the same inputs. Figure 6.2 illustrates this principle. Initially the output curve is given by O_1, but improvements in technology may allow the curve to shift upward, first to O_2, and later to O_3.

Suppose, for example, that over time, as labor is increased in agricultural production, technological improvements are being made. These improvements might include genetically engineered pest-resistant seeds, more powerful and effective fertilizers, and better farm equipment. As a result, output changes from A (with an input of 6 on curve O_1) to B (with an input of 7 on curve O_2) to C (with an input of 8 on curve O_3).

The move from A to B to C relates an increase in labor input to an increase in output and makes it appear that there are no diminishing marginal returns when in fact there are. Indeed, the shifting of the total product curve suggests that there may be no negative long-run implications for economic growth. In fact, as we can see in Example 6.1, the failure to account for long-run improvements in technology led British economist Thomas Malthus wrongly to predict dire consequences from continued population growth.

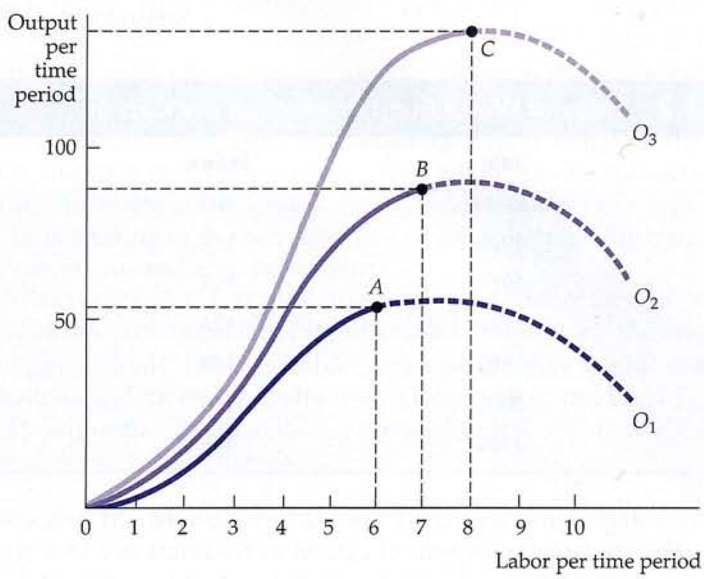


FIGURE 6.2 The Effect of Technological Improvement

Labor productivity (output per unit of labor) can increase if there are improvements in technology, even though any given production process exhibits diminishing returns to labor. As we move from point A on curve O_1 to B on curve O_2 to C on curve O_3 over time, labor productivity increases.

EXAMPLE 6.1 Malthus and the Food Crisis

The law of diminishing marginal returns was central to the thinking of political economist Thomas Malthus (1766–1834).[2] Malthus believed that the world's limited amount of land would not be able to supply enough food as the population grew. He predicted that as both the marginal and average productivity of labor fell and there were more mouths to feed, mass hunger and starvation would result. Fortunately, Malthus was wrong (although he was right about the diminishing marginal returns to labor).

Over the past century, technological improvements have dramatically altered food production in most countries (including developing countries, such as India). As a result, the average product of labor and total food output have increased. These improvements include new high-yielding, disease-resistant strains of seeds, better fertilizers, and better harvesting equipment. As the food production index in Table 6.2 shows, overall food production throughout the world has outpaced population growth continually since 1960.[3] This increase in world agricultural productivity is also illustrated in Figure 6.3, which shows average cereal yields from 1970 through 2005, along with a world price index for food.[4] Note that cereal yields have increased steadily over the period. Because growth in agricultural productivity led to increases in food supplies that outstripped the growth in demand, prices, apart from a temporary increase in the early 1970s, have been declining.

Hunger remains a severe problem in some areas, such as the Sahel region of Africa, in part because of the low productivity of labor there. Although other countries produce an agricultural surplus, mass hunger still occurs because of the difficulty of redistributing food from more to less productive regions of the world and because of the low incomes of those less productive regions.

TABLE 6.2 Index of World Food Production per Capita

Year	Index
1948–1952	100
1960	115
1970	123
1980	128
1990	138
2000	150
2005	156

[2]Thomas Malthus, *Essay on the Principle of Population*, 1798.

[3]World per capita food production data are from the United Nations Food and Agriculture Organization (FAO). See also **http://faostat.fao.org**.

[4]Data are from the United Nations Food and Agriculture Organization and the World Bank. See also **http://faostat.fao.org**.

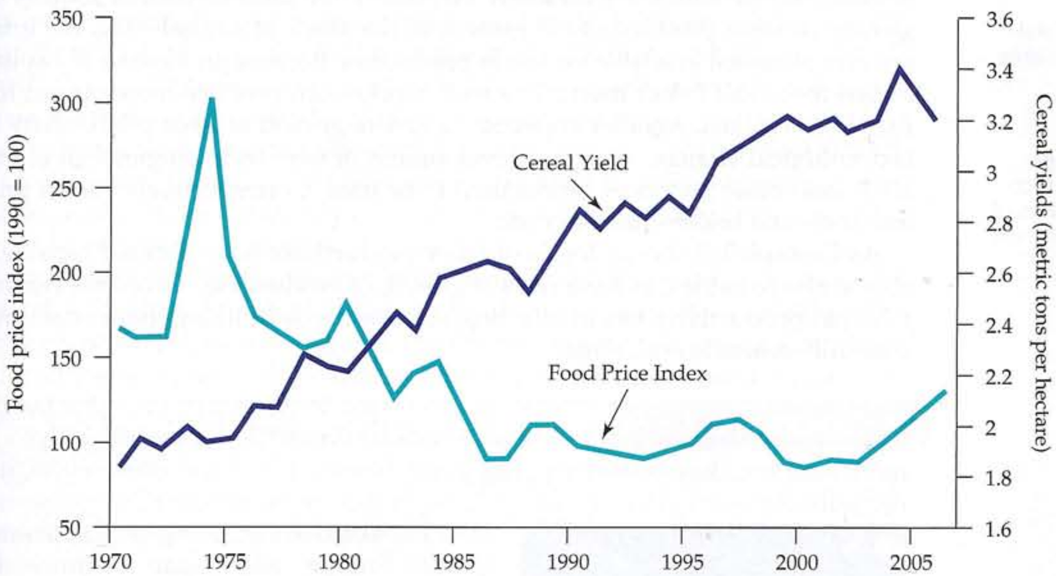


FIGURE 6.3 Cereal Yields and the World Price of Food

Cereal yields have increased. The average world price of food increased temporarily in the early 1970s but has declined since.

