In Chapter 9, we looked behind the demand curve to better understand consumer decision making. In this chapter, we look behind the supply curve to better understand firm decision making. Earlier chapters showed that supply curves are upward sloping because marginal cost increases as firms increase the quantity of a good that they supply. In this chapter, we look more closely at why this is true. In the appendix to this chapter, we extend the analysis by using isoquants and isocost lines to understand the relationship between production and costs. Once we have a good understanding of production and cost, we can proceed in the following chapters to understand how firms decide what level of output to produce and what price to charge.

10.1 LEARNING OBJECTIVE

10.1 | Define technology and give examples of technological change.

Technology: An Economic Definition

Technology The processes a firm uses to turn inputs into outputs of goods and services.

Technological change A change in the ability of a firm to produce a given level of output with a given quantity of inputs.

The basic activity of a firm is to use *inputs*, such as workers, machines, and natural resources, to produce *outputs* of goods and services. A pizza parlor, for example, uses inputs such as pizza dough, pizza sauce, cooks, and ovens to produce pizza. A firm's **technology** is the processes it uses to turn inputs into outputs of goods and services. Notice that this economic definition of technology is broader than the everyday definition. When we use the word *technology* in everyday language, we usually refer only to the development of new products. In the economic sense, a firm's technology depends on many factors, such as the skill of its managers, the training of its workers, and the speed and efficiency of its machinery and equipment. The technology of pizza production, for example, includes not only the capacity of the pizza ovens and how quickly they bake the pizza but also how quickly the cooks can prepare the pizza for baking, how well the manager motivates the workers, and how well the manager has arranged the facilities to allow the cooks to quickly prepare the pizzas and get them in the ovens.

Whenever a firm experiences positive **technological change**, it is able to produce more output using the same inputs or the same output using fewer inputs. Positive technological change can come from many sources. The firm's managers may rearrange the factory floor or the layout of a retail store, thereby increasing production and sales. The firm's workers may go through a training program. The firm may install faster or more reliable machinery or equipment. It is also possible for a firm to experience negative technological change. If a firm hires less-skilled workers or if a hurricane damages its facilities, the quantity of output it can produce from a given quantity of inputs may decline.

Better inventory controls have helped reduce firms' costs.

Making the Connection | Improving Inventory Control at Wal-Mart

Inventories are goods that have been produced but not yet sold. For a retailer such as Wal-Mart, inventories at any point in time include the goods on the store shelves as well as goods in warehouses. Inventories are an input into Wal-Mart's output of goods sold to consumers. Having money tied up in holding inventories is costly, so firms have an incentive to hold as few inventories as possible and to *turn over* their inventories as rapidly as possible by ensuring that goods do not remain on the shelves long. Holding too few inventories, however, results in *stockouts*—that is, sales being lost because the goods consumers want to buy are not on the shelf.

Improvements in inventory control meet the economic definition of positive technological change because they allow firms to produce the same output with fewer inputs. In recent years, many firms have adopted *just-in-time* inventory systems in which firms accept shipments from suppliers as close as possible to the time they will be needed. The just-in-time system was pioneered by Toyota, which used it to reduce the inventories of parts in its automobile assembly plants. Wal-Mart has been a pioneer in using similar inventory control systems in its stores.

Wal-Mart actively manages its *supply chain*, which stretches from the manufacturers of the goods it sells to its retail stores. Entrepreneur Sam Walton, the company founder, built a series of distribution centers spread across the country to supply goods to the retail stores. As goods are sold in the stores, this *point-of-sale* information is sent electronically to the firm's distribution centers to help managers determine what products will be shipped to each store. Depending on a store's location relative to a distribution center, managers can use Wal-Mart's trucks to ship goods overnight. This distribution system allows Wal-Mart to minimize its inventory holdings without running the risk of many stockouts. Because Wal-Mart sells 15 percent to 25 percent of all the toothpaste, disposable diapers, dog food, and many other products sold in the United States, it has been able to involve many manufacturers closely in its supply chain. For example, a company such as Procter & Gamble, which is one of the world's largest manufacturers of toothpaste, laundry detergent, toilet paper, and other products, receives Wal-Mart's point-of-sale and inventory information electronically. Procter & Gamble uses that information to help determine its production schedules and the quantities it should ship to Wal-Mart's distribution centers.

Technological change has been a key to Wal-Mart's becoming one of the largest firms in the world, with 1.9 million employees and revenue of more than $348 billion in 2006.

YOUR TURN: Test your understanding by doing related problem 1.5 on page 356 at the end of this chapter.

10.2 LEARNING OBJECTIVE

10.2 | Distinguish between the economic short run and the economic long run.

The Short Run and the Long Run in Economics

When firms analyze the relationship between their level of production and their costs, they separate the time period involved into the short run and the long run. In the **short run**, at least one of the firm's inputs is fixed. In particular, in the short run, the firm's technology and the size of its physical plant—its factory, store, or office—are both fixed, while the number of workers the firm hires is variable. In the **long run**, the firm is able to vary all its inputs and can adopt new technology and increase or decrease the size of its physical plant. Of course, the actual length of calendar time in the short run will be different from firm to firm. A pizza parlor may be able to increase its physical plant by adding another pizza oven and some tables and chairs in just a few weeks. BMW, in contrast, may take more than a year to increase the capacity of one of its automobile assembly plants by installing new equipment.

Short run The period of time during which at least one of a firm's inputs is fixed.

Long run The period of time in which a firm can vary all its inputs, adopt new technology, and increase or decrease the size of its physical plant.

The Difference between Fixed Costs and Variable Costs

Total cost is the cost of all the inputs a firm uses in production. We have just seen that in the short run, some inputs are fixed and others are variable. The costs of the fixed inputs are *fixed costs*, and the costs of the variable inputs are *variable costs*. We can also think of **variable costs** as the costs that change as output changes. Similarly, **fixed costs** are costs that remain constant as output changes. A typical firm's variable costs include its labor costs, raw material costs, and costs of electricity and other utilities. Typical fixed costs include lease payments for factory or retail space, payments for fire insurance, and payments for newspaper and television advertising. All of a firm's costs are either fixed or variable, so we can state the following:

Total cost The cost of all the inputs a firm uses in production.

Variable costs Costs that change as output changes.

Fixed costs Costs that remain constant as output changes.

$$\text{Total Cost} = \text{Fixed Cost} + \text{Variable Cost}$$

or, using symbols:

$$TC = FC + VC.$$

Publishers consider the salaries of editors to be a fixed cost.

Making the Connection

Fixed Costs in the Publishing Industry

An editor at Cambridge University Press gives the following estimates of the annual fixed cost for a medium-size academic book publisher.

COST	AMOUNT
Salaries and benefits	$437,500
Rent	75,000
Utilities	20,000
Supplies	6,000
Postage	4,000
Travel	8,000
Subscriptions, etc.	4,000
Miscellaneous	5,000
Total	$559,500

Academic book publishers hire editors, designers, and production and marketing managers who help prepare books for publication. Because these employees work on several books simultaneously, the number of people the company hires does not go up and down with the quantity of books the company publishes during any particular year. Publishing companies therefore consider the salaries and benefits of people in these job categories as fixed costs.

In contrast, for a company that *prints* books, the quantity of workers varies with the quantity of books printed. The wages and benefits of the workers operating the printing presses, for example, would be a variable cost.

The other costs listed in the preceding table are typical of fixed costs at many firms.

Source: Beth Luey, *Handbook for Academic Authors*, 4th ed., Cambridge, UK: Cambridge University Press, 2002, p. 244.

YOUR TURN: Test your understanding by doing related problems 2.3, 2.4, and 2.5 on page 357 at the end of this chapter.

Implicit Costs versus Explicit Costs

Opportunity cost The highest-valued alternative that must be given up to engage in an activity.

Explicit cost A cost that involves spending money.

