Chapter 6

Production and Cost

(Getty Images)

After reading this chapter,
you will understand the following:

1. How economists view the concepts of cost and profit
2. The distinction between short-run and long-run time horizons
3. How costs vary in response to changes in the quantity of a variable input
4. The graphical representation of production costs
5. The choices a firm faces in the course of long-run expansion

Before reading this chapter,
make sure you know the following concepts:

Opportunity cost
Entrepreneurship
Economic rent
Rational choice

Chapter
Outline

Business firms are among the most visible features of a market economy. Some of them are giants, like Google, that seem to reach into every corner of our lives. Others are small local operations like a lawn service or hotdog stand. As consumers, we count on business firms for the goods and services we buy; workers count on them to provide jobs; and governments count on them to pay taxes. The coming chapters will look at firms from several perspectives. This chapter looks at the way they transform inputs of labor, capital, and natural resources into useful goods and services. Chapters 7, 8, and 9 look at how firms compete with one another in the effort to earn profits and avoid losses. Chapter 10 looks at issues in public policy that arise when competition does not work smoothly. By the time we have put the whole picture together, we will see just why firms are so important to a market economy.

6.1 Costs and Profits

As we saw in Chapter 4, microeconomic models focus on objectives, constraints, and choices. What objectives and constraints shape the choices of business firms? For most firms, earning profits and avoiding losses are the key objectives. The main constraints on a firm's opportunities are its costs of production and the demand for its output. This chapter will focus on profits and costs. The next chapter will introduce the demand constraint.

6.1a The Profit Motive

The idea that firms are in business to maximize profits is familiar but not without controversy. It is open to an objection similar to that raised against the assumption of rationality in consumer choice. It implies too narrow a view of human nature, its critics say; profit is important, but it is hardly the only thing businesses are interested in. Managers of some firms seem to display other-regarding preferences. They spend large amounts on supporting the arts or aiding the homeless, and they exhibit concern for their workers, their customers, and the environment beyond any level that might increase profits by enhancing the firm's public image. Other firms are led by egotists who will risk all, including profit itself, in pursuit of building a personal empire. Still others are run by people who are content to earn a minimum profit required for survival and, if things go well, to take Wednesday afternoons off for a game of golf or a trip to the mountains, rather than toil away for a few extra dollars.

Economists have two answers to critics of the profit maximization assumption. One defense of the assumption is that economists do not mean profit maximization as a complete description of the motives of all business managers. Rather, it is a simplification designed to give a sharper structure to theories about the way changes in costs or demand affect decisions. Following the principle of Ockham's razor, we should discard a simple theory for a more complex one only when the simple theory fails to explain behavior observed in the real world. In practice, theories based on profit maximization explain a great deal of what firms do. In some special situations, we can improve the theories by considering objectives other than profit, but we do not always need to do so.

The *survivorship principle* is a second defense of the profit maximization assumption. Imagine that ownership of firms is at first distributed randomly among people who are inclined to pursue the objective of profit and others who favor the objectives of benevolence, ego satisfaction, or the easy life. Over time, the firms that maximized profit would increase their capital and grow steadily through reinvestment or acquisition. Those that pursued other objectives would at best have fewer profits to invest in expansion and at worst go out of business because of losses. Over time, those firms that maximized profit would be the most likely to survive in the market.

6.1b The Nature of Costs

Profit is the difference between revenue and costs, so we cannot get far in discussing profits without looking at costs. As we learned in Chapter 1, economists think first and foremost in terms of *opportunity cost*. There are never enough resources to satisfy all wants, so the decision to produce any one thing implies the need to forgo using the same resources to produce something else. The opportunity costs of production are a fundamental constraint on a firm's ability to maximize its profits. In this section, we will explore several aspects of production costs and explain their relationship to one another.

Implicit and Explicit Costs

Opportunity costs of production have several components. They include whatever the firm pays to workers, investors, owners of natural resources, and suppliers of intermediate goods—such as parts, semifinished materials, and business services. (As explained in Chapter 4, production may also impose external costs on bystanders through pollution, noise, or congestion, but we will wait until a later chapter to address external costs.)

We can begin by making a distinction between explicit and implicit costs. **Explicit costs** take the form of payments to suppliers of intermediate goods and factors of production. They include workers' wages, managers' salaries, salespeople's commissions, payments to banks and other suppliers of financial services, fees for legal advice, shipping charges, and many other things. Not all opportunity costs take the form of explicit payments to parties outside a firm, however. Most firms also incur **implicit costs**, by which we mean the opportunity costs of using resources that a firm's owners (or the firm itself, as a legal entity) contribute without receiving explicit payment. Firms normally do not record implicit costs in their accounts, but that does not make those costs any less real.

Explicit costs
Opportunity costs that take the form of explicit payments to suppliers of factors of production and intermediate goods

Implicit costs
Opportunity costs of using resources that a firm's owners (or the firm itself, as a legal entity) contribute without receiving explicit payment

Consider, for example, owners of small firms who work together with hired employees without receiving a salary for themselves. By doing so, they give up the opportunity to work for someone else. The correct measure of the opportunity cost of the owners' labor would be the wage or salary they could earn from the next-best employment opportunity.

Most businesses require some funding upfront to get started and to buy equipment. Broadly speaking, there are two ways to finance these startup costs: borrowing the funds or using the savings of the firm's owners (or the savings of an equity investor). Both solutions involve opportunity costs related to interest. The interest payment on a small business loan is an explicit cost: it requires that the firm actually pay money to an outside entity, the bank. For the small business owner who uses her own savings to finance the business, the opportunity cost is the interest that she *could have* earned by investing that money elsewhere but that she has forgone, making the interest an implicit cost. Thus, interest costs—like many forms of opportunity costs—can be explicit or implicit.


(Shutterstock)

A small business incurs an implicit cost when its owner works alongside employees without receiving a salary.

Costs and Profits

The distinction between explicit and implicit costs is important for understanding the meaning of the term *profit*. Economists use the term **pure economic profit** to mean the difference between a firm's total revenue and all of its opportunity costs, including both explicit and implicit costs. It is important to distinguish pure economic profit from two other uses of the term profit.

First, in the business world, we often use *profit* to mean revenue minus only explicit costs, without considering implicit costs. Economists call that concept **accounting profit** because it considers only the explicit payments that appear in the firm's written accounts.[1] The formulas for pure economic profit and accounting profit are as follows:

Pure economic profit = Revenue – Explicit costs – Implicit costs

Accounting profit = Revenue – Explicit costs

Notice that we could substitute accounting profit into the formula for pure economic profit, like this:

Pure economic profit = Accounting profit – Implicit costs

Alternatively, we could rearrange that equation:

Accounting profit = Pure economic profit + Implicit costs

Second, pure economic profit needs to be distinguished from so-called **normal profit**, a term that refers to the opportunity cost of capital contributed by the firm's owners. Suppose, for example, that you use $200,000 of your own savings as capital for a new business. You could, instead, invest in securities that pay a 10 percent rate of return, or $20,000 per year. That $20,000 would be your opportunity cost of capital. It represents the return your funds would have earned in the best alternative use. The same reasoning would apply if outside shareholders contributed the capital, rather than an entrepreneur who actively participated in managing the business.

To understand why we sometimes call the opportunity cost of owners' capital "normal profit," consider a firm whose implicit costs only include the implicit opportunity cost of capital. In order for such a firm to earn zero economic profit, its accounting profit would have to be *equal* to its implicit opportunity cost of capital. (Check the math using the formulas above!) We could call that rate of accounting profit "normal" in the sense that it is just enough to make it worthwhile for owners to invest their capital in this firm, rather than in the next-best line of business available. Lines of business that earned more than this (that is, a positive pure economic profit) would be "abnormally" profitable and would swiftly attract new investors and competitors. Those that earned less would be less than "normally" profitable and would tend to shrink as investors channeled their capital elsewhere. In terms of pure economic profit, a firm that is less than normally profitable is experiencing economic losses, even if accounting profit is slightly positive.

Pure economic profit
The sum that remains when we subtract both explicit and implicit costs from total revenue

Accounting profit
Total revenue minus explicit costs

Normal profit
The implicit opportunity cost of capital contributed by the firm's owners

If a firm has other implicit costs in addition to those of owners' capital, its accounting profit must be sufficient to cover them, too, in order to earn zero economic profit. We can express this idea in terms of any of the following equations:

Accounting profit = Pure economic profit + Implicit costs

= Pure economic profit + Implicit cost of capital + Other implicit costs

= Pure economic profit + Normal profit + Other implicit costs

Notice that if pure economic profit is equal to zero, and there are no other implicit costs, then accounting profit and normal profit will be equal. If accounting profits are less than normal profits, what can you conclude about pure economic profit?

At several points in this chapter and the next, it will be convenient to use an imaginary business as a basis for numerical examples of the concepts that we will introduce. We will call our imaginary business Fieldcom Inc. It is a small business started by a couple named Ralph and Andrea Martin. The Martins buy commonly available parts and assemble them into special-purpose smartphones that are "ruggedized" so that they can be used not only in an office or on a commuter train but also in stressful environments like a desert oil field, a tropical mining site, or an ocean racing yacht.

Figure 6–1 uses Fieldcom Inc. to illustrate the concepts of pure economic profit, accounting profit, and normal profit. The figure shows Fieldcom as having earned total revenues of $600,000 in the past year. Explicit costs—salaries paid to employees and materials purchased—came to $400,000. That left an accounting profit of $200,000. The explicit costs do not include all of the firm's opportunity costs, however. Both Andrea and Ralph Martin gave up good jobs to start the firm. Figure 6–1 lists their combined former income of $160,000 as an implicit cost of production. Another implicit cost is the $20,000 of forgone annual income that the Martins could have earned on $200,000 of personal savings if they had invested it elsewhere instead of in their business. This is the firm's opportunity cost of capital, or the normal profit or normal return on capital required to attract capital to this use rather than to the best alternative use. When we subtract both explicit and implicit costs (normal profit) from revenue, the firm has a pure economic profit of $20,000.

Figure 6–1 Accounts of Fieldcom Inc.