Labor Productivity

Although this is a textbook in microeconomics, many of the concepts developed here provide a foundation for macroeconomic analysis. Macroeconomists are particularly concerned with **labor productivity**—the average product of labor for an entire industry or for the economy as a whole. In this subsection we discuss labor productivity in the United States and a number of foreign countries. This topic is interesting in its own right, but will also help to illustrate one of the links between micro- and macroeconomics.

• **labor productivity** Average product of labor for an entire industry or for the economy as a whole.

Because the average product measures output per unit of labor input, it is relatively easy to measure (total labor input and total output are the only pieces of information you need). Labor productivity can provide useful comparisons across industries and for one industry over a long period. But labor productivity is especially important because it determines the real *standard of living* that a country can achieve for its citizens.

Productivity and the Standard of Living There is a simple link between labor productivity and the standard of living. In any particular year, the aggregate value of goods and services produced by an economy is equal to the payments made to all factors of production, including wages, rental payments to capital, and profit to firms. Consumers ultimately receive these factor payments in the form of wages, salaries, dividends, or interest payments. As a result, consumers in the aggregate can increase their rate of consumption in the long run only by increasing the total amount they produce.

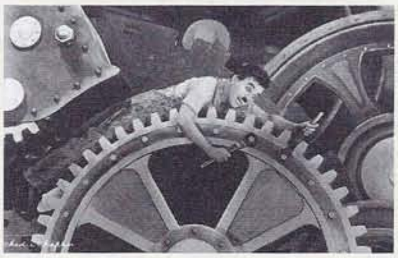
• **stock of capital** Total amount of capital available for use in production.

• **technological change** Development of new technologies allowing factors of production to be used more effectively.

Understanding the causes of productivity growth is an important area of research in economics. We do know that one of the most important sources of growth in labor productivity is growth in the **stock of capital**—i.e., the total amount of capital available for use in production. Because an increase in capital means more and better machinery, each worker can produce more output for each hour worked. Another important source of growth in labor productivity is **technological change**—i.e., the development of new technologies that allow labor (and other factors of production) to be used more effectively and to produce new and higher-quality goods.

As Example 6.2 shows, levels of labor productivity have differed considerably across countries, as have rates of growth of productivity. Given the central role that productivity has in affecting our standards of living, understanding these differences is important.

EXAMPLE 6.2 Labor Productivity and the Standard of Living



Will the standard of living in the United States, Europe, and Japan continue to improve, or will these economies barely keep future generations from being worse off than they are today? Because the real incomes of consumers in these countries increase only as fast as productivity does, the answer depends on the labor productivity of workers.

As Table 6.3 shows, the level of output per employed person in the United States in 2006 was higher than in other industrial countries. But two patterns over the post–World War II period have been disturbing. First, until the 1990s, productivity in the United States grew on average less rapidly than productivity in most other developed nations. Second, productivity growth during 1974–2006 was much lower in all developed countries than it had been in the past.[5]

TABLE 6.3 Labor Productivity in Developed Countries

	UNITED STATES	JAPAN	FRANCE	GERMANY	UNITED KINGDOM
	Real Output per Employed Person (2006)				
	$82,158	**$57,721**	**$72,949**	**$60,692**	**$65,224**
Years	*Annual Rate of Growth of Labor Productivity (%)*				
1960–1973	2.29	7.86	4.70	3.98	2.84
1974–1982	0.22	2.29	1.73	2.28	1.53
1983–1991	1.54	2.64	1.50	2.07	1.57
1992–2000	1.94	1.08	1.40	1.64	2.22
2001–2006	1.78	1.73	1.02	1.10	1.47

[5]Recent growth numbers on GDP, employment, and PPP data are from the OECD. For more information, visit **http://www.oecd.org**: select Frequently Requested Statistics within the Statistics directory.

Throughout most of the 1960–1991 period, Japan had the highest rate of productivity growth, followed by Germany and France. U.S. productivity growth was the lowest, even somewhat lower than that of the United Kingdom. This is partly due to differences in rates of investment and growth in the stock of capital in each country. The greatest capital growth during the postwar period was in Japan, France, and Germany, which were rebuilt substantially after World War II. To some extent, therefore, the lower rate of growth of productivity in the United States, when compared to that of Japan, France, and Germany, is the result of these countries catching up after the war.

Productivity growth is also tied to the natural resource sector of the economy. As oil and other resources began to be depleted, output per worker fell. Environmental regulations (e.g., the need to restore land to its original condition after strip-mining for coal) magnified this effect as the public became more concerned with the importance of cleaner air and water.

Observe from Table 6.3 that productivity growth in the United States accelerated in the 1990s. Some economists believe that information and communication technology (ICT) has been the key impetus for this growth. However, sluggish growth in more recent years suggests that ICT's contribution may have already peaked.

6.3 PRODUCTION WITH TWO VARIABLE INPUTS

We have completed our analysis of the short-run production function in which one input, labor, is variable, and the other, capital, is fixed. Now we turn to the long run, for which both labor and capital are variable. The firm can now produce its output in a variety of ways by combining different amounts of labor and capital. In this section, we will see how a firm can choose among combinations of labor and capital that generate the same output. In the first subsection, we will examine the scale of the production process, analyzing how output changes as input combinations are doubled, tripled, and so on.

Isoquants

Let's begin by examining the production technology of a firm that uses two inputs and can vary both of them. Suppose that the inputs are labor and capital and that they are used to produce food. Table 6.4 tabulates the output achievable for various combinations of inputs.

TABLE 6.4 Production with Two Variable Inputs

	LABOR INPUT				
Capital Input	**1**	**2**	**3**	**4**	**5**
1	20	40	55	65	75
2	40	60	75	85	90
3	55	75	90	100	105
4	65	85	100	110	115
5	75	90	105	115	120

Labor inputs are listed across the top row, capital inputs down the column on the left. Each entry in the table is the maximum (technically efficient) output that can be produced each year with each combination of labor and capital used over that year. For example, 4 units of labor per year and 2 units of capital per year yield 85 units of food per year. Reading along each row, we see that output increases as labor inputs are increased, while capital inputs remain fixed. Reading down each column, we see that output also increases as capital inputs are increased, while labor inputs remain fixed.