Implicit cost A nonmonetary opportunity cost.

It is important to remember that economists always measure costs as *opportunity costs*. The **opportunity cost** of any activity is the highest-valued alternative that must be given up to engage in that activity. As we saw in Chapter 7, costs are either *explicit* or *implicit*. When a firm spends money, it incurs an **explicit cost**. When a firm experiences a nonmonetary opportunity cost, it incurs an **implicit cost**.

For example, suppose that Jill Johnson owns a pizza restaurant. In operating her store, Jill has explicit costs, such as the wages she pays her workers and the payments she makes for rent and electricity. But some of Jill's most important costs are implicit. Before opening her own restaurant, Jill earned a salary of $30,000 per year managing a restaurant for someone else. To start her restaurant, Jill quit her job, withdrew $50,000 from her bank account—where it earned her interest of $3,000 per year—and used the funds to equip her restaurant with tables, chairs, a cash register, and other equipment. To open her own business, Jill had to give up the $30,000 salary and the $3,000 in interest. This $33,000 is an implicit cost because it does not represent payments that Jill has to make. All the same, giving up this $33,000 per year is a real cost to Jill. In addition, during the course of the year, the $50,000 worth of tables, chairs, and other physical capital in Jill's store will lose some of its value due partly to wear and tear and partly to better furniture, cash registers, and so forth becoming available. *Economic depreciation* is the difference between what Jill paid for her capital at the beginning of the year and what she could sell the capital for at the end of the year. If Jill could sell the capital for $40,000 at the end of the year, then the $10,000 in economic depreciation represents another implicit cost.

TABLE 10-1

Jill Johnson's Costs per Year

Pizza dough, tomato sauce, and other ingredients	$20,000
Wages	48,000
Interest payments on loan to buy pizza ovens	10,000
Electricity	6,000
Lease payment for store	24,000
Foregone salary	30,000
Foregone interest	3,000
Economic depreciation	10,000
Total	$151,000

(Note that the whole $50,000 she spent on the capital is not a cost because she still has the equipment at the end of the year, although it is now worth only $40,000.)

Table 10-1 lists Jill's costs. The entries in red are explicit costs, and the entries in blue are implicit costs. As we saw in Chapter 7, the rules of accounting generally require that only explicit costs be used for purposes of keeping the company's financial records and for paying taxes. Therefore, explicit costs are sometimes called *accounting costs*. *Economic costs* include both accounting costs and implicit costs.

The Production Function

Let's look at the relationship between the level of production and costs in the short run for Jill Johnson's restaurant. To keep things simpler than in the more realistic situation in Table 10-1, let's assume that Jill uses only labor—workers—and one type of capital—pizza ovens—to produce a single good: pizzas. Many firms use more than two inputs and produce more than one good, but it is easier to understand the relationship between output and cost by focusing on the case of a firm using only two inputs and producing only one good. In the short run, Jill doesn't have time to build a larger restaurant, install additional pizza ovens, or redesign the layout of her restaurant. So, in the short run, she can increase or decrease the quantity of pizzas she produces only by increasing or decreasing the quantity of workers she employs.

The first three columns of Table 10-2 show the relationship between the quantity of workers and ovens Jill uses each week and the quantity of pizzas she can produce. The relationship between the inputs employed by a firm and the maximum output it can

TABLE 10-2 | Short-Run Production and Cost at Jill Johnson's Restaurant

QUANTITY OF WORKERS	QUANTITY OF PIZZA OVENS	QUANTITY OF PIZZAS PER WEEK	COST OF PIZZA OVENS (FIXED COST)	COST OF WORKERS (VARIABLE COST)	TOTAL COST OF PIZZAS	COST PER PIZZA (AVERAGE TOTAL COST)
0	2	0	$800	$0	$800	—
1	2	200	800	650	1,450	$7.25
2	2	450	800	1,300	2,100	4.67
3	2	550	800	1,950	2,750	5.00
4	2	600	800	2,600	3,400	5.67
5	2	625	800	3,250	4,050	6.48
6	2	640	800	3,900	4,700	7.34

Production function The relationship between the inputs employed by a firm and the maximum output it can produce with those inputs.

produce with those inputs is called the firm's **production function.** Because a firm's technology is the processes it uses to turn inputs into output, the production function represents the firm's technology. In this case, Table 10-2 shows Jill's *short-run* production function because we are assuming that the time period is too short for Jill to increase or decrease the quantity of ovens she is using.

A First Look at the Relationship between Production and Cost

Table 10-2 gives us information on Jill's costs. We can determine the total cost of producing a given quantity of pizzas if we know how many workers and ovens are required to produce that quantity of pizzas and what Jill has to pay for those workers and pizzas. Suppose Jill has taken out a bank loan to buy two pizza ovens. The cost of the loan is \$800 per week. Therefore, her fixed costs are \$800 per week. If Jill pays \$650 per week to each worker, her variable costs depend on how many workers she hires. In the short run, Jill can increase the quantity of pizzas she produces only by hiring more workers. The table shows that if she hires 1 worker, she produces 200 pizzas during the week; if she hires 2 workers, she produces 450 pizzas; and so on. For a particular week, Jill's total cost of producing pizzas is equal to the \$800 she pays on the loan for the ovens plus the amount she pays to hire workers. If Jill decides to hire 4 workers and produce 600 pizzas, her total cost is \$3,400: \$800 to lease the ovens and \$2,600 to hire the workers. Her cost per pizza is equal to her total cost of producing pizzas divided by the quantity of pizzas produced. If she produces 600 pizzas at a total cost of \$3,400, her cost per pizza, or *average total cost*, is \$3,400/600 = \$5.67. A firm's **average total cost** is always equal to its total cost divided by the quantity of output produced.

Average total cost Total cost divided by the quantity of output produced.

Panel (a) of Figure 10-1 uses the numbers in the next-to-last column of Table 10-2 to graph Jill's total cost. Panel (b) uses the numbers in the last column to graph her average