Total revenue		$600,000
	Less explicit costs:	
Wages and salaries		300,000
Materials and other		100,000

Equals accounting profit		$200,000
	Less implicit costs:	
Forgone salary, Andrea Martin		80,000
Forgone salary, Ralph Martin		80,000
Opportunity cost of capital		20,000

Equals pure economic profit		$20,000

This figure shows the implicit and explicit costs of the imaginary firm, Fieldcom Inc., owned by entrepreneurs Ralph and Andrea Martin. Total revenue minus explicit costs equals accounting profit. Subtracting implicit costs from this quantity yields pure economic profit. Another term for the opportunity cost of the capital and time contributed by the Martins is *normal profit*.

Costs Are Subjective

A final word is in order regarding costs. Comparing the theory of consumer choice in Chapter 5 to the theory of production costs, it may at first appear that we are moving from an area of economics governed by the *subjective* valuations of personal tastes and preferences that underlie demand to one of *objective* valuations governed by the hard realities of production; but that is true only in part, if at all.

Yes, business managers and their accountants do try to record costs using consistent methods that are as free as possible from wishful thinking and intentional bias. In that sense, cost accounting is objective. In a deeper sense, though, the theory of cost is just as subjective as is the theory of consumer choice. All costs are *opportunity costs* that reflect the value that a firm would have earned by putting resources to their best alternative use—that is, on counterfactual estimates of what people would be willing to pay for alternative products. For that reason, there can be no clear line between "objective" determinants of cost and "subjective" determinants of demand.

Opinions can differ as to what is the best alternative. For example, just what is the opportunity cost to the Martins of investing their $200,000 savings in their firm? Ralph might think that the best alternative use would have been to purchase a portfolio of blue-chip stocks paying a 5 percent rate of return. Andrea might think the best alternative use would have been to buy shares in an aggressive hedge fund—a riskier use of their savings, but one yielding an expected return of 15 percent. Who is to say which one is right? Which alternative use of the $200,000 is best depends not only on subjective estimates of the likely return from alternative investments but also on the subjective attitude toward risk of the person making the investment. Further, it's very possible that owning one's own business and being one's own boss provides subjective psychological benefits (or costs) that would be absent if the $200,000 were passively invested into a stock portfolio.

The same is true of the opportunity costs of resources other than capital. For example, an estimate of the opportunity cost of assigning a worker to one task must take into account not only the worker's pay but also what he or she could have contributed by doing another task instead. A manager will not always be able to measure the worker's productivity objectively in both tasks, so the decision will rest on a subjective judgment. In short, because choices are subjective, costs are subjective, too.

6.1c Profit, Rents, and Entrepreneurship

Pure economic profit, as we have defined it, is the difference between what a firm receives for the products it sells and the opportunity cost of producing them. We first introduced the notion of payments in excess of opportunity costs in Chapter 4, where we called them *economic rents*. Pure economic profit, then, is a type of economic rent, but the two terms are not fully interchangeable.

For one thing, economic rent is a broader notion than profit. We usually use *profit* when discussing the net income of business firms, but it is possible for any factor of production to earn *rents*. Consider, for example, the income of rock stars, sports professionals, and other people with exceptional talents. The opportunity cost of pursuing their chosen line of work may be low, in the sense that the income from their best-paying alternative occupation (say, selling insurance or working as a lifeguard) may be far lower than what they now earn. The amount by which their extraordinary income as a rock star, sports professional, or whatever exceeds their income from their best alternative occupation is economic rent, but we would not usually call it profit.

We can also make a distinction between *profit seeking* and *rent seeking*. Profit seeking is what entrepreneurs do when they look for ways to create goods and services that are worth more than the inputs they require. Henry Ford, Steve Jobs, and Sam Walton are examples of people who devoted their lives to finding new ways of satisfying customer needs. *Profit seeking*, then, means finding ways to create new value.

On the other hand, some firms try to increase their revenues by seeking restrictions on competition rather than through innovation and cost reduction. For example, US cotton and sugar farmers have increased their business earnings not so much by cutting costs or improving their products as by persuading Congress to restrict imports. We can best think of this as the result not of profit seeking but of political rent seeking.[2]

(Keeton Gale/Shutterstock)

The incomes of professional athletes often include a large share of economic rent.

The distinction between profits and rents is certainly not watertight. In both cases, we are dealing with revenues that exceed opportunity costs. Data like those presented in Figure 6–1 do not tell us all we might want to know about the origin of the $20,000 of pure economic profit. Did the Martins earn that $20,000 by creating a superior new product or by lobbying Congress to obstruct the efforts of competitors? It is not always possible to tell just by looking at the kind of cost and revenue data that we will deal with in this chapter and the one that follows.

6.1d Fixed Costs, Variable Costs, and Sunk Costs

The implicit-explicit distinction is not the only way to classify costs. Another approach focuses on the time horizon for production decisions.

The amounts of inputs a firm uses vary as the amount of output changes. It is possible to adjust the amount of some inputs quickly. For example, a firm can increase inputs of electricity instantly by flipping a switch. Certain raw materials and hourly labor are other examples of inputs that the firm may be able to adjust quickly. We call these **variable inputs** and the costs of obtaining them **variable costs**. Other inputs take longer to adjust. For example, building a new office building takes many months. In general, inputs that take longer to adjust are those that define the size of the firm's plant, including structures and production equipment. We call these **fixed inputs** and the cost of providing them **fixed costs**.

Which inputs are fixed and which are variable depends not so much on the physical properties of the inputs themselves as on the context of decisions about them. For example, a firm that hires workers on an hourly basis may treat wages as a variable cost. Another firm that hires workers on a yearly contract, subject to a "no layoff" agreement with a labor union, would treat wages as a fixed cost. Public policies may also affect which inputs represent fixed or variable costs. For example, in the United States, most firms follow the principle of *employment at will*, meaning that they can lay off workers whenever they think doing so would improve their profits. That makes wages a variable cost. In contrast, many European countries have strong labor protection laws that require advance notice, complex paperwork, and large severance payments when a firm lays off workers. That makes wages more of a fixed cost.

Fixed and variable inputs, in turn, are the basis for a distinction between two time horizons: the short run and the long run. These terms are operational concepts, not standardized periods of calendar time. The **short run** is a length of time over which the firm can vary output by using more or fewer of its variable inputs, but one that is too short for changes in the firm's fixed inputs, such as the size of its plant. The **long run** is a time horizon that is long enough for the firm to change its fixed as well as its variable inputs. For example, an automaker can vary output from month to month by adding extra shifts

Variable input
Input that can be varied within a short time in order to increase or decrease output

Variable cost
The explicit and implicit costs of providing variable inputs

Fixed input
Input that cannot be increased or decreased in a short time in order to increase or decrease output

Fixed cost
The explicit and implicit opportunity costs associated with providing fixed inputs

Short run
A time horizon within which a firm can adjust output only by changing the amounts of variable inputs it uses while fixed inputs remain unchanged

Long run
A time horizon that is long enough to permit changes in both fixed and variable inputs

of workers without adding or expanding plants. That is the short run. Over a period of a few years, the same firm can increase capacity to meet expected growth of demand by building new plants or expanding old ones. That is the long run.

Producer Surplus and Profit

In the last chapter, we introduced the idea of producer surplus—the counterpart of consumer surplus that shows how much producers benefit from participating in the market. Now that we've discussed the difference between fixed and variable costs, as well as the difference between the short and long run, we can discuss the relationship between profit and producer surplus.

Remember that we defined producer surplus as the difference between what producers receive for a unit of a good and the minimum that they would be willing to accept. In the short run, some costs are fixed, which means they cannot be avoided. So, in the short run, producers only consider variable costs when deciding whether to produce. For example, suppose you are a private-practice lactation consultant and that you are required to carry malpractice insurance and have your own scale for weighing babies. The costs of the insurance and the scale are fixed costs, say $100 per year for the insurance and $900 for the scale. When deciding whether to take on another client, you ignore the costs of the insurance and the scale, because those costs do not disappear if you turn down the new client. What you will consider are your variable costs: the printed materials you may leave behind, the non-reusable medical devices you will use during the consult, and, most important, your time.

(Shutterstock)

An accurate baby scale is a necessary fixed cost for a lactation consultant.

When we think about the minimum that a producer would be willing to accept, we have to remember that a producer will only consider the costs that can be avoided. In the short run, this means that the minimum price that a producer is willing to accept is that price that will just cover the variable costs. In the calculation of producer surplus, then, only variable costs are subtracted from revenue. This means that in the short run, when some costs are fixed, producer surplus will be greater than profit by the amount of the fixed costs. Given this, it is possible for producer surplus to be positive (participating in the market made the producer better off) even when economic profit is negative (there are better opportunities): When this occurs, a producer has an incentive to stop producing and leave the market in the long run, but in the short run it is better to produce than not to. We will return to this scenario in a later chapter.

Implicit and Explicit Fixed Costs

In all cases, "cost" means opportunity cost, including both implicit and explicit costs. However, *implicit fixed costs* deserve special attention.

Fixed costs, by their nature, do not vary with the firm's rate of output. They must be borne by the firm as long as it stays in business, regardless of how much it produces in the short run. Some fixed costs, such as the malpractice insurance for a lactation consultant, take the form of periodic payments, which means they are explicit fixed costs. Implicit fixed costs are opportunity costs associated with facilities owned by the firm itself but not reflected in ongoing payments.


(Shutterstock)

A warehouse is a fixed cost for a trucking firm that cannot operate without it.

Consider a trucking firm that cannot operate without a warehouse. The warehouse is a fixed cost because—within wide limits, at least—the firm needs it regardless of how much freight it hauls in a given month. The cost of the warehouse could be either explicit or implicit depending on who owns it. The firm might lease the warehouse for an annual payment of $12,000 paid in installments of $1,000 per month. The lease payments would be an explicit fixed cost. Instead, it might use cash reserves to buy the warehouse for a price of $120,000. The $120,000 used to buy the warehouse could have been used for some other purpose—say, to buy securities yielding 10 percent annual interest. The income (10 percent of $120,000, which is $12,000 a year or $1,000 a month) that could have been earned with those funds if they had not been used to buy the warehouse is an opportunity cost of owning the warehouse—an implicit fixed cost.