• **isoquant** Curve showing all possible combinations of inputs that yield the same output.

The information in Table 6.4 can also be represented graphically using isoquants. An **isoquant** is a *curve that shows all the possible combinations of inputs that yield the same output*. Figure 6.4 shows three isoquants. (Each axis in the figure measures the quantity of inputs.) These isoquants are based on the data in Table 6.4, but are drawn as smooth curves to allow for the use of fractional amounts of inputs.

For example, isoquant q_1 shows all combinations of labor and capital per year that together yield 55 units of output per year. Two of these points, *A* and *D*, correspond to Table 6.4. At *A*, 1 unit of labor and 3 units of capital yield 55 units of output; at *D*, the same output is produced from 3 units of labor and 1 unit of capital. Isoquant q_2 shows all combinations of inputs that yield 75 units of output and corresponds to the four combinations of labor and capital circled in the table (e.g., at *B*, where 2 units of labor and 3 units of capital are combined). Isoquant q_2 lies above and to the right of q_1 because obtaining a higher level of output requires more labor and capital. Finally, isoquant q_3 shows labor-capital combinations that yield 90 units of output. Point C, for example, involves 3 units

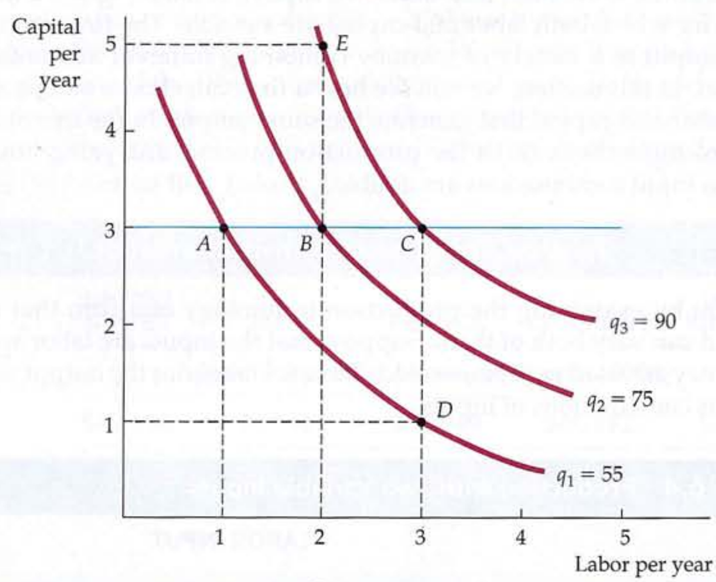


FIGURE 6.4 Production with Two Variable Inputs

Production isoquants show the various combinations of inputs necessary for the firm to produce a given output. A set of isoquants, or *isoquant map*, describes the firm's production function. Output increases as we move from isoquant q_1 (at which 55 units per year are produced at points such as *A* and *D*), to isoquant q_2 (75 units per year at points such as *B*) and to isoquant q_3 (90 units per year at points such as C and *E*).

of labor and 3 units of capital, whereas Point E involves 2 units of labor and 5 units of capital.

Isoquant Maps When a number of isoquants are combined in a single graph, we call the graph an **isoquant map**. Figure 6.4 shows three of the many isoquants that make up an isoquant map. An isoquant map is another way of describing a production function, just as an indifference map is a way of describing a utility function. Each isoquant corresponds to a different level of output, and the level of output increases as we move up and to the right in the figure.

• **isoquant map** Graph combining a number of isoquants, used to describe a production function.

Input Flexibility

Isoquants show the flexibility that firms have when making production decisions: They can usually obtain a particular output by substituting one input for another. It is important for managers to understand the nature of this flexibility. For example, fast-food restaurants have recently faced shortages of young, low-wage employees. Companies have responded by automating—adding self-service salad bars and introducing more sophisticated cooking equipment. They have also recruited older people to fill positions. As we will see in Chapters 7 and 8, by taking into account this flexibility in the production process, managers can choose input combinations that minimize cost and maximize profit.

Diminishing Marginal Returns

Even though both labor and capital are variable in the long run, it is useful for a firm that is choosing the optimal mix of inputs to ask what happens to output as each input is increased, with the other input held fixed. The outcome of this exercise is described in Figure 6.4, which reflects diminishing marginal returns to both labor and capital. We can see why there is diminishing marginal returns to labor by drawing a horizontal line at a particular level of capital—say, 3. Reading the levels of output from each isoquant as labor is increased, we note that each additional unit of labor generates less and less additional output. For example, when labor is increased from 1 unit to 2 (from A to B), output increases by 20 (from 55 to 75). However, when labor is increased by an additional unit (from B to C), output increases by only 15 (from 75 to 90). Thus there are diminishing marginal returns to labor both in the long and short run. Because adding one factor while holding the other factor constant eventually leads to lower and lower incremental output, the isoquant must become steeper as more capital is added in place of labor and flatter when labor is added in place of capital.

There are also diminishing marginal returns to capital. With labor fixed, the marginal product of capital decreases as capital is increased. For example, when capital is increased from 1 to 2 and labor is held constant at 3, the marginal product of capital is initially 20 (75 – 55) but falls to 15 (90 – 75) when capital is increased from 2 to 3.

Substitution Among Inputs

With two inputs that can be varied, a manager will want to consider substituting one input for another. The slope of each isoquant indicates how the quantity of one input can be traded off against the quantity of the other, while output

• **marginal rate of technical substitution (MRTS)** Amount by which the quantity of one input can be reduced when one extra unit of another input is used, so that output remains constant.