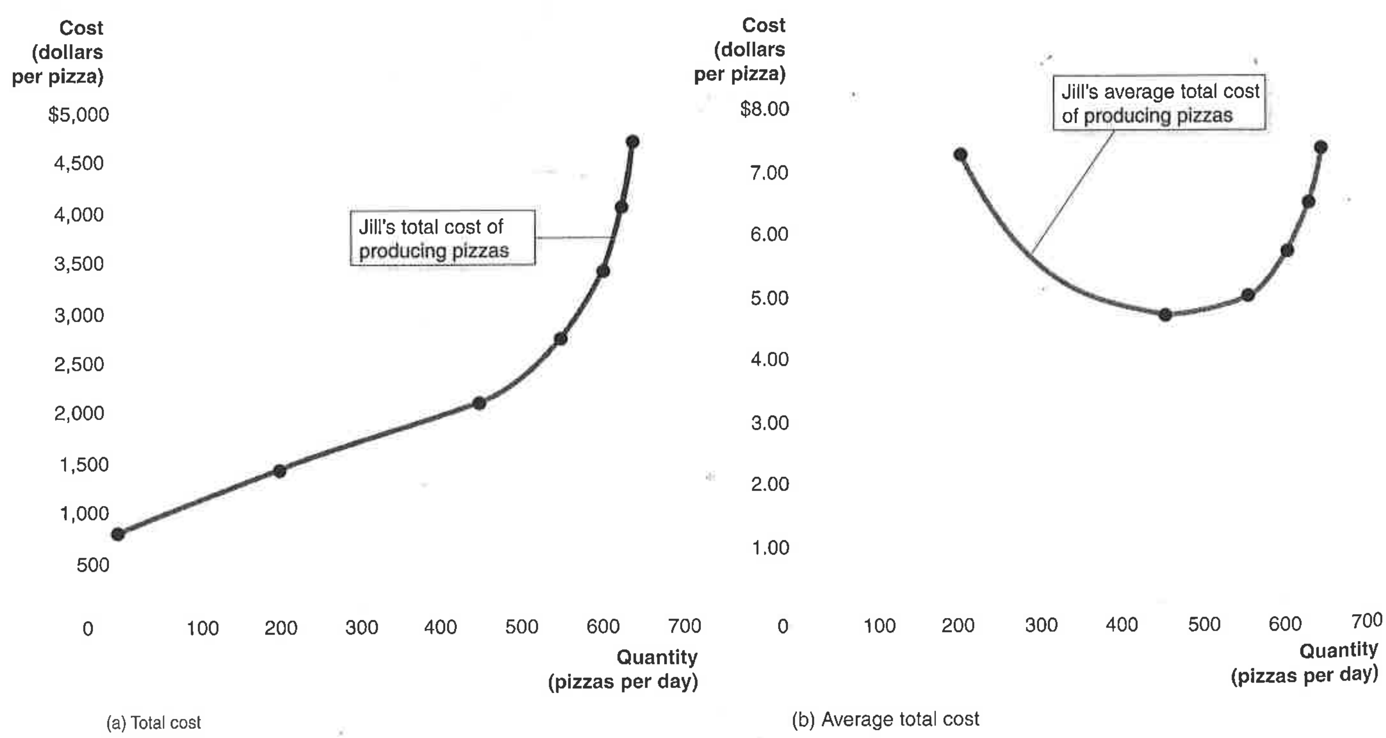


Figure 10-1 | Graphing Total Cost and Average Total Cost of Jill Johnson's Restaurant

We can use the information from Table 10-2 to graph the relationship between the quantity of pizzas Jill produces and her total cost and average total cost. Panel (a) shows that total cost increases as the level of production increases. In panel (b), we see that the average total cost is roughly U-shaped: As production increases from low levels, average cost falls before rising at higher levels of production. To understand why average cost has this shape, we must look more closely at the technology of producing pizzas, as shown by the production function.

total cost. Notice in panel (b) that Jill's average cost has roughly the same U shape as the average cost curve we saw Akio Morita calculate for Sony transistor radios at the beginning of this chapter. As production increases from low levels, average cost falls. Average cost then becomes fairly flat, before rising at higher levels of production. To understand why average cost has this U shape, we first need to look more closely at the technology of producing pizzas, as shown by the production function for Jill's restaurant. Then we need to look at how this technology determines the relationship between production and cost.

10.3 LEARNING OBJECTIVE

10.3 | Understand the relationship between the marginal product of labor and the average product of labor.

The Marginal Product of Labor and the Average Product of Labor

To better understand the choices Jill faces, given the technology available to her, think first about what happens if she hires only one worker. That one worker will have to perform several different activities, including taking orders from customers, baking the pizzas, bringing the pizzas to the customers' tables, and ringing up sales on the cash register. If Jill hires two workers, some of these activities can be divided up: One worker could take the orders and ring up the sales, and one worker could bake the pizzas. With this division of tasks, Jill will find that hiring two workers actually allows her to produce more than twice as many pizzas as she could produce with just one worker.

The additional output a firm produces as a result of hiring one more worker is called the **marginal product of labor**. We can calculate the marginal product of labor by determining how much total output increases as each additional worker is hired. We do this for Jill's restaurant in Table 10-3.

Marginal product of labor The additional output a firm produces as a result of hiring one more worker.

When Jill hires only 1 worker, she produces 200 pizzas per week. When she hires 2 workers, she produces 450 pizzas per week. Hiring the second worker increases her production by 250 pizzas per week. So, the marginal product of labor for 1 worker is 200 pizzas. For 2 workers, the marginal product of labor rises to 250 pizzas. This increase in marginal product results from the *division of labor* and from *specialization*. By dividing the tasks to be performed—the division of labor—Jill reduces the time workers lose moving from one activity to the next. She also allows them to become more specialized at their tasks. For example, a worker who concentrates on baking pizzas will become skilled at doing so quickly and efficiently.

The Law of Diminishing Returns

In the short run, the quantity of pizza ovens Jill leases is fixed, so as she hires more workers, the marginal product of labor eventually begins to decline. This happens because at some point, Jill uses up all the gains from the division of labor and from specialization

TABLE 10-3

The Marginal Product of Labor at Jill Johnson's Restaurant

QUANTITY OF WORKERS	QUANTITY OF PIZZA OVENS	QUANTITY OF PIZZAS	MARGINAL PRODUCT OF LABOR
0	2	0	—
1	2	200	200
2	2	450	250
3	2	550	100
4	2	600	50
5	2	625	25
6	2	640	15

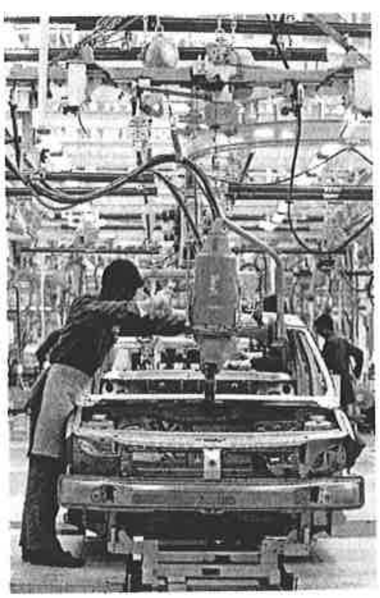
Law of diminishing returns The principle that, at some point, adding more of a variable input, such as labor, to the same amount of a fixed input, such as capital, will cause the marginal product of the variable input to decline.

and starts to experience the effects of the **law of diminishing returns**. This law states that adding more of a variable input, such as labor, to the same amount of a fixed input, such as capital, will eventually cause the marginal product of the variable input to decline. For Jill, the marginal product of labor begins to decline when she hires the third worker. Hiring three workers raises the quantity of pizzas she produces from 450 per week to 550. But the increase in the quantity of pizzas—100—is less than the increase when she hired the second worker—250.

If Jill kept adding more and more workers to the same quantity of pizza ovens, eventually workers would begin to get in each other's way, and the marginal product of labor would actually become negative. When the marginal product is negative, the level of total output declines. No firm would actually hire so many workers as to experience a negative marginal product of labor and falling total output.

Graphing Production

Panel (a) in Figure 10-2 shows the relationship between the quantity of workers Jill hires and her total output of pizzas, using the numbers from Table 10-3. Panel (b) shows the marginal product of labor. In panel (a), output increases as more workers are hired, but the increase in output does not occur at a constant rate. Because of specialization and the division of labor, output at first increases at an increasing rate, with each additional worker hired causing production to increase by a *greater* amount than did the hiring of the previous worker. But after the second worker has been hired, hiring more workers while keeping the quantity of ovens constant results in diminishing returns. When the point of diminishing returns is reached, production increases at a decreasing rate. Each additional worker hired after the second worker causes production to increase by a *smaller* amount than did the hiring of the previous worker. In panel (b), the marginal product of labor curve rises initially because of the effects of specialization and division of labor, and then it falls due to the effects of diminishing returns.



The gains from division of labor and specialization are as important to firms today as they were in the eighteenth century, when Adam Smith first discussed them.

Making the Connection | Adam Smith's Famous Account of the Division of Labor in a Pin Factory

In *The Wealth of Nations*, Adam Smith uses production in a pin factory as an example of the gains in output resulting from the division of labor. The following is an excerpt from his account of how pin making was divided into a series of tasks:

> One man draws out the wire, another straightens it, a third cuts it, a fourth points it, a fifth grinds it at the top for receiving the head; to make the head requires two or three distinct operations; to put it on is a [distinct operation], to whiten the pins is another; it is even a trade by itself to put them into the paper; and the important business of making a pin is, in this manner, divided into eighteen distinct operations.