Whether the warehouse cost is explicit (as under a lease) or implicit (as under ownership), it continues as long as the firm stays in business, even if it goes a month without carrying any freight at all. The same is true for the lactation consultant who buys a $900 baby scale for her business. Though smaller than a warehouse, the baby scale still carries with it an implicit fixed cost, because the $900 could have been invested elsewhere. And the lactation consultant bears this cost whether she sees clients or not, unless she decides to shut down her practice and sell the scale.

Sunk Costs

It is important not to confuse fixed costs, especially implicit fixed costs, with **sunk costs**. Not all fixed costs are sunk. Sunk costs reflect once-and-for-all expenditures that a firm cannot recover even if it leaves its line of business. For example, the trucking firm in our example might have paid $1,000 to have "Taylor Trucking" painted on the wall of its warehouse. That is a sunk cost. If the firm sells the warehouse (or terminates its lease), the sign becomes worthless. There is no way to recover the $1,000 because the next owner or tenant will want a different sign.

Are sunk costs opportunity costs? That depends on circumstances. If a firm is entering a new line of business or expanding its operations, the sunk costs of doing so are an opportunity cost. For example, if our trucking firm is considering service to a new city, its owners must think, "The opportunity cost of entering the new market equals $120,000 to buy a warehouse plus $1,000 to paint the sign." On the other hand, once the firm has incurred a sunk cost, it is no longer an opportunity cost relevant to any future decision because the firm has, once and for all, lost the opportunity to do anything else with the funds in question. If the firm later considers abandoning service to the city in question, its owners should think only, "We could get $120,000 by selling the warehouse." The $1,000 paid for the sign is relevant before entering the market but irrelevant after.

The remainder of this chapter will be concerned only with firms' ongoing (nonsunk) fixed and variable costs of doing business. Sunk costs will come back into the picture in later chapters.

Sunk costs
Once-and-for-all costs that a firm cannot recover once it incurs them

6.2 Production and Costs in the Short Run

Now that we have pinned down the meaning of cost, our next job is to build a theory to explain how a firm's costs vary with its level of output. A firm's cost of production is one of the basic constraints that shape a firm's decisions (along with demand, which we will take up in the next chapter). We will divide the discussion of cost theory into two parts: short run and long run.

6.2a Production with One Variable Input in the Short Run

Although most firms can vary any of several inputs in the short run, we will keep things simple at the outset by considering the case in which there is only one variable input—the quantity of labor. To do so, we will turn once again to our imaginary firm, Fieldcom.

Figure 6–2 shows what happens to daily units of output, or **total physical product**, as the firm varies the number of workers from zero to eight. If it employs no one, or just one worker, it cannot produce anything because some parts of the job require at least two people working together. Two workers can get production moving; but because they use a lot of time setting up jobs and changing from one job to another, they are able to produce at a rate of only one smartphone per day. When a third worker joins them, specialization becomes possible and production increases to three units a day. A fourth worker gets things moving really smoothly, and output goes up to seven units a day. Adding workers five, six, and seven boosts the plant's output to its maximum of thirteen units a day. Beyond that point, it does no good to add more workers; all the tools and equipment are in use, and any extra workers would have to stand around waiting for a turn to use them.

In practice, the firm could increase output beyond thirteen units per day by adding other inputs along with more workers—more assembly tables, more testing equipment, and so on. And it may decide to do so in the long run. For the moment, however, we are looking at the short-run effects of increasing labor alone, other things being equal.

Marginal Physical Product

The chart in Panel (b) and columns (1) and (2) in Panel (a) of Figure 6–2 show the relationship between labor inputs and daily output. In the range of one to seven workers, output rises as labor input increases, but not at a constant rate. Column (3) of the table and the chart in Panel (c) of the figure show how much output changes for each added worker. We call the change in output produced by an added unit of a variable input the **marginal physical product** of that input. (As elsewhere, the adjective *marginal* refers to the effect of a small change in a quantity.)

Total physical product
The total output of a firm, measured in physical units

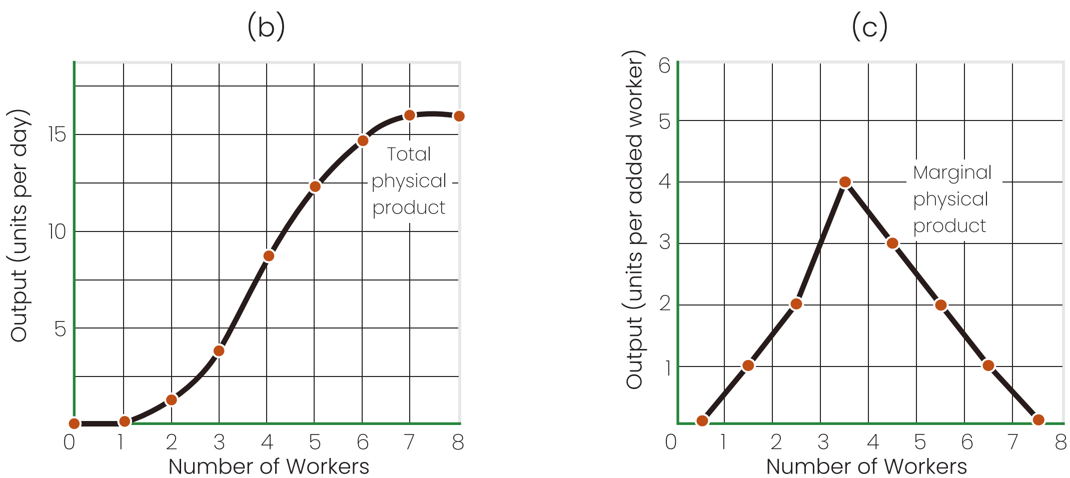
Marginal physical product
The amount by which output, expressed in physical units, increases as a result of adding one unit of a variable input, other things being equal

In our example, the marginal physical product is one unit of output when labor input increases from one unit to two; from two to three workers, marginal physical product rises to two units; and so on. The increase in staff from three workers to four gives the greatest boost to output. After that, output increases at a diminishing rate with each added worker. Once the staff reaches seven workers, the marginal physical product drops to zero.

Figure 6–2 Response of Output to Changes in One Variable Input

(a)

(1) Input (workers per day)	(2) Total Physical Product (units per day)	(3) Marginal Physical Product (units per worker)
0	0	
1	0	0
2	1	1
3	3	2
4	7	4
5	10	3
6	12	2
7	13	1
8	14	0



This figure shows how Fieldcom Inc. responds to changes in labor inputs. All other inputs remain constant while the number of workers is varied. One worker can produce nothing. After that, output increases as more workers are used. After seven workers are on the job, more workers add nothing to output. Column 3 of Panel (a) and the chart in Panel (c) show the amount of output added by each worker, a quantity we call the marginal physical product of the variable input.

The Law of Diminishing Returns

Our example shows a pattern that economists consider typical for the marginal product of a single-variable input. At first, marginal product increases as the firm adds workers. Increasing marginal product reflects the advantages of cooperation, the superiority of team production, and the benefits of specialization by comparative advantage. After a point, however, marginal product stops rising and begins to fall. In the case of a single-variable input, the principal reason for the eventual decline in marginal physical product is the overcrowding of complementary fixed inputs—in our example, such things as workspace, tools, and testing equipment.

Panel (c) of Figure 6–2 uses a graph called the *marginal physical product* curve to show the relationship of marginal physical product to the number of workers. The part of the curve with a negative slope illustrates a principle known as the **law of diminishing returns**. According to that principle, as the amount of one variable input increases while the amounts of all other inputs remain fixed, the firm will eventually reach a point beyond which the marginal physical product of the input will decrease.

Law of diminishing returns
The principle that, as one variable input increases while all others remain fixed, a firm will eventually reach a point beyond which the marginal physical product of the variable input will begin to decrease

The law of diminishing returns applies to all production processes and to all variable inputs. The example just given comes from manufacturing; however, we could illustrate the law just as well with farming, using fertilizer as the variable input. Applying more fertilizer increases output, but beyond some point the gain in output from one more ton of fertilizer tapers off. (Too much fertilizer could even poison the plants, in which case marginal physical product would become negative.) Oil refineries, power plants, barber shops, government bureaus—indeed, *all* production processes—are subject to the law of diminishing returns.

6.2b From Marginal Physical Product to Marginal Costs

The relationship between inputs and outputs in terms of physical units is an important constraint on a firm's profit-maximizing activities. However, many business decisions focus not on physical units but on money. Our next step, then, is to restate the constraint information given in the marginal physical product curve in terms of money; we can do this by asking how much each added unit of output costs.

We call the change in cost associated with a one-unit change in output **marginal cost**. Here is how we can make the transition from marginal physical product to marginal cost, still using the Fieldcom example: First, we rearrange the data given in Figure 6–2 in terms of input per unit of output, as in Figure 6–3. The table in Panel (a) of Figure 6–3 reverses the order of the first two columns of the table in Panel (a) of Figure 6–2 and gives the minimum number of workers required to produce each level of output. Also, we have flipped around the charts in Panels (b) and (c) so that they now have units of output, rather than units of labor input, on the horizontal axis.

The next step is to convert physical units of input into costs. To do so, we need to know the cost per unit of input. For our example, we will assume that the cost of hiring one worker is $100 per day. Multiplying the labor inputs in column (2) of Figure 6–3 by the $100-per-day wage yields total labor costs, which appear in column (3). Those data then form the basis for the total labor cost curve in Panel (b) of the figure. Taking the rearrangement of the axes and the change in units into account, we can recognize the total labor cost curve as the mirror image of the total physical product curve.

Finally, column (4) of the table in Figure 6–3 shows marginal cost—that is, the change in cost for each added unit of output, calculated as the increase in labor costs divided by the increase in output. Increasing output from zero to one requires adding two workers, so the added cost per unit in that range is $200; increasing output by two more units (from one to three) requires one more worker, so the cost per added unit of output in that range is $50; and so on. The marginal cost curve shown in Panel (c) of the figure comes from columns (1) and (4) of the table. Again, considering the change in units and rearrangement of the axes, Panel (c) of Figure 6–3 looks much like a mirror image of the marginal physical product curve that was shown in Panel (c) of Figure 6–2.