In §3.1, we explain that the marginal rate of substitution is the maximum amount of one good that the consumer is willing to give up to obtain one unit of another good.

is held constant. When the negative sign is removed, we call the slope the **marginal rate of technical substitution (MRTS)**. The *marginal rate of technical substitution of labor for capital* is the amount by which the input of capital can be reduced when one extra unit of labor is used, so that output remains constant. This is analogous to the marginal rate of substitution (MRS) in consumer theory. Recall from Section 3.1 that the MRS describes how consumers substitute among two goods while holding the level of satisfaction constant. Like the MRS, the MRTS is always measured as a positive quantity:

$$\text{MRTS} = -\text{Change in capital input/change in labor input}$$
$$= -\Delta K/\Delta L \text{ (for a fixed level of } q)$$

where ΔK and ΔL are small changes in capital and labor along an isoquant.

In Figure 6.5 the MRTS is equal to 2 when labor increases from 1 unit to 2 and output is fixed at 75. However, the MRTS falls to 1 when labor is increased from 2 units to 3, and then declines to 2/3 and to 1/3. Clearly, as more and more labor replaces capital, labor becomes less productive and capital becomes relatively more productive. Therefore, we need less capital to keep output constant, and the isoquant becomes flatter.

Diminishing MRTS We assume that there is a *diminishing MRTS*. In other words, the MRTS falls as we move down along an isoquant. The mathematical implication is that isoquants, like indifference curves, are *convex*, or bowed inward. This is indeed the case for most production technologies. The diminishing MRTS tells us that the productivity of any one input is limited. As more

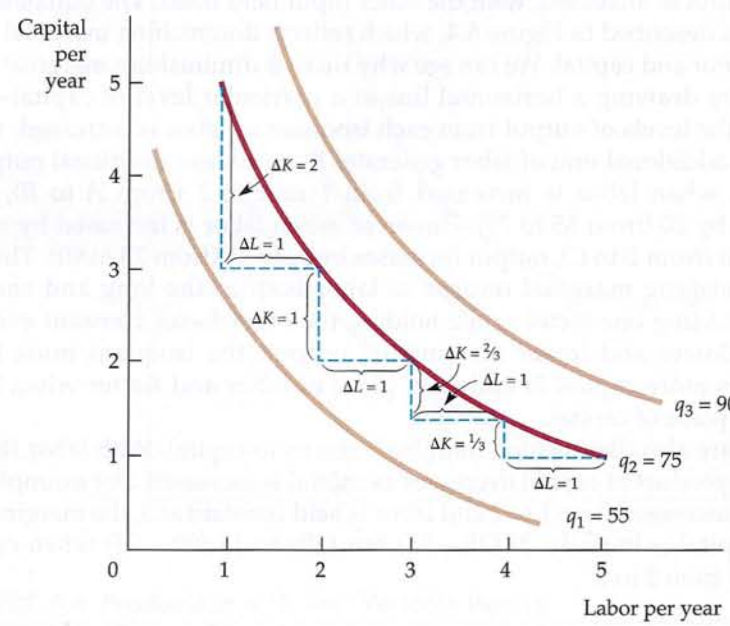


FIGURE 6.5 Marginal Rate of Technical Substitution

Like indifference curves, isoquants are downward sloping and convex. The slope of the isoquant at any point measures the marginal rate of technical substitution—the ability of the firm to replace capital with labor while maintaining the same level of output. On isoquant q_2, the MRTS falls from 2 to 1 to 2/3 to 1/3.

and more labor is added to the production process in place of capital, the productivity of labor falls. Similarly, when more capital is added in place of labor, the productivity of capital falls. Production needs a balanced mix of both inputs.

In §3.1, we explain that an indifference curve is convex if the marginal rate of substitution diminishes as we move down along the curve.

As our discussion has just suggested, the MRTS is closely related to the marginal products of labor MP_L and capital MP_K. To see how, imagine adding some labor and reducing the amount of capital sufficient to keep output constant. The additional output resulting from the increased labor input is equal to the additional output per unit of additional labor (the marginal product of labor) times the number of units of additional labor:

$$\text{Additional output from increased use of labor} = (MP_L)(\Delta L)$$

Similarly, the decrease in output resulting from the reduction in capital is the loss of output per unit reduction in capital (the marginal product of capital) times the number of units of capital reduction:

$$\text{Reduction in output from decreased use of capital} = (MP_K)(\Delta K)$$

Because we are keeping output constant by moving along an isoquant, the total change in output must be zero. Thus,

$$(MP_L)(\Delta L) + (MP_K)(\Delta K) = 0$$

Now, by rearranging terms we see that

$$(MP_L)/(MP_K) = -(\Delta K/\Delta L) = MRTS \qquad \textbf{(6.2)}$$

Equation (6.2) tells us that *the marginal rate of technical substitution between two inputs is equal to the ratio of the marginal products of the inputs*. This formula will be useful when we look at the firm's cost-minimizing choice of inputs in Chapter 7.

Production Functions—Two Special Cases

Two extreme cases of production functions show the possible range of input substitution in the production process. In the first case, shown in Figure 6.6, inputs to production are *perfect substitutes* for one another. Here the MRTS is constant at all points on an isoquant. As a result, the same output (say q_3) can be produced with mostly capital (at *A*), with mostly labor (at *C*), or with a balanced combination of both (at *B*). For example, musical instruments can be manufactured almost entirely with machine tools or with very few tools and highly skilled labor.

In §3.1, we explain that two goods are perfect substitutes if the marginal rate of substitution of one for the other is a constant.

Figure 6.7 illustrates the opposite extreme, the **fixed-proportions production function**, sometimes called a *Leonitief production function*. In this case, it is impossible to make any substitution among inputs. Each level of output requires a specific combination of labor and capital: Additional output cannot be obtained unless more capital and labor are added in specific proportions. As a result, the isoquants are L-shaped, just as indifference curves are L-shaped when two goods are perfect complements. An example is the reconstruction of concrete sidewalks using jackhammers. It takes one person to use a jackhammer—neither two people and one jackhammer nor one person and two jackhammers will increase production. As another example, suppose that a cereal company offers a new breakfast cereal, Nutty Oat Crunch, whose two inputs, not surprisingly, are oats and nuts. The secret formula for the cereal requires exactly one ounce of nuts for every four ounces of oats in every serving. If the company were to purchase additional nuts but not additional oats, the output of cereal would remain unchanged, since the

• **fixed-proportions production function** Production function with L-shaped isoquants, so that only one combination of labor and capital can be used to produce each level of output.