Because the labor of pin making was divided up in this way, the average worker was able to produce about 4,800 pins per day. Smith speculated that a single worker using the pin-making machinery alone would make only about 20 pins per day. This lesson from more than 225 years ago, showing the tremendous gains from division of labor and specialization, remains relevant to most business situations today.

Source: Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations*, Vol. 1, Oxford, UK: Oxford University Press edition, 1976, pp. 14–15.

YOUR TURN: Test your understanding by doing related problem 3.6 on page 358 at the end of this chapter.

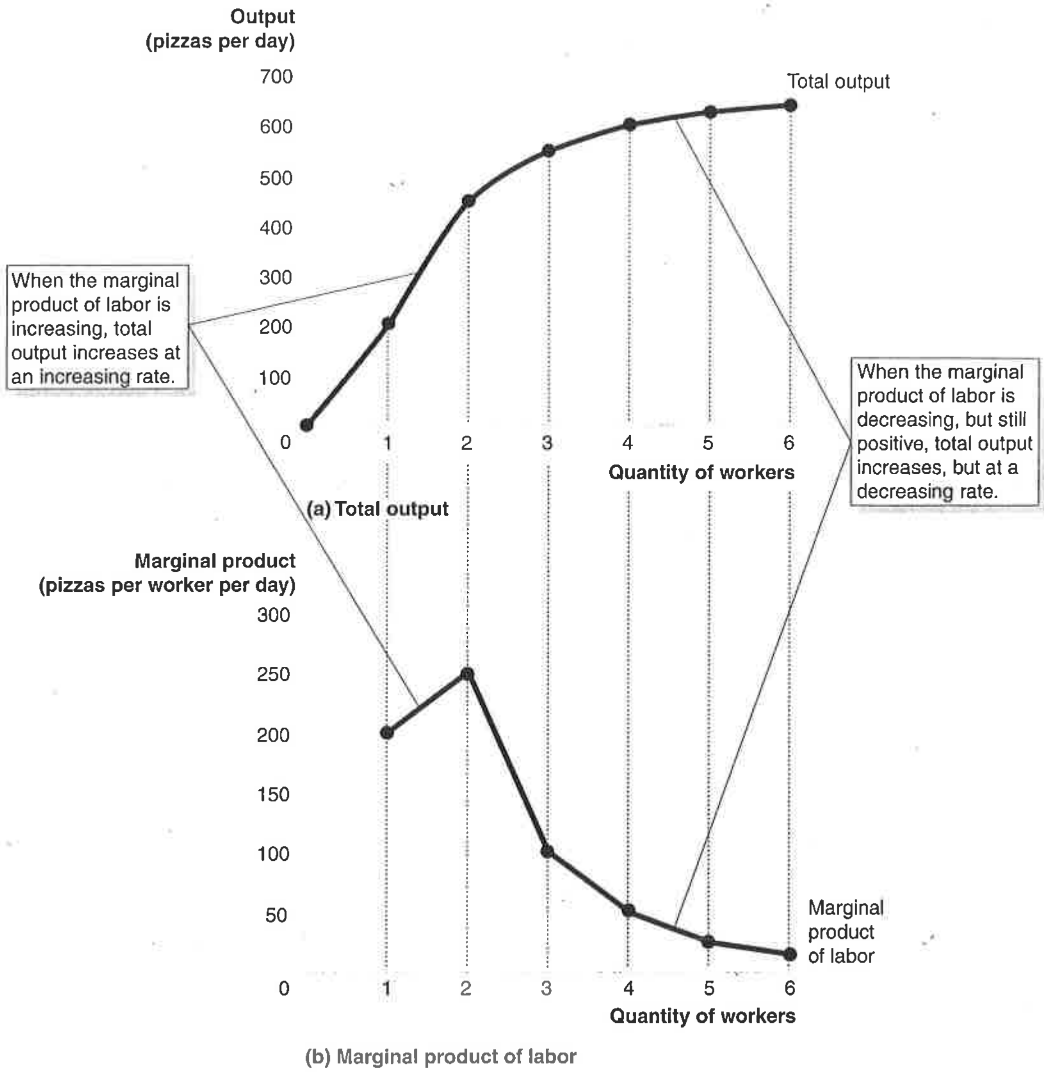


Figure 10-2 | Total Output and the Marginal Product of Labor

In panel (a), output increases as more workers are hired, but the increase in output does not occur at a constant rate. Because of specialization and the division of labor, output at first increases at an increasing rate, with each additional worker hired causing production to increase by a *greater* amount than did the hiring of the previous worker. After the third worker has been hired, hiring more workers while keeping the number of pizza ovens constant results in diminishing returns. When the point of diminishing returns is reached, production increases at a decreasing rate. Each additional worker hired after the third worker causes production to increase by a *smaller* amount than did the hiring of the previous worker. In panel (b), the *marginal product of labor* is the additional output produced as a result of hiring one more worker. The marginal product of labor rises initially because of the effects of specialization and division of labor, and then it falls due to the effects of diminishing returns.

The Relationship between Marginal and Average Product

The marginal product of labor tells us how much total output changes as the quantity of workers hired changes. We can also calculate how many pizzas workers produce on average. The **average product of labor** is the total output produced by a firm divided by the quantity of workers. For example, using the numbers in Table 10-3, if Jill hires 4 workers to produce 600 pizzas, the average product of labor is 600/4 = 150.

Average product of labor The total output produced by a firm divided by the quantity of workers.

We can state the relationship between the marginal and average products of labor this way: *The average product of labor is the average of the marginal products of labor.* For example, the numbers from Table 10-3 show that the marginal product of the first worker Jill hires is 200, the marginal product of the second worker is 250, and the marginal product of the third worker is 100. Therefore, the average product of labor for three workers is 183.3:

$$183.3 = (200 + 250 + 100)/3$$

- 183.3: Average product of labor for three workers
- 200: Marginal product of labor of first worker
- 250: Marginal product of labor of second worker
- 100: Marginal product of labor of third worker

By taking the average of the marginal products of the first three workers, we have the average product of the three workers.

Whenever the marginal product of labor is greater than the average product of labor, the average product of labor must be increasing. This statement is true for the same reason that a person 6 feet, 2 inches tall entering a room where the average height is 5 feet, 9 inches raises the average height of people in the room. Whenever the marginal product of labor is less than the average product of labor, the average product of labor must be decreasing. The marginal product of labor equals the average product of labor for the quantity of workers where the average product of labor is at its maximum.

An Example of Marginal and Average Values: College Grades

The relationship between the marginal product of labor and the average product of labor is the same as the relationship between the marginal and average values of any variable. To see this more clearly, think about the familiar relationship between a student's grade point average (GPA) in one semester and his overall, or cumulative, GPA. The table in Figure 10-3 shows Paul's college grades for each semester, beginning with fall 2005. The graph in Figure 10-3 plots the grades from the table. Just as each additional worker hired adds to a firm's total production, each additional semester adds to Paul's total grade points. We can calculate what each individual worker hired adds to total production (marginal product), and we can calculate the average production of the workers hired so far (average product).

Similarly, we can calculate the GPA Paul earns in a particular semester (his "marginal GPA"), and we can calculate his cumulative GPA for all the semesters he has completed so far (his "average GPA"). As the table shows, Paul gets off to a weak start in the fall semester of his freshman year, earning only a 1.50 GPA. In each subsequent semester through the fall of his junior year, his GPA for the semester increases from the previous semester—raising his cumulative GPA. As the graph shows, however, his cumulative GPA does not increase as rapidly as his semester-by-semester GPA because his cumulative GPA is held back by the low GPAs of his first few semesters. Notice that in Paul's junior year, even though his semester GPA declines from fall to spring, his cumulative GPA rises. Only in the fall of his senior year, when his semester GPA drops below his cumulative GPA, does his cumulative GPA decline.

10.4 LEARNING OBJECTIVE

10.4 | Explain and illustrate the relationship between marginal cost and average total cost.