Marginal cost (MC)
The increase in cost required to increase the output of some good or service by one unit

Figure 6–3 Cost and Output with One Variable Input

(a)

(1) Output (units per day)	(2) Labor Input	(3) Total Labor Cost (dollars per day)	(4) Marginal Cost (dollars per unit)
0	0	0	
			200
1	2	200	
			50
3	3	300	
			25
7	4	400	
			33
10	5	500	
			50
12	6	600	
			100
13	7	700	

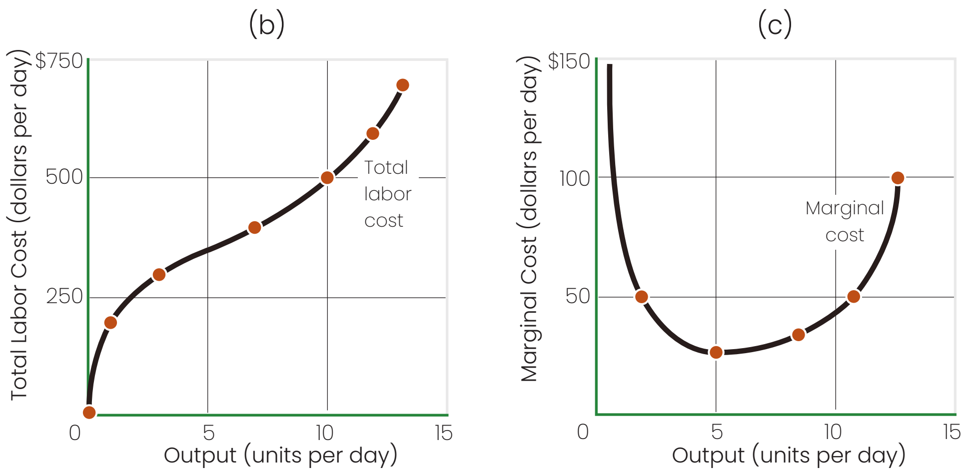


Figure 6–3 shows how the cost of production at Fieldcom Inc. changes as output varies. The table and graphs use data from Figure 6–2; however, here we recast them to stress cost, assuming a daily wage of $100 per worker. Column (3) of the table and the chart in Panel (b) show total labor cost for various output levels. Column (4) of the table and the chart in Panel (c) show marginal cost—the amount by which cost increases per added unit of output. For example, increasing the number of workers from three to four raises output by four units, from three units to seven per day. Over this range, then, the cost of each added unit is one-fourth of a day's wage, or $25. (Note that $25 is plotted halfway between three units of output and seven units of output, at five units of output.)

6.2c More Than One Variable Input

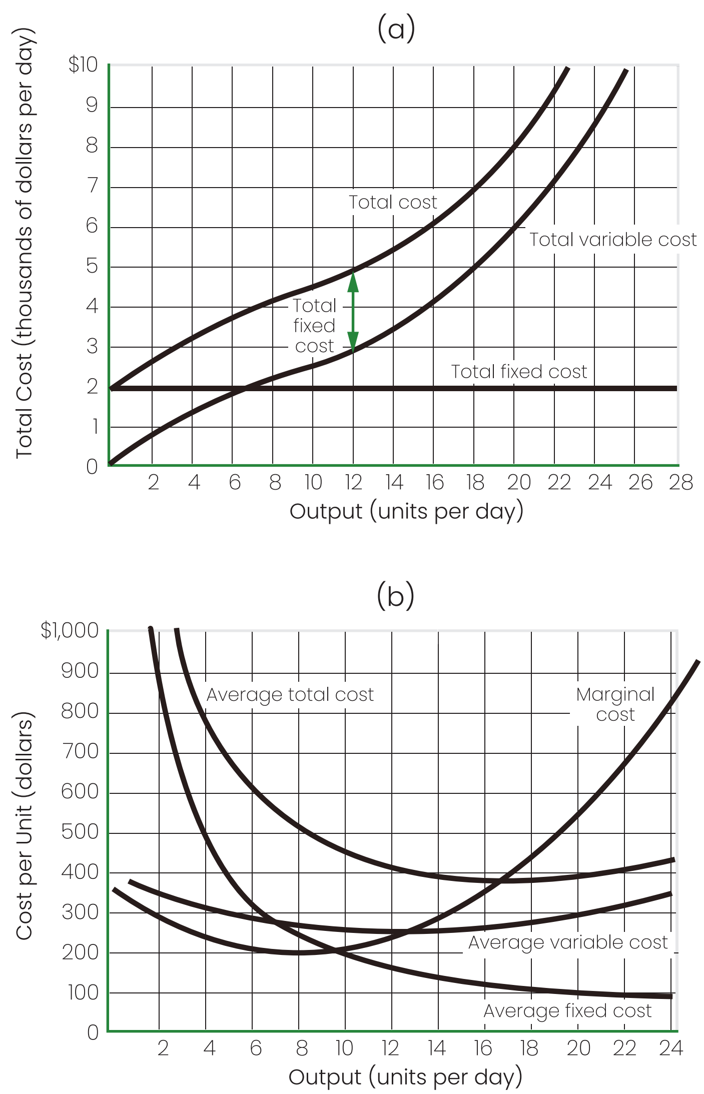
The Fieldcom example assumes that only one input is varied. In practice, short-run increases or decreases in output often involve changes in many—although not all—inputs. For example, if Fieldcom wanted to raise its output, it is possible that it would not only have to hire more workers but also use more fuel to keep the shop heated longer each day and double the rate at which it orders parts.

The appendix to this chapter outlines a way of analyzing changes in two or more variable inputs. Without going into detail, we can say that, as long as at least some inputs remain fixed, the law of diminishing returns continues to apply. Also, a region of increasing marginal physical product will often exist at low levels of output. Together, those features tend to give total cost curves with a reverse-S shape and marginal cost curves with a U shape—even when there is more than one variable input—just as in the case of the simpler one-input example.

6.2d A Set of Short-Run Cost Curves

Figure 6–4 shows a complete set of short-run cost curves for the firm in our example, together with a table showing the data used in drawing the curves. Panel (a) of the figure shows three total cost curves. The first is a total variable cost curve similar to the one shown earlier in Figure 6–3, but it is based on the data in column (2) of the table (c) in Figure 6–4. The second curve shows total fixed costs, from column (3) of the table (c) in Figure 6–4, which are $2,000 per day. These include all the costs of office staff, testing equipment, rent, and so on that are the same regardless of the firm's level of output. Adding columns (2) and (3) gives total cost, in column (4), which corresponds to the third curve in Panel (a). The total cost and total variable cost curves are parallel. The vertical distance between them equals total fixed cost.

Figure 6–4 A Set of Short-Run Cost Curves



(continues)

Figure 6–4 A Set of Short-Run Cost Curves *(continued)*

(c)

(1) Quantity of Output (units per day)	(2) Total Variable Cost (dollars per day)	(3) Total Fixed Cost (dollars per day)	(4) Total Cost (dollars per day)	(5) Marginal Cost (dollars per day)	(6) Average Variable Cost (dollars per day)	(7) Average Fixed Cost (dollars per day)	(8) Average Total Cost (dollars per day)
0	$ 0	$ 2,000	$ 2,000		—	—	—
1	380	2,000	2,380	$ 380	$ 380	$ 2,000	$ 2,380
2	720	2,000	2,720	340	360	1,000	1,360
3	1,025	2,000	3,025	305	342	667	1,008
4	1,300	2,000	3,300	275	325	500	825
5	1,550	2,000	3,550	250	310	400	710
6	1,780	2,000	3,780	230	297	333	630
7	1,995	2,000	3,995	215	285	286	571
8	2,200	2,000	4,200	205	275	250	525
9	2,400	2,000	4,400	200	267	222	489
10	2,605	2,000	4,605	205	261	200	461
11	2,820	2,000	4,820	215	256	182	438
12	3,050	2,000	5,050	230	254	167	421
13	3,300	2,000	5,300	250	254	154	408
14	3,575	2,000	5,575	275	255	143	398
15	3,880	2,000	5,880	305	259	133	392
16	4,220	2,000	6,220	340	264	125	389
17	4,600	2,000	6,600	380	271	118	388
18	5,025	2,000	7,025	425	279	111	390
19	5,500	2,000	7,500	475	289	105	395
20	6,030	2,000	8,030	530	302	100	402
21	6,620	2,000	8,620	590	315	95	410
22	7,275	2,000	9,275	655	331	91	422
23	8,000	2,000	10,000	725	348	87	435
24	8,800	2,000	10,800	800	367	83	450

(d)

Common Abbreviations		Useful Formulas
Q	Quantity of output	TC = TFC + TVC
TC	Total cost	$MC = \frac{\text{Change in TC}}{\text{Change in Q}} = \frac{\text{Change in TVC}}{\text{Change in Q}}$
TFC	Total fixed cost	
TVC	Total variable cost	$AVC = \frac{TVC}{Q}$
MC	Marginal cost	
AVC	Average variable cost	$AFC = \frac{TFC}{Q}$
AFC	Average fixed cost	
ATC	Average total cost	$AFC = \frac{TC}{Q}$

We can derive a whole set of short-run cost curves from data on fixed and variable costs, as this figure shows. The figure presents the data in the form of a table and a pair of graphs. The figure also lists a number of useful abbreviations and formulas.

6.2e Some Geometric Relationships

Parnels (a) and (b) of Figure 6–4 illustrate some important geometric relationships among the cost curves. First, notice that the minimum point of the marginal cost (MC) curve in Panel (b) corresponds to the inflection point of the total variable cost (TVC) curve in Panel (a)—that is, the quantity of output where it stops flattening out and starts getting steeper. That relationship holds because the slope of the TVC—that is, the rate at which total variable cost changes as output changes—is equal to marginal cost. Another way to express the relationship is to say that the height of the MC curve equals the slope of the TVC curve. So the slope of the TVC curve is at its *lowest* when the MC curve (which shows the slope of the TVC curve) reaches its *minimum*.

Next, notice that the MC curve intersects both the average variable cost (AVC) curve and the average total cost (ATC) curve at their lowest points. That relationship reflects the **marginal-average rule**. To understand the rule, ask what the cost of making one more unit of output will be, starting from any point. The answer is equal to the marginal cost of that unit. Then ask whether that cost is more or less than the average cost of all units produced up to that point. If the added cost of the next unit produced is *less* than the average cost of all the previous units, then producing it will have the effect of pulling *down* the average. If the next unit costs *more*, then producing it will pull the average *up*. It follows that whenever marginal cost is below average variable cost, the AVC curve must be falling (negatively sloped), and whenever marginal cost is above average variable cost, the AVC curve must be rising (positively sloped). That, in turn, implies that the MC curve cuts the AVC curve at its lowest point. The same reasoning holds for the relationship between marginal cost and average total cost.