In §3.1, we explain that two goods are perfect complements when the indifference curves for the goods are shaped as right angles.

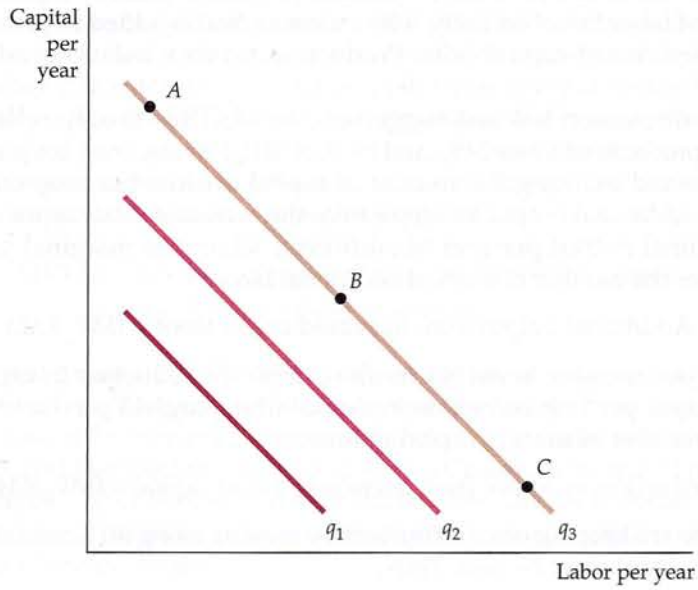


FIGURE 6.6 Isoquants When Inputs Are Perfect Substitutes

When the isoquants are straight lines, the MRTS is constant. Thus the rate at which capital and labor can be substituted for each other is the same no matter what level of inputs is being used. Points *A*, *B*, and *C* represent three different capital-labor combinations that generate the same output q_3.

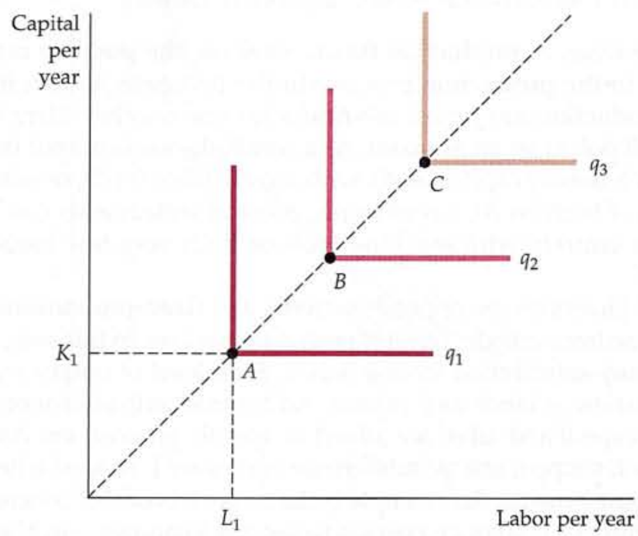


FIGURE 6.7 Fixed-Proportions Production Function

When the isoquants are L-shaped, only one combination of labor and capital can be used to produce a given output (as at point *A* on isoquant q_1, point *B* on isoquant q_2, and point C on isoquant q_3). Adding more labor alone does not increase output, nor does adding more capital alone.

nuts must be combined with the oats in a fixed proportion. Similarly, purchasing additional oats without additional nuts would also be unproductive.

In Figure 6.7 points A, B, and C represent technically efficient combinations of inputs. For example, to produce output q_1, a quantity of labor L_1 and capital K_1 can be used, as at A. If capital stays fixed at K_1, adding more labor does not change output. Nor does adding capital with labor fixed at L_1. Thus, on the vertical and the horizontal segments of the L-shaped isoquants, either the marginal product of capital or the marginal product of labor is zero. Higher output results only when both labor and capital are added, as in the move from input combination A to input combination B.

The fixed-proportions production function describes situations in which methods of production are limited. For example, the production of a television show might involve a certain mix of capital (camera and sound equipment, etc.) and labor (producer, director, actors, etc.). To make more television shows, all inputs to production must be increased proportionally. In particular, it would be difficult to increase capital inputs at the expense of labor, because actors are necessary inputs to production (except perhaps for animated films). Likewise, it would be difficult to substitute labor for capital, because filmmaking today requires sophisticated film equipment.

EXAMPLE 6.3 A Production Function for Wheat

Crops can be produced using different methods. Food grown on large farms in the United States is usually produced with a *capital-intensive technology*, which involves substantial investments in capital, such as buildings and equipment, and relatively little input of labor. However, food can also be produced using very little capital (a hoe) and a lot of labor (several people with the patience and stamina to work the soil). One way to describe the agricultural production process is to show one isoquant (or more) that describes the combination of inputs which generates a given level of output (or several output levels). The description that follows comes from a production function for wheat that was estimated statistically.[6]

Figure 6.8 shows one isoquant, associated with the production function, corresponding to an output of 13,800 bushels of wheat per year. The manager of the farm can use this isoquant to decide whether it is profitable to hire more labor or use more machinery. Assume the farm is currently operating at A, with a labor input L of 500 hours and a capital input K of 100 machine hours. The manager decides to experiment by using only 90 hours of machine time. To produce the same crop per year, he finds that he needs to replace this machine time by adding 260 hours of labor.

The results of this experiment tell the manager about the shape of the wheat production isoquant. When he compares points A (where $L = 500$ and $K = 100$) and B (where $L = 760$ and $K = 90$) in Figure 6.8, both of which are on the same isoquant, the manager finds that the marginal rate of technical substitution is equal to 0.04 ($-\Delta K/\Delta L = -(-10)/260 = .04$).

[6]The food production function on which this example is based is given by the equation $q = 100(K^{.8}L^{.2})$, where q is the rate of output in bushels of wheat per year, K is the quantity of machines in use per year, and L is the number of hours of labor per year.