The Relationship between Short-Run Production and Short-Run Cost

We have seen that technology determines the values of the marginal product of labor and the average product of labor. In turn, the marginal and average products of labor affect the firm's costs. Keep in mind that the relationships we are discussing are *short-run* relationships: We are assuming that the time period is too short for the firm to change its technology or the size of its physical plant.

	Semester GPA (Marginal) GPA	Cumulative GPA (Average) GPA
Freshman Year		
Fall	1.50	1.50
Spring	2.00	1.75
Sophomore Year		
Fall	2.20	1.90
Spring	3.00	2.18
Junior Year		
Fall	3.20	2.38
Spring	3.00	2.48
Senior Year		
Fall	2.40	2.47
Spring	2.00	2.41

Average GPA continues to rise, although marginal GPA falls.

With the marginal GPA below the average, the average GPA falls.

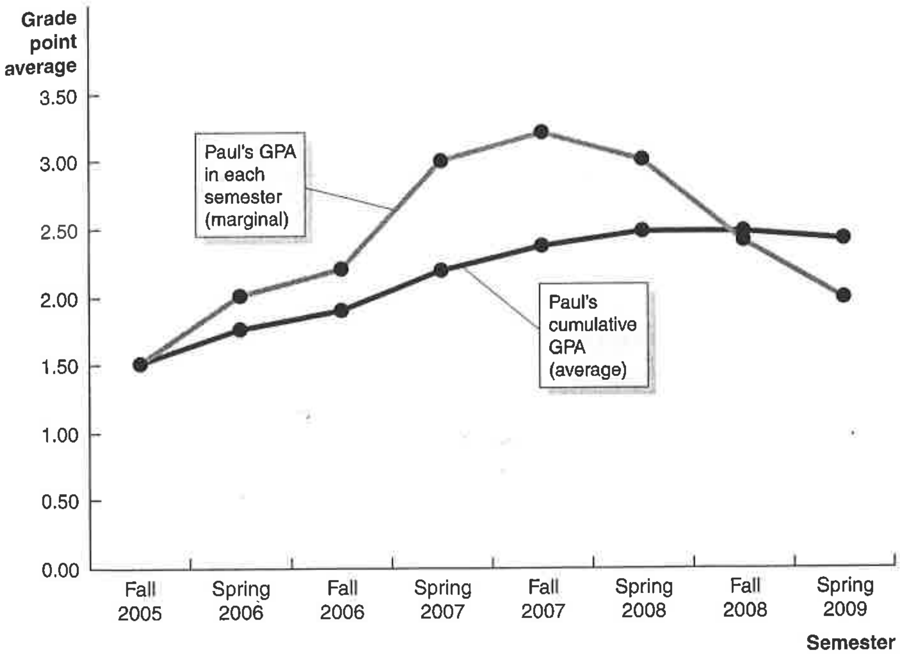


Figure 10-3

Marginal and Average GPAs

The relationship between marginal and average values for a variable can be illustrated using GPAs. We can calculate the GPA Paul earns in a particular semester (his "marginal GPA"), and we can calculate his cumulative GPA for all the semesters he has completed so far (his "average GPA"). Paul's GPA is only 1.50 in the fall semester of his freshman year. In each following semester through fall of his junior year, his GPA for the semester increases—raising his cumulative GPA. In Paul's junior year, even though his semester GPA declines from fall to spring, his cumulative GPA rises. Only in the fall of his senior year, when his semester GPA drops below his cumulative GPA, does his cumulative GPA decline.

At the beginning of this chapter, we saw how Akio Morita used an average total cost curve to determine the price of radios. The average total cost curve Morita used and the average total cost curve in Figure 10-1 for Jill Johnson's restaurant both have a U shape. As we will soon see, the U shape of the average total cost curve is determined by the shape of the curve that shows the relationship between *marginal cost* and the level of production.

Marginal Cost

As we saw in Chapter 1, one of the key ideas in economics is that optimal decisions are made at the margin. Consumers, firms, and government officials usually make decisions about doing a little more or a little less. As Jill Johnson considers whether to hire additional workers to produce additional pizzas, she needs to consider how much she will add to her total cost by producing the additional pizzas. **Marginal cost** is the change in a firm's total cost from producing one more unit of a good or service. We can calculate marginal cost for a particular increase in output by dividing the change in cost by the

Marginal cost The change in a firm's total cost from producing one more unit of a good or service.

change in output. We can express this idea mathematically (remembering that the Greek letter delta, Δ, means "change in"):

$$MC = \frac{\Delta TC}{\Delta Q}.$$

In the table in Figure 10-4, we use this equation to calculate Jill's marginal cost of producing pizzas.

Why Are the Marginal and Average Cost Curves U-Shaped?

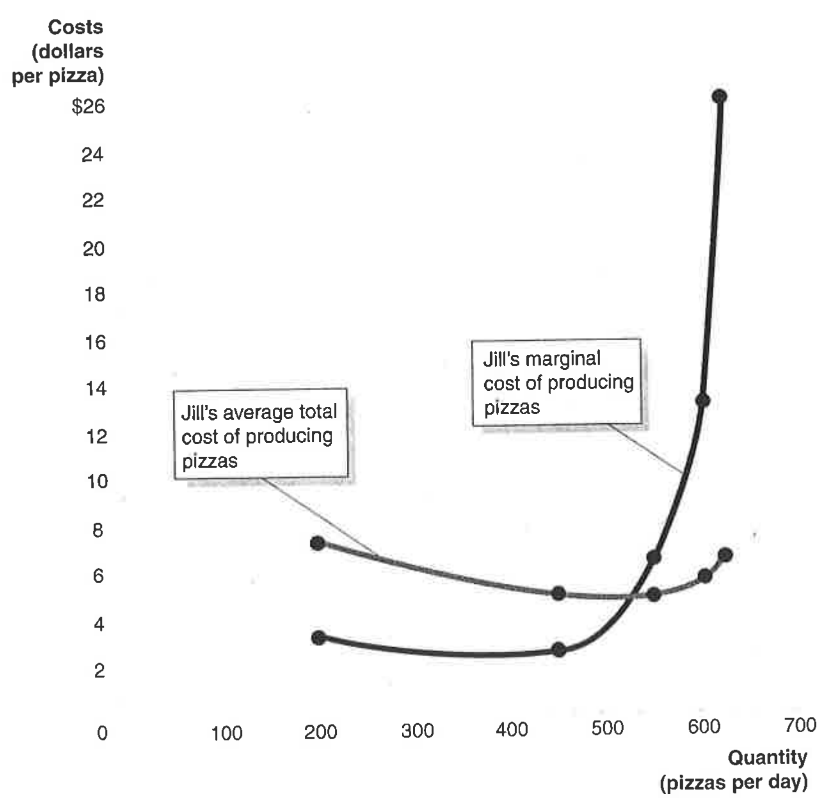
Notice in the graph in Figure 10-4 that Jill's marginal cost of producing pizzas declines at first and then increases, giving the marginal cost curve a U shape. The table in Figure 10-4 also shows the marginal product of labor. This table helps us see the important relationship between the marginal product of labor and the marginal cost of production: The marginal product of labor is *rising* for the first two workers, but the marginal cost of the pizzas produced by these workers is *falling*. The marginal product of labor is *falling* for the last four workers, but the marginal cost of pizzas produced by these workers is *rising*. To summarize this point: *When the marginal product of labor is rising, the marginal cost of output is falling. When the marginal product of labor is falling, the marginal cost of production is rising.*

Figure 10-4

Jill Johnson's Marginal Cost and Average Total Cost of Producing Pizzas

We can use the information in the table to calculate Jill's marginal cost and average total cost of producing pizzas. For the first two workers hired, the marginal product of labor is increasing. This increase causes the marginal cost of production to fall. For the last four workers hired, the marginal product of labor is falling. This causes the marginal cost of production to increase. Therefore, the marginal cost curve falls and then rises—that is, has a U shape—because the marginal product of labor rises and then falls. As long as marginal cost is below average total cost, average total cost will be falling. When marginal cost is above average total cost, average total cost will be rising. The relationship between marginal cost and average total cost explains why the average total cost curve also has a U shape.