The marginal-average rule is not unique to economics; we encounter it in many everyday situations. Consider, for example, the effect of your grade in this course on your grade point average (GPA). You could call your econ grade your "marginal grade" because it represents the additional grade points earned by taking this particular course. If your econ grade is higher than current GPA, it will pull your average up. If you do worse than average in this course, your GPA will fall. The relationship between your marginal grade and your cumulative GPA is the same as that between marginal cost and average cost.

6.3 Long-Run Costs and Economies of Scale

Up to this point, we have focused on short-run decisions—for example, how many smartphones to produce in Fieldcom's workshop or how much corn to grow on a certain piece of land. These short-run decisions involve changes in variable inputs only. They correspond to movements along a firm's short-run cost curves.

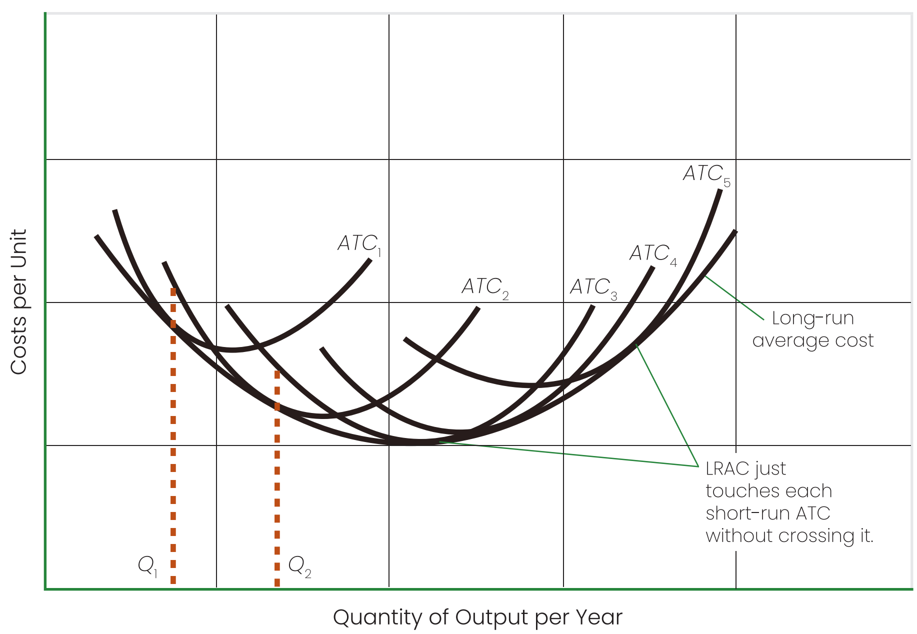
Now we turn our attention to decisions regarding a lasting expansion or contraction of the firm's stock of fixed inputs. For example, farmers might adjust to new subsidies for corn-based ethanol by buying additional land or farm equipment. Such decisions bring about changes in long-run costs. For the time being, we consider only fixed costs that are recoverable in the event that the firm leaves its line of business or permanently scales back its operations. We assume that there are no sunk costs.

Marginal-average rule
The rule that marginal cost must equal average cost when average cost is at its minimum

6.3a Planning for Expansion

Put yourself in the position of an entrepreneur about to set up a small firm. Like many start-ups, you are going to start small, working out of your own garage; but you want to do some long-range planning, too. In consultation with specialists, you put together plans for plants of five possible sizes, each of which could represent a stage in the future growth of your firm. Figure 6–5 shows short-run ATC curves for each of the plants. The first one, ATC_1, shows short-run average total costs at various outputs that you could produce in your first plant, the garage. The second curve corresponds to a slightly larger plant, and so on.

Figure 6–5 Short- and Long-Run Average Cost Curves



Short-run average total costs, and the position of the short-run ATC curves, depend on the size of the plant. In the long run, the firm has a choice of operating with any size of plant it chooses. Each plant size corresponds to a different U-shaped, short-run ATC curve. This graph shows five such curves. A new firm might begin in the owner's garage, corresponding to short-run average total cost curve ATC_1. Then, as demand for its product expands, the firm might move to one of the ATC curves farther to the right. Its long-run average cost (LRAC) curve is the "envelope" of these and other possible short-run ATC curves; that is, it is a smooth curve drawn so that it just touches the short-run curves without intersecting any of them.

Choosing a plant of a certain size does not commit a firm to that plant size forever, but the choice is not a trivial one, either. A small firm cannot afford to take on the costs of a permanently larger plant just to fill a single order. It will not make sense to expand the size of your plant unless you can spread fixed costs over a large enough total output. For example, in the 1950s, Sony Corporation was a tiny firm just starting to produce transistor radios. A buyer for a large American retail chain asked for prices on quantities ranging from ten thousand to one hundred thousand radios. Akio Morita, Sony's chairman, surprised the buyer by giving a higher price per unit on the larger order. He explained that one hundred thousand units exceeded the company's plant capacity. It would be a big risk to invest in a larger plant just for one large order. The buyer ended up placing an order for ten thousand radios, which was just right for Sony at the time.[3] The moral of the story: Only when the firm expects a sustained long-term increase in its output should it move from one of the short-run curves shown in Figure 6–5 to the next.

The five short-run cost curves in the figure represent only a sample of possible plant sizes. Taking into account the short-run curves that correspond to plants of sizes between those in the figure, we can draw a *long-run average cost (LRAC) curve* as the "envelope" of all the possible short-run average cost curves. By that, we mean the LRAC curve just touches each of the possible short-run curves without crossing them. The optimal size of plant for any given level of output in the long run will have a short-run ATC curve that is just tangent to the long-run ATC curve at the chosen level of output. Stated differently, we could say that on every possible ATC curve, there is a point where that ATC curve is lower than all of the others; it is these points that are on the LRAC curve.

It may be possible to produce a given level of output in a plant larger or smaller than the optimal one, but doing so would carry a penalty in terms of cost per unit. For example, in Figure 6–5 the firm can produce output level Q_1 at least cost in a plant of the size corresponding to the short-run curve ATC_1. It could instead produce the same level of output in the larger plant corresponding to ATC_2, but only at a higher cost per unit. On the other hand, the larger plant shown by ATC_2 is the best plant size for output Q_2. Producing that larger quantity of output in the smaller plant would mean running it above its design capacity. The cost penalty of doing so is evident from the fact that ATC_1 lies above ATC_2 at the output level Q_2.

If a firm wants to produce at an unusually high or low rate for a short time, it may make sense to do so by moving along the short-run ATC curve corresponding to its present plant size. An example would be a firm that decides to run overtime to fill an exceptionally large order, or one that cuts back to half-shifts to weather a temporary business downturn. When sustained increases in output level are under consideration, a firm minimizes costs by building a larger plant. Likewise, a firm that is planning to reduce its output permanently will eliminate or downsize its plant rather than keep production facilities operating at lower levels of output than those for which they were designed. Decisions of that kind represent movements along the firm's LRAC curve.

Economies of scale
A situation in which long-run average cost decreases as output increases

Diseconomies of scale
A situation in which long-run average cost increases as output increases

Constant returns to scale
A situation in which there are neither economies nor diseconomies of scale

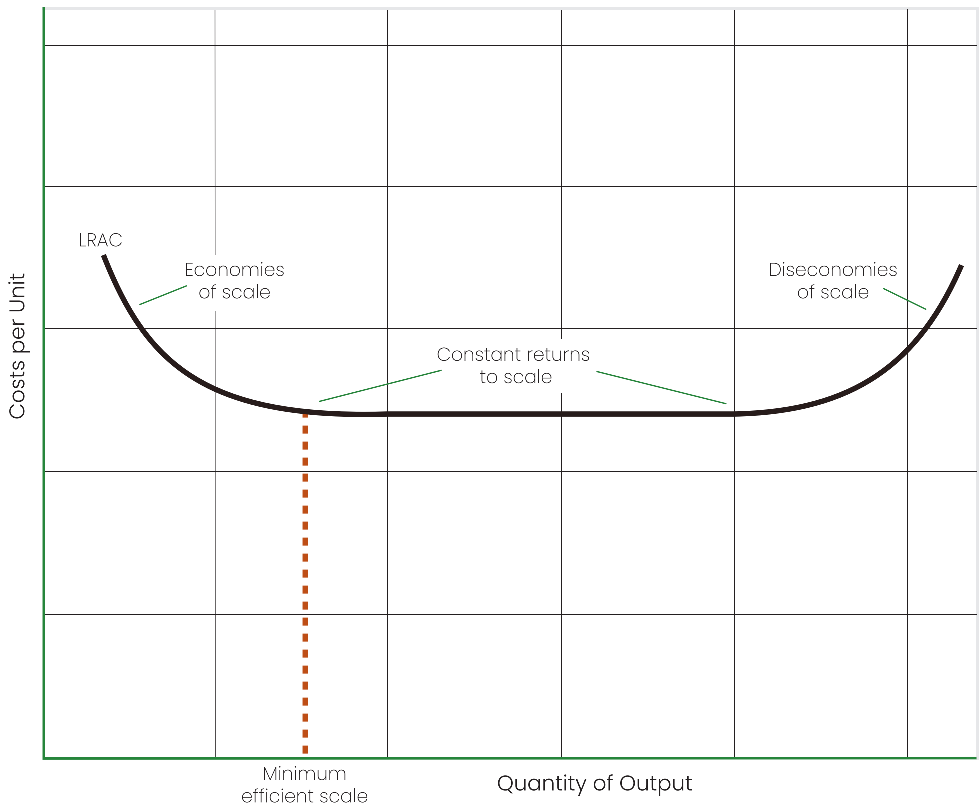
Minimum efficient scale
The output level at which economies of scale cease

6.3b Economies of Scale

We refer to movements along a firm's LRAC curve, during which it is free to adjust quantities of all the inputs it uses, as changes in the *scale* of production. Some special terminology applies to the way long-run average cost changes as the scale of production changes. In any output range in which long-run average cost *decreases* as output increases, the firm experiences **economies of scale**. In any output range in which long-run average cost *increases* as output increases, the firm experiences **diseconomies of scale**. Finally, if there is any range of output for which long-run average cost does not change as output varies, the firm experiences **constant returns to scale** in that range.