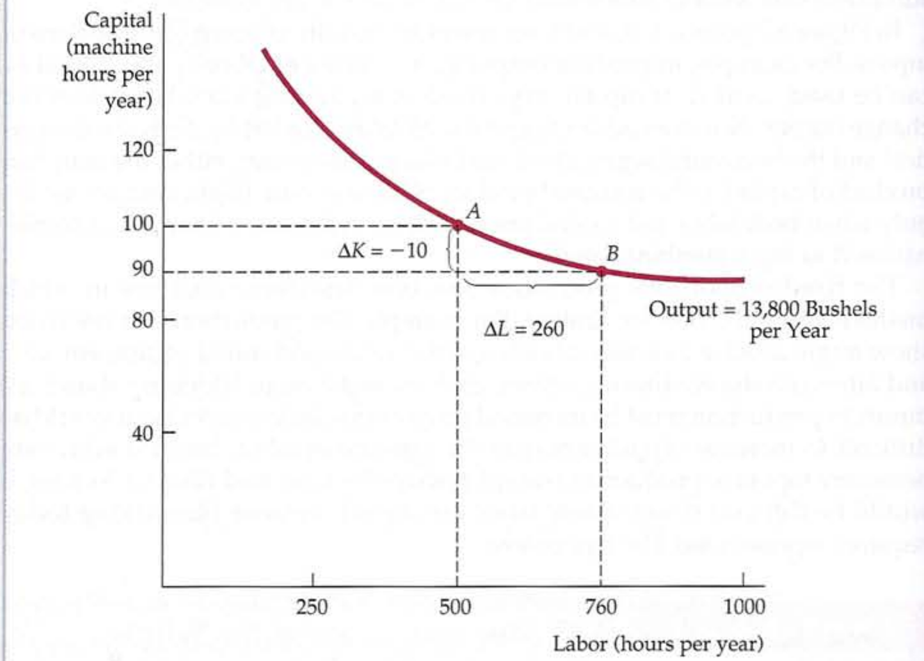


FIGURE 6.8 Isoquant Describing the Production of Wheat

A wheat output of 13,800 bushels per year can be produced with different combinations of labor and capital. The more capital-intensive production process is shown as point *A*, the more labor-intensive process as point *B*. The marginal rate of technical substitution between *A* and *B* is 10/260 = 0.04.

The MRTS reveals the nature of the trade-off involved in adding labor and reducing the use of farm machinery. Because the MRTS is substantially less than 1 in value, the manager knows that when the wage of a laborer is equal to the cost of running a machine, he ought to use more capital. (At his current level of production, he needs 260 units of labor to substitute for 10 units of capital.) In fact, he knows that unless labor is much less expensive than the use of a machine, his production process ought to become more capital-intensive.

The decision about how many laborers to hire and machines to use cannot be fully resolved until we discuss the costs of production in the next chapter. However, this example illustrates how knowledge about production isoquants and the marginal rate of technical substitution can help a manager. It also suggests why most farms in the United States and Canada, where labor is relatively expensive, operate in the range of production in which the MRTS is relatively high (with a high capital-to-labor ratio), whereas farms in developing countries, in which labor is cheap, operate with a lower MRTS (and a lower capital-to-labor ratio).[7] The exact labor/capital combination to use depends on input prices, a subject that we discuss in Chapter 7.

[7]With the production function given in footnote 6, it is not difficult (using calculus) to show that the marginal rate of technical substitution is given by MRTS = $(MP_L/MP_K) = (1/4)(K/L)$. Thus, the MRTS decreases as the capital-to-labor ratio falls. For an interesting study of agricultural production in Israel, see Richard E. Just, David Zilberman, and Eithan Hochman, "Estimation of Multicrop Production Functions," *American Journal of Agricultural Economics* 65 (1983): 770–80.

6.4 RETURNS TO SCALE

Our analysis of input substitution in the production process has shown us what happens when a firm substitutes one input for another while keeping output constant. However, in the long run, with all inputs variable, the firm must also consider the best way to increase output. One way to do so is to change the *scale* of the operation by increasing *all of the inputs to production in proportion*. If it takes one farmer working with one harvesting machine on one acre of land to produce 100 bushels of wheat, what will happen to output if we put two farmers to work with two machines on two acres of land? Output will almost certainly increase, but will it double, more than double, or less than double? **Returns to scale** is the rate at which output increases as inputs are increased proportionately. We will examine three different cases: increasing, constant, and decreasing returns to scale.

• **returns to scale** Rate at which output increases as inputs are increased proportionately.

Increasing Returns to Scale If output more than doubles when inputs are doubled, there are **increasing returns to scale**. This might arise because the larger scale of operation allows managers and workers to specialize in their tasks and to make use of more sophisticated, large-scale factories and equipment. The automobile assembly line is a famous example of increasing returns.

• **increasing returns to scale** Situation in which output more than doubles when all inputs are doubled.

The prospect of increasing returns to scale is an important issue from a public-policy perspective. If there are increasing returns, then it is economically advantageous to have one large firm producing (at relatively low cost) rather than to have many small firms (at relatively high cost). Because this large firm can control the price that it sets, it may need to be regulated. For example, increasing returns in the provision of electricity is one reason why we have large, regulated power companies.

Constant Returns to Scale A second possibility with respect to the scale of production is that output may double when inputs are doubled. In this case, we say there are **constant returns to scale**. With constant returns to scale, the size of the firm's operation does not affect the productivity of its factors: Because one plant using a particular production process can easily be replicated, two plants produce twice as much output. For example, a large travel agency might provide the same service per client and use the same ratio of capital (office space) and labor (travel agents) as a small agency that services fewer clients.

• **constant returns to scale** Situation in which output doubles when all inputs are doubled.

Decreasing Returns to Scale Finally, output may less than double when all inputs double. This case of **decreasing returns to scale** applies to some firms with large-scale operations. Eventually, difficulties in organizing and running a large-scale operation may lead to decreased productivity of both labor and capital. Communication between workers and managers can become difficult to monitor as the workplace becomes more impersonal. Thus, the decreasing-returns case is likely to be associated with the problems of coordinating tasks and maintaining a useful line of communication between management and workers.

• **decreasing returns to scale** Situation in which output less than doubles when all inputs are doubled.

Describing Returns to Scale

Returns to scale need not be uniform across all possible levels of output. For example, at lower levels of output, the firm could have increasing returns to scale, but constant and eventually decreasing returns at higher levels of output.