Quantity of Workers	Quantity of Ovens	Marginal Product of Labor	Total Cost of Pizzas	Marginal Cost of Pizzas	Average Total Cost of Pizzas
0	0	–	$800	–	–
1	200	200	1,450	$3.25	$7.25
2	450	250	2,100	2.60	4.67
3	550	100	2,750	6.50	5.00
4	600	50	3,400	13.00	5.67
5	625	25	4,050	26.00	6.48
6	640	15	4,700	43.33	7.34



One way to understand why this point is true is first to notice that the only additional cost to Jill from producing more pizzas is the additional wages she pays to hire more workers. She pays each new worker the same $650 per week. So the marginal cost of the additional pizzas each worker makes depends on that worker's additional output, or marginal product. As long as the additional output from each new worker is rising, the marginal cost of that output is falling. When the additional output from each new worker is falling, the marginal cost of that output is rising. *We can conclude that the marginal cost of production falls and then rises—forming a U shape—because the marginal product of labor rises and then falls.*

The relationship between marginal cost and average total cost follows the usual relationship between marginal and average values. As long as marginal cost is below average total cost, average total cost falls. When marginal cost is above average total cost, average total cost rises. Marginal cost equals average total cost when average total cost is at its lowest point. Therefore, the average total cost curve has a U shape because the marginal cost curve has a U shape.

Solved Problem | 10-4

The Relationship between Marginal Cost and Average Cost

Is Jill Johnson right or wrong when she says the following? "I am currently producing 10,000 pizzas per month at a total cost of $500.00. If I produce 10,001 pizzas, my total cost will rise to $500.11. Therefore, my marginal cost of producing pizzas must be increasing." Draw a graph to illustrate your answer.

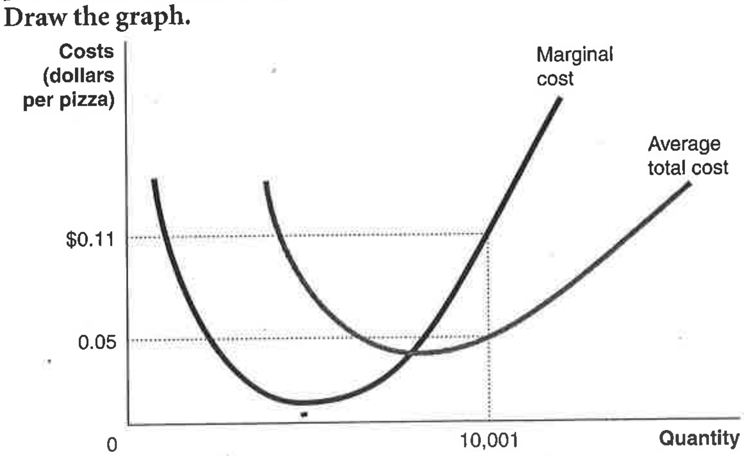
SOLVING THE PROBLEM:

Step 1: **Review the chapter material.** This problem requires understanding the relationship between marginal and average cost, so you may want to review the section "Why Are the Marginal and Average Cost Curves U-Shaped?" which begins on page 344.

Step 2: **Calculate average total cost and marginal cost.** Average total cost is total cost divided by total output. In this case, average total cost is $500.11/10,001 = $0.05. Marginal cost is the change in total cost divided by the change in output. In this case, marginal cost is $0.11/1 = $0.11.

Step 3: **Use the relationship between marginal cost and average total cost to answer the question.** When marginal cost is greater than average total cost, marginal cost must be increasing. You have shown in step 2 that marginal cost is greater than average total cost. Therefore, Jill is right: Her marginal cost of producing pizzas must be increasing.

Step 4: **Draw the graph.**



YOUR TURN: For more practice, do related problems 4.5 and 4.6 on page 359 at the end of this chapter.

>> End Solved Problem 10-4

10.5 LEARNING OBJECTIVE

10.5 | Graph average total cost, average variable cost, average fixed cost, and marginal cost.

Graphing Cost Curves

Average fixed cost Fixed cost divided by the quantity of output produced.

Average variable cost Variable cost divided by the quantity of output produced.

We have seen that we calculate average total cost by dividing total cost by the quantity of output produced. Similarly, we can calculate **average fixed cost** by dividing fixed cost by the quantity of output produced. And we can calculate **average variable cost** by dividing variable cost by the quantity of output produced. Or, mathematically, with Q being the level of output, we have:

$$\text{Average total cost} = ATC = \frac{TC}{Q}$$

$$\text{Average fixed cost} = AFC = \frac{FC}{Q}$$

$$\text{Average variable cost} = AVC = \frac{VC}{Q}.$$

Finally, notice that average total cost is the sum of average fixed cost plus average variable cost:

$$ATC = AFC + AVC.$$

The only fixed cost Jill incurs in operating her restaurant is the $800 per week she pays on the bank loan for her pizza ovens. Her variable costs are the wages she pays her workers. The table and graph in Figure 10-5 show Jill's costs.

We will use graphs like the one in Figure 10-5 in the next several chapters to analyze how firms decide the level of output to produce and the price to charge. Before going further, be sure you understand the following three key facts about Figure 10-5:

1. The marginal cost (MC), average total cost (ATC), and average variable cost (AVC) curves are all U-shaped, and the marginal cost curve intersects the average variable cost and average total cost curves at their minimum points. When marginal cost is less than either average variable cost or average total cost, it causes them to decrease. When marginal cost is above average variable cost or average total cost, it causes them to increase. Therefore, when marginal cost equals average variable cost or average total cost, they must be at their minimum points.
2. As output increases, average fixed cost gets smaller and smaller. This happens because in calculating average fixed cost, we are dividing something that gets larger and larger—output—into something that remains constant—fixed cost. Firms often refer to this process of lowering average fixed cost by selling more output as "spreading the overhead." By "overhead" they mean fixed costs.
3. As output increases, the difference between average total cost and average variable cost decreases. This happens because the difference between average total cost and average variable cost is average fixed cost, which gets smaller as output increases.

10.6 LEARNING OBJECTIVE

10.6 | Understand how firms use the long-run average cost curve in their planning.

Costs in the Long Run

The distinction between fixed cost and variable cost that we just discussed applies to the short run but *not* to the long run. For example, in the short run, Jill Johnson has fixed costs of $800 per week because she signed a loan agreement with a bank when she bought her pizza ovens. In the long run, the cost of purchasing more pizza ovens becomes variable because Jill can choose whether to expand her business by buying

Quantity of Workers	Quantity of Ovens	Quantity of Pizzas	Cost of Ovens (Fixed Cost)	Cost of Workers (Variable Cost)	Total Cost of Pizzas	*ATC*	*AFC*	*AVC*	*MC*
0	2	0	$800	$0	$800	—	—	—	—
1	2	200	800	650	1,450	$7.25	$4.00	$3.25	$3.25
2	2	450	800	1,300	2,100	4.67	1.78	2.89	2.60
3	2	550	800	1,950	2,750	5.00	1.45	3.55	6.50
4	2	600	800	2,600	3,400	5.67	1.33	4.33	13.00
5	2	625	800	3,250	4,050	6.48	1.28	5.2	26.00
6	2	640	800	3,900	4,700	7.34	1.25	6.09	43.33