The LRAC curve in Figure 6–5 is smoothly U-shaped, so there is no range of constant returns to scale. However, empirical studies suggest that the LRAC curves of actual firms may have long flat sections in a middle range of output over which average cost changes little as output changes, as shown in Figure 6–6. Economies of scale for such a firm appear only at very low outputs, and diseconomies appear only at very high outputs. For a firm with such an LRAC curve, the level of output at which economies of scale end and constant returns to scale begin is called the firm's **minimum efficient scale**.

Figure 6–6 Long-Run Average Costs and Economies of Scale



When long-run average cost decreases as output increases, the firm experiences economies of scale. When long-run average cost increases as output increases, the firm experiences diseconomies of scale. In the range of output for which long-run average cost does not change as output varies, the firm experiences constant returns to scale. The level of output at which economies of scale end and constant returns to scale begin is called the firm's minimum efficient scale.

Sources of Economies of Scale

Where do economies of scale come from? If firms grew simply by increasing fixed and variable inputs in exact proportion, so that a large plant amounted to nothing more than a lot of small plants built side by side, we might expect changes in scale to have no effect on average cost. That is not the way firms expand, however. As firms grow, they tend to change the technologies they use and their methods of internal organization to take advantage of new opportunities offered by higher output levels.

In part, economies of scale stem from human factors like the advantages of team production and specialization according to comparative advantage. In a small firm, for example, the marketing function may be something the owner does from 3:00 p.m. to 4:00 p.m., after touring the plant floor and perhaps taking a turn running a machine. A somewhat larger firm can afford to hire a marketing manager who devotes full time to the job. In a still larger firm, subspecialties develop—a sales manager, a director of market research, an advertising specialist—all under the direction of the marketing manager.

Other economies of scale have origins in technology. In many lines of production, for example, a larger machine that is capable of doing twice the work of a smaller one costs *less than twice as much* to build and operate. For a firm that is too small to make full use of a large piece of equipment, the smaller model can be the appropriate choice. As the firm grows, technological economies lower its average costs.

Growth of a firm does not just mean constant expansion of a single plant. Operation of multiple plants can yield further economies of scale even after each plant reaches the minimum efficient scale. McDonald's provides an example. The minimum efficient scale for a single plant (restaurant) is very low in the fast-food industry. Yet McDonald's gains some important economies by running a large number of restaurants as a system: Individual food items and ingredients can be made in central kitchens and shipped to individual locations, managers can be trained together at "Hamburger University," and so on. A multiplant firm such as McDonald's also realizes economies of scale in such areas as finance and marketing.

Sources of Diseconomies of Scale

Sources of economies of scale are not limitless. As a firm expands, it encounters *diseconomies* of scale as well.

The most important diseconomies of scale are organizational. As a firm grows, it finds itself depending more and more on hierarchical means of coordinating its employees' activities. As a hierarchy grows, the cost of channeling information to key decision-makers tends to rise. Moreover, individual incentives become hard to maintain in a large hierarchical organization. More and more managerial skill has to be devoted to employee loyalty and motivation. There is an increasing risk that departments and divisions will pursue parochial interests that diverge from those of the firm as a whole.

In some lines of business, firms can grow to a very large size before the diseconomies start to outweigh the economies. Huge firms—such as Walmart, Toyota, and Exxon Mobil—successfully manage hierarchies that are bigger than the governments of many countries.

In other lines of business, comparatively small firms seem to have the edge. In farming, services, and many sectors of retail trade, small units predominate. Still other industries use franchising to combine economies of scale for a few functions, such as marketing and product development, with the operating flexibility of small units.

Summary

1. How do economists view the concepts of cost and profit?

Explicit costs are opportunity costs that take the form of explicit payments to suppliers of factors of production and intermediate goods. *Implicit costs* are the opportunity costs associated with using resources contributed by the firm's owners (or owned by the firm itself as a legal entity) that are not obtained under contracts calling for explicit payments. Implicit costs include the opportunity cost of capital needed to attract owners' capital to the firm. Revenue minus explicit costs gives accounting profit. Revenue minus all costs, both implicit and explicit, gives *pure economic profit*.

2. What is the distinction between short-run and long-run time horizons?

Fixed inputs cannot be increased or decreased in a short time. We call the costs of those inputs *fixed costs*. *Variable inputs* are those that firms can add or reduce quickly in order to increase or decrease output; they include hourly labor, energy, and raw materials. Those inputs give rise to *variable costs*. *Sunk costs* are once-and-for-all expenditures that a firm cannot recover once it has made them. The *short run* is a period within which a firm can adjust only variable inputs. In the *long run,* a firm can make changes in fixed inputs, thereby changing its plant size.

3. How do costs vary in response to changes in the quantity of a variable input?

When the amount of one input to a production process increases while the amounts of all other inputs remain fixed, output will increase, at least over some range. The amount that each one-unit increase in the variable input adds to output is the *marginal physical product* of that input. According to the *law of diminishing returns*, as the amount of one variable input increases (with the amounts of all other inputs remaining fixed), beyond some point, the amount of output added per unit of added variable input (that is, the marginal physical product of the variable input) will begin to decrease.

4. How can we represent a firm's cost structure in geometric terms?

We can construct a whole set of cost curves for a firm from data on its fixed and variable costs: total cost, total fixed cost, total variable cost, average fixed cost, average variable cost, average total cost, and marginal cost. According to the *marginal-average rule*, the marginal cost curve intersects the average variable cost and average total cost curves at their lowest points.

5. What choices does a firm face in the course of long-run expansion?

In the long run, a firm can adjust the amounts of fixed inputs that it uses by expanding or reducing its plant. Each possible plant size has a U-shaped short-run average total cost curve. The firm's long-run average cost curve is the envelope of its short-run curves. When long-run average cost decreases as output increases, the firm experiences *economies of scale*. When long-run average cost increases as output increases, it experiences *diseconomies of scale*. If there are neither economies nor diseconomies of scale, the firm has *constant returns to scale*.

Key Terms

Accounting profit	148
Constant returns to scale	162
Diseconomies of scale	162
Economies of scale	162
Expansion path	175
Explicit costs	147
Fixed cost	151
Fixed input	151
Implicit costs	147
Isoquantity line (isoquant)	171
Law of diminishing returns	155
Long run	151
Marginal-average rule	160
Marginal cost (MC)	156
Marginal physical product	154
Minimum efficient scale	162
Normal profit	148
Pure economic profit	148
Short run	151
Sunk costs	153
Total physical product	154
Variable cost	151
Variable input	151

Problems and Topics for Discussion

1. **Entrepreneurship and risk**
   One of the opportunity costs borne by anyone who starts a new business, whether it is Akio Morita of Sony or our imaginary Ralph and Andrea Martin, is that of exchanging the secure life of employees of large firms for the risky life of entrepreneurs. Would you be willing to make that transition if you expected to earn no more than your previous salaries plus a "normal profit" on the capital you invested in your firm? Would you require some pure economic profit as compensation for the risks and responsibilities of being an entrepreneur? Or would the freedom of running your own business be so attractive you would do it even if your total income were less than what you could earn working for someone else? Discuss.

2. **Implicit and explicit costs**
   List the basic costs of owning and operating an automobile. Which are explicit costs? Which are implicit costs? Does driving an automobile impose any external costs on the economy as a whole that do not show up on your list as either implicit or explicit costs? If so, what are they?

3. **Fixed and variable costs**
   Divide the costs of owning and operating an automobile into fixed and variable costs. Suppose that you were deciding whether to drive to a football game at a nearby college or to take the bus instead. Would you consider both fixed and variable costs? Suppose that you were deciding whether to buy a house in a neighborhood where you could walk to work or a house in a neighborhood where you would have to buy a second car to drive to work every day. Would you consider both fixed and variable costs of the second car? Explain the difference between the two situations.

4. **Economies and diseconomies of scale**
   Do you think the business of running a college is subject to economies or diseconomies of scale? Which parts of the college's operation (such as library, dormitories, faculty salaries, moving students between classes, and so on) are subject to economies of scale, diseconomies of scale, or constant returns to scale?

5. **Total cost curves**

Draw a set of coordinate axes on a piece of graph paper. Label the x axis "Output" (0 to 20 units) and the y axis "Cost" (0 to 20 units). Plot the following (x, y) points on your graph: $(0, 4)$; $(2, 6)$; $(4, 7)$; $(7, 8)$; $(9, 9)$; $(11, 11)$; $(13, 14)$. Connect these points with a smooth curve and label it "Total Cost." Working from this curve, construct a total fixed cost curve and a total variable cost curve for the same firm.

6. **Marginal and average cost curves**

Draw a second set of coordinate axes on another piece of graph paper. Label the horizontal axis "Output" (0 to 20 units) and the vertical axis "Cost per Unit" (0 to 2 units, in tenths of a unit). Using as a basis the total cost, total variable cost, and total fixed cost curves you drew for Problem 5, construct the following curves on your new graph: marginal cost, average total cost, average variable cost, and average fixed cost.

7. **Relating the long- and short-run cost curves**

Turn to Figure 6–5 and copy the diagram onto a sheet of graph paper, drawing the long-run average total cost curve and one of the short-run average total cost curves. Use these curves to construct the corresponding long- and short-run *total* cost curves. Both total cost curves should be reverse-S shaped and tangent to each other at the same output level for which the average total cost curves are tangent.

8. **Diminishing returns**

Suppose that you examine the relationship between the amount of coal burned per week in a certain power plant and the amount of electricity generated per week. You find that for small amounts of coal—too small even to bring the boiler up to the temperature needed to make steam—no electricity can be produced. After burning a certain minimum amount of coal, the plant begins to operate. From that point on, the added amount of electricity generated per added ton of coal burned is constant over a wide range. Then after a point, burning more coal produces no more electricity. Sketch the total physical product curve for this plant, and draw a graph showing how marginal physical product varies as output changes. Does this production process obey the law of diminishing returns?