The presence or absence of returns to scale is seen graphically in the two parts of Figure 6.9. The line 0*A* from the origin in each panel describes a production process in which labor and capital are used as inputs to produce various levels of output in the ratio of 5 hours of labor to 2 hours of machine time. In Figure 6.9(a), the firm's production function exhibits constant returns to scale. When 5 hours of labor and 2 hours of machine time are used, an output of 10 units is produced. When both inputs double, output doubles from 10 to 20 units; when both inputs triple, output triples, from 10 to 30 units. Put differently, twice as much of both inputs is needed to produce 20 units, and three times as much is needed to produce 30 units.

In Figure 6.9(b), the firm's production function exhibits increasing returns to scale. Now the isoquants come closer together as we move away from the origin along 0*A*. As a result, *less* than twice the amount of both inputs is needed to increase production from 10 units to 20; substantially less than three times the inputs are needed to produce 30 units. The reverse would be true if the production function exhibited decreasing returns to scale (not shown here). With decreasing returns, the isoquants are increasingly distant from one another as output levels increase proportionally.

Returns to scale vary considerably across firms and industries. Other things being equal, the greater the returns to scale, the larger the firms in an industry are likely to be. Because manufacturing involves large investments in capital equipment, manufacturing industries are more likely to have increasing returns to scale than service-oriented industries. Services are more labor-intensive and can usually be provided as efficiently in small quantities as they can on a large scale.

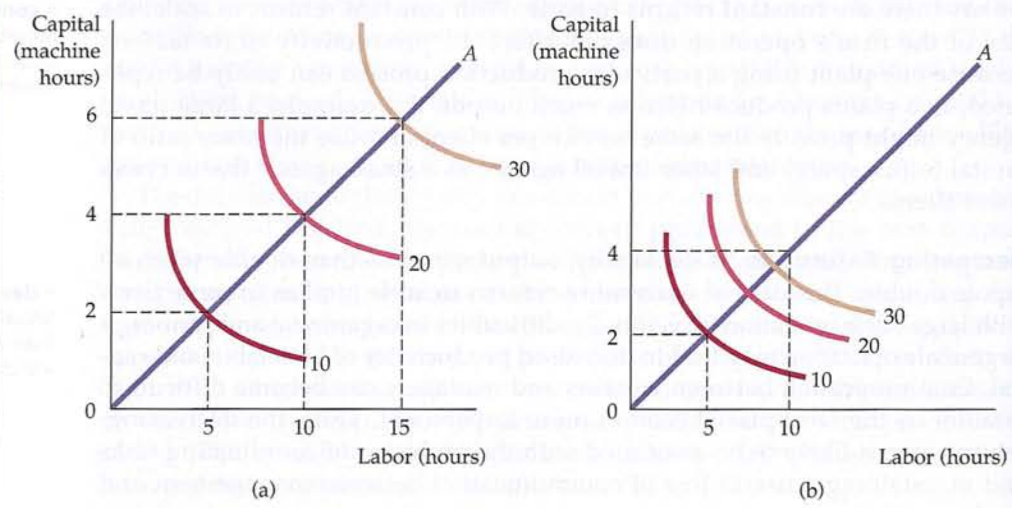


FIGURE 6.9 Returns to Scale

When a firm's production process exhibits constant returns to scale as shown by a movement along line 0*A* in part **(a)**, the isoquants are equally spaced as output increases proportionally. However, when there are increasing returns to scale as shown in **(b)**, the isoquants move closer together as inputs are increased along the line.

EXAMPLE 6.4 Returns to Scale in the Carpet Industry

The carpet industry in the United States centers around the town of Dalton in northern Georgia. From a relatively small industry with many small firms in the first half of the twentieth century, it grew rapidly and became a major industry with a large number of firms of all sizes. For example, the top five carpet manufacturers, ranked by shipments in millions of dollars in 2005, are shown in Table 6.5.[8]

Currently, there are three relatively large manufacturers (Shaw, Mohawk, and Beaulieu), along with a number of smaller producers. There are also many retailers, wholesale distributors, buying groups, and national retail chains. The carpet industry has grown rapidly for several reasons. Consumer demand for wool, nylon, and polypropylene carpets in commercial and residential uses has skyrocketed. In addition, innovations such as the introduction of larger, faster, and more efficient carpet-tufting machines have reduced costs and greatly increased carpet production. Along with the increase in production, innovation and competition have worked together to reduce real carpet prices.

To what extent, if any, can the growth of the carpet industry be explained by the presence of returns to scale? There have certainly been substantial improvements in the processing of key production inputs (such as stain-resistant yarn) and in the distribution of carpets to retailers and consumers. But what about the production of carpets? Carpet production is capital intensive—manufacturing plants require heavy investments in high-speed tufting machines that turn various types of yarn into carpet, as well as machines that put the backings onto the carpets, cut the carpets into appropriate sizes, and package, label, and distribute them.

Overall, physical capital (including plant and equipment) accounts for about 77 percent of a typical carpet manufacturer's costs, while labor accounts for the remaining 23 percent. Over time, the major carpet manufacturers have increased the scale of their operations by putting larger and more efficient tufting machines into larger plants. At the same time, the use of labor in these plants has also increased significantly. The result? Proportional increases in inputs have

TABLE 6.5 The U.S. Carpet Industry

Carpet Sales, 2005 (Millions of Dollars per Year)	
1. Shaw	4346
2. Mohawk	3779
3. Beaulieu	1115
4. Interface	421
5. Royalty	298

[8]*Floor Focus*, May 2005.

resulted in a more than proportional increase in output for these larger plants. For example, a doubling of capital and labor inputs might lead to a 110-percent increase in output. This pattern has not, however, been uniform across the industry. Most smaller carpet manufacturers have found that small changes in scale have little or no effect on output; i.e., small proportional increases in inputs have only increased output proportionally.

We can therefore characterize the carpet industry as one in which there are constant returns to scale for relatively small plants but increasing returns to scale for larger plants. These increasing returns, however, are limited, and we can expect that if plant size were increased further, there would eventually be decreasing returns to scale.