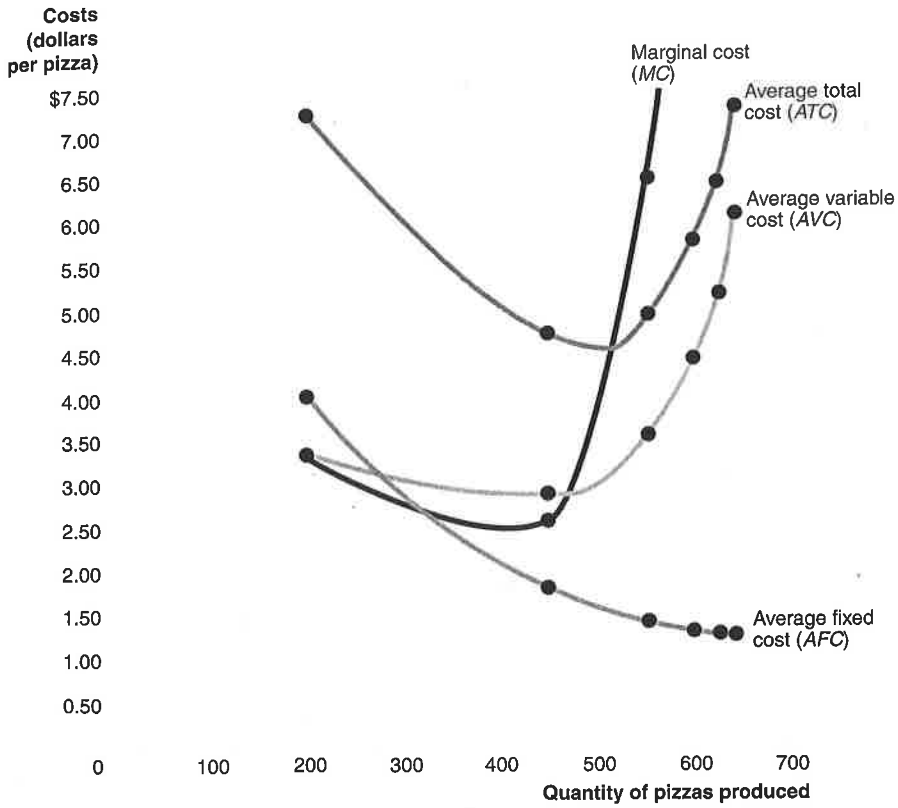


Figure 10-5

Costs of Jill Johnson's Restaurant

Jill's costs of making pizzas are shown in the table and plotted in the graph. Notice three important facts about the graph: (1) The marginal cost (*MC*), average total cost (*ATC*), and average variable cost (*AVC*) curves are all U-shaped, and the marginal cost curve intersects both the average variable cost curve and average total cost curve at their minimum points. (2) As output increases, average fixed cost (*AFC*) gets smaller and smaller. (3) As output increases, the difference between average total cost and average variable cost decreases. Make sure you can explain why each of these three facts is true. You should spend time becoming familiar with this graph because it is one of the most important graphs in microeconomics.

more ovens. The same would be true of any other fixed costs a company like Jill's might have. Once a company has purchased a fire insurance policy, the cost of the policy is fixed. But when the policy expires, the company must decide whether to renew it, and the cost becomes variable. The important point here is this: *In the long run, all costs are variable. There are no fixed costs in the long run.* In other words, in the long run, total cost equals variable cost, and average total cost equals average variable cost.

Managers of successful firms simultaneously consider how they can most profitably run their current store, factory, or office and also whether in the long run they would be more profitable if they became larger or, possibly, smaller. Jill must consider how to run her current restaurant, which has only two pizza ovens, and she must also plan what to do when her current bank loan is paid off and the lease on her store ends. Should she buy more pizza ovens? Should she lease a larger restaurant?

Economies of Scale

Short-run average cost curves represent the costs a firm faces when some input, such as the quantity of machines it uses, is fixed. The **long-run average cost curve** shows the lowest cost at which a firm is able to produce a given level of output in the long run, when no inputs are fixed. Many firms experience **economies of scale**, which means the

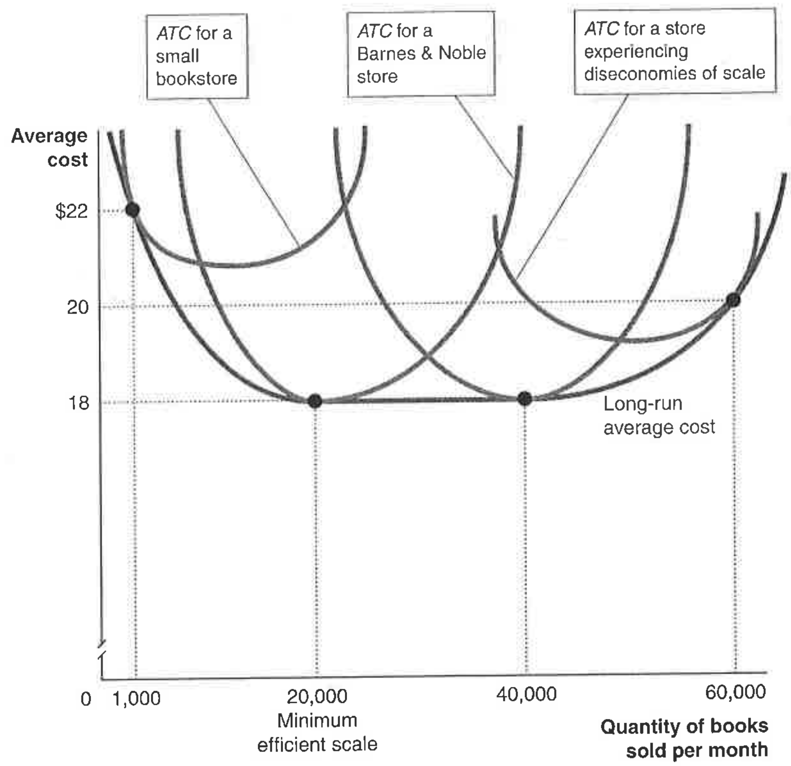
Long-run average cost curve A curve showing the lowest cost at which a firm is able to produce a given quantity of output in the long run, when no inputs are fixed.

Economies of scale The situation when a firm's long-run average costs fall as it increases output.

Figure 10-6

The Relationship between Short-Run Average Cost and Long-Run Average Cost

If a small bookstore expects to sell only 1,000 books per month, then it will be able to sell that quantity of books at the lowest average cost of $22 per book if it builds the small store represented by the *ATC* curve on the left of the figure. A larger bookstore will be able to sell 20,000 books per month at a lower cost of $18 per book. A bookstore selling 20,000 books per month and a bookstore selling 40,000 books per month will experience constant returns to scale and have the same average cost. A bookstore selling 20,000 books per month will have reached minimum efficient scale. Very large bookstores will experience diseconomies of scale, and their average costs will rise as sales increase beyond 40,000 books per month.



firm's long-run average costs fall as it increases the quantity of output it produces. We can see the effects of economies of scale in Figure 10-6, which shows the relationship between short-run and long-run average cost curves. Managers can use long-run average cost curves for planning because they show the effect on cost of expanding output by, for example, building a larger factory or store.

Long-Run Average Total Cost Curves for Bookstores

Figure 10-6 shows long-run average cost in the retail bookstore industry. If a small bookstore expects to sell only 1,000 books per month, then it will be able to sell that quantity of books at the lowest average cost of $22 per book if it builds the small store represented by the *ATC* curve on the left of the figure. A much larger bookstore, such as one run by a national chain like Barnes & Noble, will be able to sell 20,000 books per month at a lower average cost of $18 per book. This decline in average cost from $22 to $18 represents the economies of scale that exist in bookselling. Why would the larger bookstore have lower average costs? One important reason is that the Barnes & Noble store is selling 20 times as many books per month as the small store but might need only six times as many workers. This saving in labor cost would reduce Barnes & Noble's average cost of selling books.