9. **More on diminishing returns**

"If not for the law of diminishing returns, all the food that the world needs could be grown in a flowerpot." Do you agree, disagree, or agree in part? Suggestion: Think of land as the only fixed factor and fertilizer as the only variable factor. How much food could you grow in the flowerpot if the marginal physical product of fertilizer were constant regardless of the amount per unit of land?

Case for Discussion

Tennis at the Grand Slam

The Grand Slam Sport and Health Club is a large, modern facility in the suburbs of a medium-sized American city. The club offers many activities, including swimming, weight training, and aerobics; but its leading attractions are its excellent indoor tennis courts. Members may play on clay or two types of hard-surface courts. To add to members' enjoyment, the club offers private and group lessons; tournament, ladder, and team competitions; and numerous other social events.

To join the club, a single individual pays a $1,000 nonrefundable initiation fee. In addition, there is an $88 monthly membership charge, which members must pay whether or not they use the facilities. Those two fees cover most of the club's costs, so it is able to keep the charge for actual playing time quite low. The fee for an hour's use of a court is only $2.

When it first started operation, the low hourly court fee created a problem for the club. The fee was so low that members would not bother to call to cancel a court reservation if they changed their minds about playing. Other members would then think that there was no space to play, when in fact the courts stood empty.

To overcome that problem, the club introduced a new rule: Members who make reservations and use the court pay the usual $2 per hour, but a member who makes a reservation and does not show up pays a penalty rate of $10 per hour for the unused time. A member may cancel a reservation nine hours or more in advance with no charge at all. The new rule has proved successful in reducing abuses of the reservation system and making court time more readily available to all members.

Questions

1. Classify the costs of membership in the Grand Slam as fixed, variable, and sunk.
2. Suppose that you are thinking about joining the Grand Slam to play indoor tennis. Which of the costs of membership are opportunity costs that would be relevant to your decision?
3. Suppose you are a member of the Grand Slam but are considering dropping your membership so that you can afford to do other things. Which of the costs of membership in the club are opportunity costs that would be relevant to your decision?
4. Suppose that you are a member of the club and are deciding whether to spend next Saturday afternoon playing tennis there. Which of the costs of membership are opportunity costs that would be relevant to your decision?
5. Suppose that it is noon on Saturday. You have made a reservation for an hour of court time at 5:00 p.m. A friend asks you to join a pickup basketball game at that time instead. What is your opportunity cost of abandoning the tennis reservation to join the basketball game? How would the answer be different if you had received the basketball invitation the afternoon of the day before your court reservation?

Endnotes

1 If you have studied accounting, you will recognize that this description of "accounting profit" is somewhat oversimplified. Accountants and economists have different objectives in analyzing the operations of a business firm. As a result, their concepts of costs and profits do not always allow precise comparison. Although the comparison is not exact, what economists call "accounting profit" most closely corresponds to what corporate accountants would call "net operating profit after taxes" (NOPAT) plus interest expense.

2 See James M. Buchanan, "Rent Seeking and Profit Seeking," in *Toward a Theory of the Rent-Seeking Society,* eds. James M. Buchanan, Robert D. Tollison, and Gordon Tullock (College Station: Texas A&M University Press, 1980), 3–15.

3 The anecdote is told by Akio Morita in "When Sony Was an Up and Comer," *Forbes,* October 6, 1986, 98–102.

Appendix to Chapter 6

Cost and Output with Two Variable Inputs

In the body of this chapter, we looked at the relationship between cost and output when just one input is varied and all other inputs are kept constant. Now, in this appendix, we extend the theory to the case of more than one variable input.

Substitution of Inputs

Having more than one variable input raises the possibility of substituting one input for another. Consider the case of Alex Hathaway, a farmer who grows corn. Hathaway works full time on the farm and does not hire any additional help. For Hathaway, the amount of labor used in growing corn is a fixed input, and the machinery they own is also a fixed input. In addition to those, there are two variable inputs: fertilizer, which is bought by the ton, and land, which is leased by the acre from a nearby landowner.

Hathaway can grow a given quantity of corn—say, two hundred bushels—in many different ways. Figure 6–7 shows some of the possibilities. One way to grow two hundred bushels of corn is to use 2.5 tons of fertilizer and 10.0 acres of land, corresponding to point P on the graph. If Hathaway wants to grow the same amount of corn on less land, the solution is to substitute fertilizer for land. For example, at point Q Hathaway can grow two hundred bushels of corn on 5.0 acres by using 5.0 tons of fertilizer. By substituting still more fertilizer for land, Hathaway can move to point R, where the two hundred bushels are grown on just 2.5 acres using 10.0 tons of fertilizer.

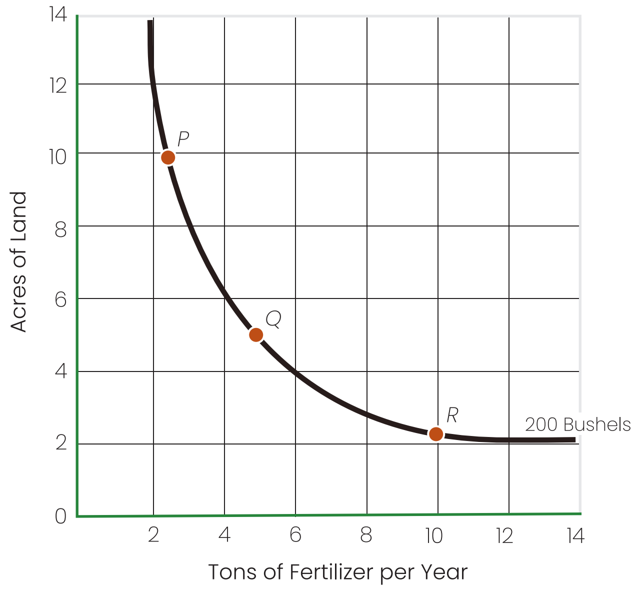
Diminishing Returns in Substitution

In this chapter, we defined the law of diminishing returns as it applies to a situation where one input varies while all others remain constant. In such a case, beyond some point, the amount of the variable input needed to make an extra unit of output increases. (That is another way of saying that the marginal physical product of the variable input decreases.) A similar principle applies when one input is substituted for another in such a way as to keep output at a constant level: As the amount of input x is increased, the amount of x needed to replace one unit of y increases.

Figure 6–7 illustrates this principle. In moving from point P to point Q, 2.5 tons of fertilizer replaces 5.0 acres of land while output stays constant at two hundred bushels. In moving from point Q to point R, however, we must apply 5.0 more tons of fertilizer to replace just 2.5 acres of land.

Because of diminishing returns in substituting one input for another, the curve connecting points P, Q, and R becomes flatter as one moves downward and to the right along it. That reflects the decreasing ratio of the marginal physical product of fertilizer to the marginal physical product of land as more fertilizer and less land are used.

Figure 6–7 An Isoquantity Line



This graph shows an isoquantity line, or isoquant, for the production of two hundred bushels of corn. The variable inputs are land and fertilizer; the other inputs, labor and machinery, are fixed. Points *P*, *Q*, and *R* represent various ways of growing the given quantity of corn. A movement downward along the isoquant represents the substitution of fertilizer for land while maintaining output at two hundred bushels per year. As more and more fertilizer is used and less land, the isoquant becomes flatter because of diminishing returns.

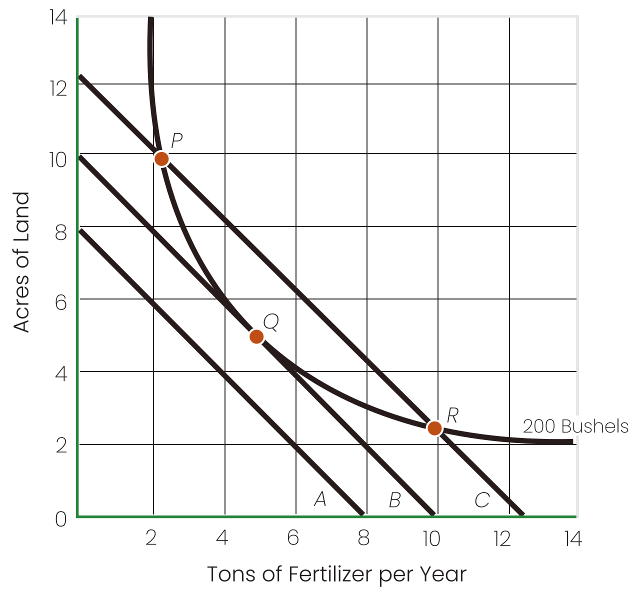
Choosing the Least-Cost Production Method

We call the line connecting points *P*, *Q*, and *R* in Figure 6–7 an **isoquantity line**, or *isoquant*, because it shows the combinations of inputs that are sufficient to produce a given amount of output. (The prefix *iso* comes from a Greek word meaning "equal.") Although all the points on the isoquant are equal in terms of output, they are not equal in terms of cost. To see how a producer can choose the least-cost method of producing a given level of output, we need to know the prices of the inputs.

In the appendix to Chapter 5, we used budget lines to indicate the prices of consumer goods. Figure 6–8 shows how we can use the same technique to represent the prices of inputs. The graph assumes a cost of $50 a ton for fertilizer and a rental price of $50 per acre per year for land. At those prices, $400 can buy 8.0 tons of fertilizer and no land, 8.0 acres of land with no fertilizer, or any of the other points on line *A*; $500 will buy 10.0 tons of fertilizer, 10.0 acres of land, or any of the other points on line *B*; and so on.

Isoquantity line (isoquant)
A line showing the various combinations of inputs that are sufficient to produce a given quantity of output

Figure 6–8 Finding the Least-Cost Production Method



This graph shows how to find the least-cost method of production on an isoquant, given the input prices. In this example, the price of fertilizer is $50 a ton and the rental price of land is $50. The figure shows three budget lines representing various levels of spending on inputs. Line *A*, which corresponds to a total variable cost of $400, does not provide enough inputs to produce two hundred bushels of corn, because it does not touch the isoquant. Line *C*, which corresponds to a total variable cost of $625, provides enough inputs to grow two hundred bushels of corn using methods *P* or *R*. Line *B*, which corresponds to a total variable cost of $500, gives the least costly method of growing two hundred bushels, using method *Q*.