SUMMARY

1. A *production function* describes the maximum output that a firm can produce for each specified combination of inputs.
2. In the short run, one or more inputs to the production process are fixed. In the long run, all inputs are potentially variable.
3. Production with one variable input, labor, can be usefully described in terms of the *average product of labor* (which measures output per unit of labor input) and the *marginal product of labor* (which measures the additional output as labor is increased by 1 unit).
4. According to the *law of diminishing marginal returns*, when one or more inputs are fixed, a variable input (usually labor) is likely to have a marginal product that eventually diminishes as the level of input increases.
5. An *isoquant* is a curve that shows all combinations of inputs that yield a given level of output. A firm's production function can be represented by a series of isoquants associated with different levels of output.
6. Isoquants always slope downward because the marginal product of all inputs is positive. The shape of each isoquant can be described by the marginal rate of technical substitution at each point on the isoquant. The *marginal rate of technical substitution of labor for capital* (MRTS) is the amount by which the input of capital can be reduced when one extra unit of labor is used so that output remains constant.
7. The standard of living that a country can attain for its citizens is closely related to its level of labor productivity. Decreases in the rate of productivity growth in developed countries are due in part to the lack of growth of capital investment.
8. The possibilities for substitution among inputs in the production process range from a production function in which inputs are *perfect substitutes* to one in which the proportions of inputs to be used are fixed (a *fixed-proportions production function*).
9. In long-run analysis, we tend to focus on the firm's choice of its scale or size of operation. *Constant returns to scale* means that doubling all inputs leads to doubling output. *Increasing returns to scale* occurs when output more than doubles when inputs are doubled; *decreasing returns to scale* applies when output less than doubles.

QUESTIONS FOR REVIEW

1. What is a production function? How does a long-run production function differ from a short-run production function?
2. Why is the marginal product of labor likely to increase initially in the short run as more of the variable input is hired?
3. Why does production eventually experience diminishing marginal returns to labor in the short run?
4. You are an employer seeking to fill a vacant position on an assembly line. Are you more concerned with the average product of labor or the marginal product of labor for the last person hired? If you observe that your average product is just beginning to decline, should you hire any more workers? What does this situation imply about the marginal product of your last worker hired?

5. What is the difference between a production function and an isoquant?
6. Faced with constantly changing conditions, why would a firm ever keep *any* factors fixed? What criteria determine whether a factor is fixed or variable?
7. Isoquants can be convex, linear, or L-shaped. What does each of these shapes tell you about the nature of the production function? What does each of these shapes tell you about the MRTS?
8. Can an isoquant ever slope upward? Explain.
9. Explain the term "marginal rate of technical substitution." What does a MRTS = 4 mean?
10. Explain why the marginal rate of technical substitution is likely to diminish as more and more labor is substituted for capital.
11. It is possible to have diminishing returns to a single factor of production and constant returns to scale at the same time. Discuss.
12. Can a firm have a production function that exhibits increasing returns to scale, constant returns to scale, and decreasing returns to scale as output increases? Discuss.
13. Give an example of a production process in which the short run involves a day or a week and the long run any period longer than a week.

EXERCISES

1. The menu at Joe's coffee shop consists of a variety of coffee drinks, pastries, and sandwiches. The marginal product of an additional worker can be defined as the number of customers that can be served by that worker in a given time period. Joe has been employing one worker, but is considering hiring a second and a third. Explain why the marginal product of the second and third workers might be higher than the first. Why might you expect the marginal product of additional workers to diminish eventually?
2. Suppose a chair manufacturer is producing in the short run (with its existing plant and equipment). The manufacturer has observed the following levels of production corresponding to different numbers of workers:

Number of Workers	Number of Chairs
1	10
2	18
3	24
4	28
5	30
6	28
7	25

   a. Calculate the marginal and average product of labor for this production function.
   b. Does this production function exhibit diminishing returns to labor? Explain.
   c. Explain intuitively what might cause the marginal product of labor to become negative.

3. Fill in the gaps in the table below.

Quantity of Variable Input	Total Output	Marginal Product of Variable Input	Average Product of Variable Input
0	0	—	—
1	225		
2			300
3		300	
4	1140		
5		225	
6			225

4. A political campaign manager must decide whether to emphasize television advertisements or letters to potential voters in a reelection campaign. Describe the production function for campaign votes. How might information about this function (such as the shape of the isoquants) help the campaign manager to plan strategy?
5. For each of the following examples, draw a representative isoquant. What can you say about the marginal rate of technical substitution in each case?
   a. A firm can hire only full-time employees to produce its output, or it can hire some combination of full-time and part-time employees. For each full-time worker let go, the firm must hire an increasing number of temporary employees to maintain the same level of output.
   b. A firm finds that it can always trade two units of labor for one unit of capital and still keep output constant.

c. A firm requires exactly two full-time workers to operate each piece of machinery in the factory.

6. A firm has a production process in which the inputs to production are perfectly substitutable in the long run. Can you tell whether the marginal rate of technical substitution is high or low, or is further information necessary? Discuss.

7. The marginal product of labor in the production of computer chips is 50 chips per hour. The marginal rate of technical substitution of hours of labor for hours of machine capital is 1/4. What is the marginal product of capital?

8. Do the following functions exhibit increasing, constant, or decreasing returns to scale? What happens to the marginal product of each individual factor as that factor is increased and the other factor held constant?
   a. $q = 3L + 2K$
   b. $q = (2L + 2K)^{1/2}$
   c. $q = 3LK^2$
   d. $q = L^{1/2}K^{1/2}$
   e. $q = 4L^{1/2} + 4K$

9. The production function for the personal computers of DISK, Inc., is given by

$$q = 10K^{0.5}L^{0.5}$$

where q is the number of computers produced per day, K is hours of machine time, and L is hours of labor input. DISK's competitor, FLOPPY, Inc., is using the production function

$$q = 10K^{0.6}L^{0.4}$$

   a. If both companies use the same amounts of capital and labor, which will generate more output?
   b. Assume that capital is limited to 9 machine hours, but labor is unlimited in supply. In which company is the marginal product of labor greater? Explain.

10. In Example 6.3, wheat is produced according to the production function

$$q = 100(K^{0.8}L^{0.2})$$

   a. Beginning with a capital input of 4 and a labor input of 49, show that the marginal product of labor and the marginal product of capital are both decreasing.
   b. Does this production function exhibit increasing, decreasing, or constant returns to scale?