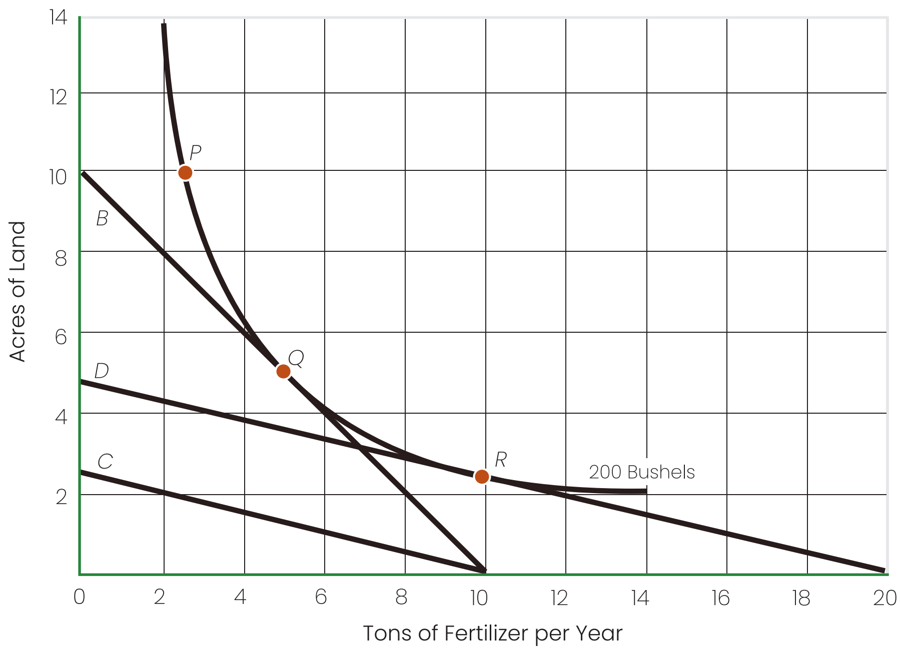
When we add the isoquant for two hundred bushels of corn to the set of budget lines, it becomes easy to find the least-cost method of production—namely, the method that uses 5.0 tons of fertilizer and 5.0 acres of land. That corresponds to point *Q* on the graph, where the isoquant just touches budget line *B*. Points *P* and *R* are possible ways of growing two hundred bushels of corn, but they lie on budget line *C*, which has input costs of $625. Note any budget of less than $500 (say, $400, as shown by budget line *A*) is not enough to reach the two hundred-bushel isoquant no matter how much goes to fertilizer and how much to land.

Responses to Changes in Input Prices

If input prices change, the least-cost combination of inputs will change as well. Suppose that suburbs begin to expand in the direction of Hathaway's farm, driving up the price of land. Land that used to rent for $50 per acre per year now costs $200 per acre. The price of fertilizer remains unchanged at $50 a ton.

Figure 6–9 shows the effects of the higher price of land. Now $500 is not enough to buy the combinations of inputs that fall along budget line *B*. Even if Hathaway spends all the money on land, the farmer can rent only 2.5 acres. The new $500 budget line is *C*, which does not reach the two hundred-bushel isoquant at any point.

Figure 6–9 Effects of a Change in Input Prices



If the rental price of land increases from \$50 to \$200 while the price of fertilizer remains fixed at \$50 a ton, it is no longer possible to produce two hundred bushels of corn for \$500. The \$500 budget line shifts from position *B* to position *C* and now falls short of the two hundred-bushel isoquant. Increasing the amount spent on variable inputs to \$1,000 shifts the budget line up to position *D*, where it just touches the isoquant at point *R*. The increase in the price of land not only raises the total variable cost of growing two hundred bushels of corn but also causes a substitution of fertilizer for land, which is now relatively more costly.

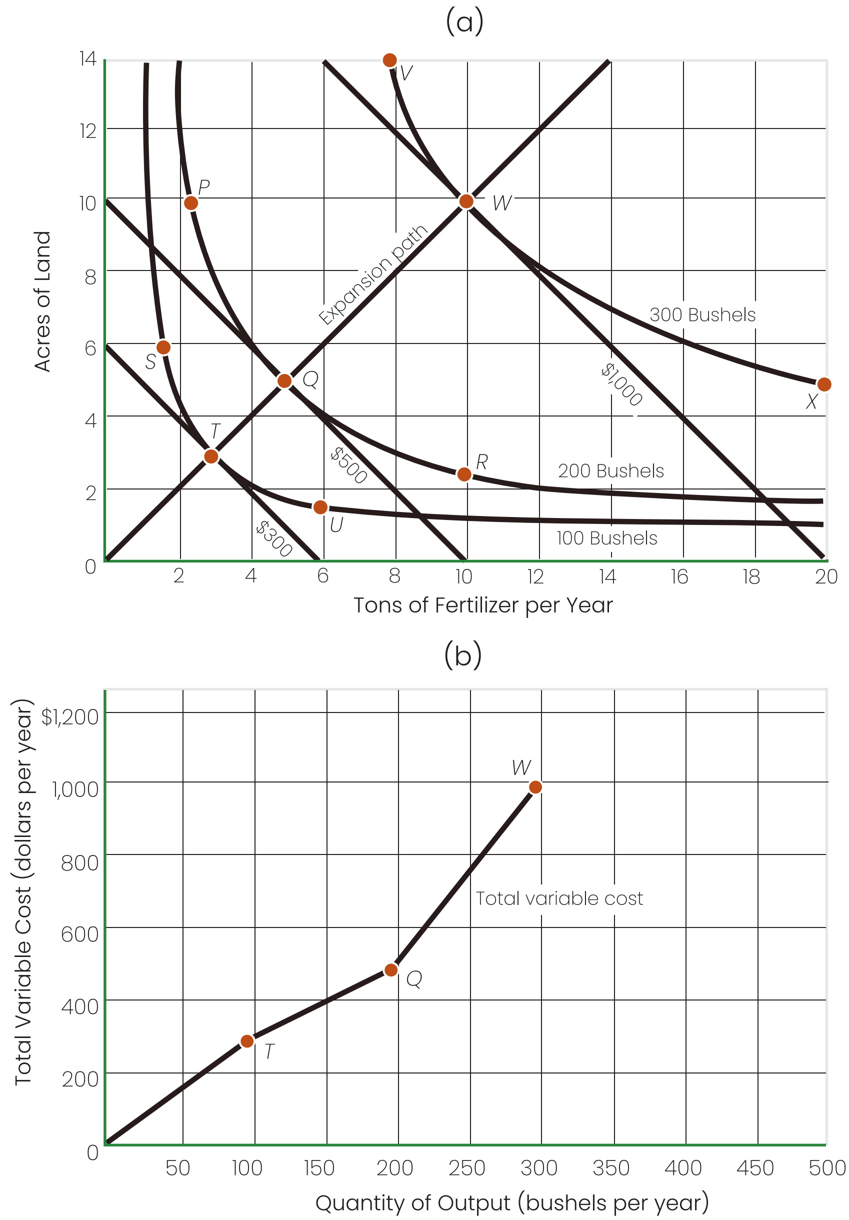
To grow two hundred bushels, Hathaway must now spend more than \$500. As the farmer increases spending on land and fertilizer, the budget line shifts upward but stays parallel to *C*. When the budget line reaches *D*, which corresponds to spending \$1,000 on inputs, it just touches the isoquant at *R*. We see that now \$1,000 is the lowest cost for growing two hundred bushels of corn, given a price of \$50 a ton for fertilizer and \$200 per acre for land. With those prices, *R* is the least-cost combination of inputs.

Notice that the increase in the price of land not only increases the cost of producing a given quantity of corn but also causes a substitution of fertilizer for land. We will return to this topic of substitution among factors of production in later chapters.

Varying Output

So far, we have assumed a fixed level of output. We can extend the isoquant technique to analyze variations in output. Panel (a) of Figure 6–10 shows three isoquants, each corresponding to a different level of output. Points *P*, *Q*, and *R* represent three ways of growing two hundred bushels of corn; points *S*, *T*, and *U*, three ways of growing one hundred bushels; and points *V*, *W*, and *X*, three ways of growing three hundred bushels.

Figure 6–10 Expansion of Output and Total Variable Costs



Panel (a) of this figure shows three isoquants for the production of corn, corresponding to outputs of one hundred, two hundred, and three hundred bushels per year. Assuming input prices of $50 an acre for land and $50 a ton for fertilizer, we can draw budget lines to show the minimum total variable cost for each output level. As output expands, the firm will move from *T* to *Q* and then to *W* along a line that we call an expansion path. Panel (b) of the figure plots the amount of output and the total variable cost for points along the expansion path. The result is a reverse-S-shaped total variable cost curve that shows diminishing returns for output levels above two hundred bushels per year.

In this example, we return to prices of $50 an acre for land and $50 a ton for fertilizer. Using those prices, we draw a set of budget lines, corresponding to total variable costs of $300, $500, and $1,000.

There is a least-cost method for producing each output level given these prices. Point *T* shows the best combination of land and fertilizer to produce one hundred bushels, *Q* shows the best combination for two hundred bushels, and *W* shows the best combination for three hundred bushels. If we added more isoquants, we could show the least-cost production points for other output levels. All of them would lie along the line drawn from the origin through points *T*, *Q*, and *W*. We call such a line the firm's **expansion path**. As the firm moves along its expansion path, it uses more of both the variable inputs, while holding fixed inputs (labor and machinery, in our example) constant.

Deriving a Cost Curve from the Isoquant Map

Once we have found the expansion path, we can easily construct a total variable cost curve for the firm. Panel (b) of Figure 6–10 shows how we can do that. At the origin, both output and total variable cost are zero. At point *T*, output is one hundred bushels per year and total variable cost is $300 per year; at *Q*, we have two hundred bushels and $500; and at *W*, three hundred bushels and $1,000. The firm's total variable cost curve is a line connecting those points.

Note that the cost curve has the reverse-S shape discussed earlier in the chapter. That shape is a result of the law of diminishing returns, as applied to the case in which two inputs vary while all others remain fixed. Beyond point *Q*, the amounts of inputs needed to produce each added unit of output begin to rise, just as they did in the case of just one variable input. Only if all inputs vary while none remains fixed can a firm escape the effects of the law of diminishing returns.

Expansion path
A line on an isoquant diagram showing the least-cost production points for various levels of output