Firms may experience economies of scale for several reasons. First, as in the case of Barnes & Noble, the firm's technology may make it possible to increase production with a smaller proportional increase in at least one input. Second, both workers and managers can become more specialized, enabling them to become more productive, as output expands. Third, large firms, like Barnes & Noble, Wal-Mart, and General Motors, may be able to purchase inputs at lower costs than smaller competitors. In fact, as Wal-Mart expanded, its bargaining power with its suppliers increased, and its average costs fell. Finally, as a firm expands, it may be able to borrow money more inexpensively, thereby lowering its costs.

Economies of scale do not continue forever. The long-run average cost curve in most industries has a flat segment that often stretches over a substantial range of output. As Figure 10-6 shows, a bookstore selling 20,000 books per month and a bookstore selling 40,000 books per month have the same average cost. Over this range of output, firms in the industry experience **constant returns to scale**. As these firms increase their output, they have to increase their inputs, such as the size of the store and the quantity of

Constant returns to scale The situation when a firm's long-run average costs remain unchanged as it increases output.

workers, proportionally. The level of output at which all economies of scale are exhausted is known as **minimum efficient scale**. A bookstore selling 20,000 books per month has reached minimum efficient scale.

Very large bookstores experience increasing average costs as managers begin to have difficulty coordinating the operation of the store. Figure 10-6 shows that for sales above 40,000 books per month, firms in the industry experience **diseconomies of scale**. Toyota ran into diseconomies of scale in assembling automobiles. The firm found that as it expanded production at its Georgetown, Kentucky, plant and its plants in China, its managers had difficulty keeping costs from rising. The president of Toyota's Georgetown plant was quoted as saying, "Demand for . . . high volumes saps your energy. Over a period of time, it eroded our focus . . . [and] thinned out the expertise and knowledge we painstakingly built up over the years." One analysis of the problems Toyota faced in expanding production concluded: "It is the kind of paradox many highly successful companies face: Getting bigger doesn't always mean getting better."

Minimum efficient scale The level of output at which all economies of scale are exhausted.

Diseconomies of scale The situation when a firm's long-run average costs rise as the firm increases output.

Solved Problem | 10-6

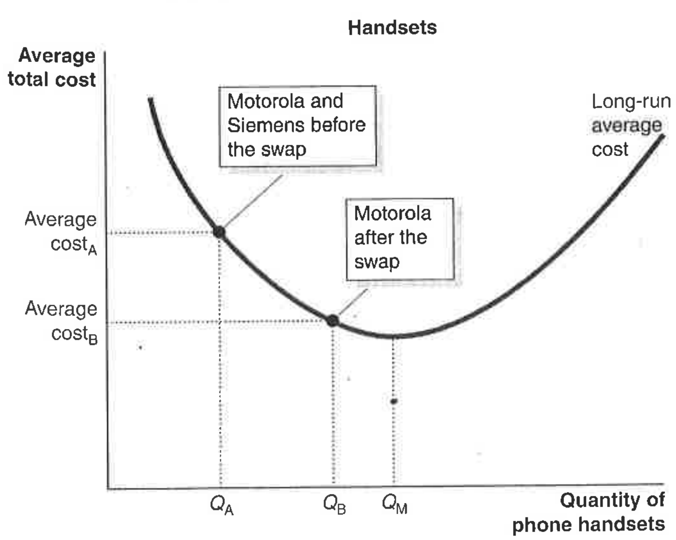
Using Long-Run Average Cost Curves to Understand Business Strategy

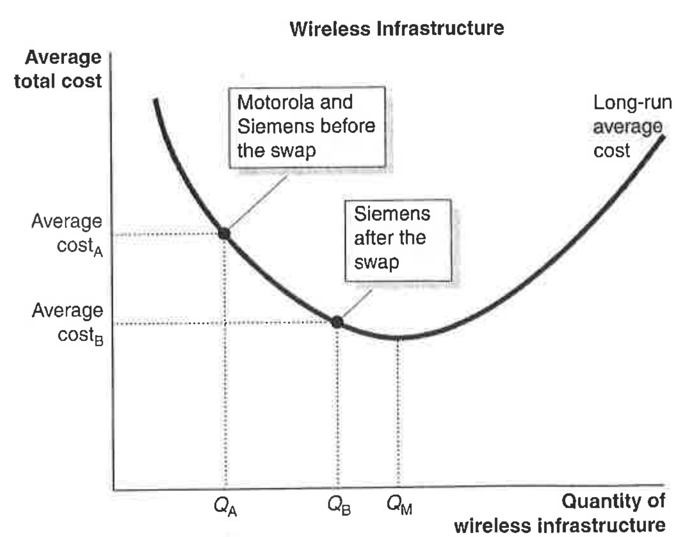
In fall 2002, Motorola and Siemens were each manufacturing both mobile phone handsets and wireless infrastructure—the base stations needed to operate a wireless communications network. The firms discussed the following arrangement: Motorola would give Siemens its wireless infrastructure business in exchange for Siemens giving Motorola its mobile phone handsets business. The main factor motivating the trade was the hope of taking advantage of economies of scale in each business. Use long-run average total cost curves to explain why this trade might make sense for Motorola and Siemens.

SOLVING THE PROBLEM:

Step 1: Review the chapter material. This problem is about the long-run average cost curve, so you may want to review the material in the section "Costs in the Long Run," which begins on page 346.

Step 2: Draw long-run average cost graphs for Motorola and Siemens. The question does not provide us with the details of the quantity of each product each firm is producing before the trade or the firms' average costs of production. If economies of scale were an important reason for the trade, we can assume that Motorola and Siemens were not yet at minimum efficient scale in the wireless infrastructure and phone handset businesses. Therefore, we can draw the following graphs:





Step 3: Explain the curves in the graphs. Before the proposed trade, Motorola and Siemens are producing both products at less than the minimum efficient scale, which is Q_M in both graphs. After the trade, Motorola's production of handsets will increase, moving it from Q_A to Q_B in the first graph. This increase in production will allow it to take advantage of economies of scale and reduce its average cost from Average $Cost_A$ to Average $Cost_B$. Similarly, production of wireless infrastructure by Siemens will increase from Q_A to Q_B, lowering its average cost from Average $Cost_A$ to Average $Cost_B$. As drawn, the graphs show that both firms will still be short of minimum efficient scale after the trade, although their average costs will have fallen.

EXTRA CREDIT: These were new technologies at the time Motorola and Siemens discussed the trade. As a result, companies making these products were only beginning to understand how large minimum efficient scale was. To survive in the industry, the managements of both companies wanted to lower their costs by taking advantage of economies of scale. As one industry analyst put it: "Motorola and Siemens may be driven by the conviction that they have little choice. Most observers believe consolidation in both the [wireless] networking and handset areas is inevitable."

Source for quote: Ray Hegarty, *Rumored Motorola–Siemens Business Unit Swap? A Compelling M&A Story*, www.thefeature.com.

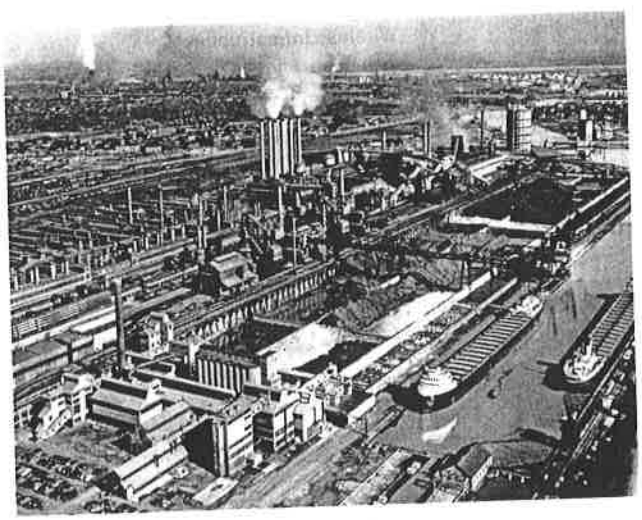
YOUR TURN: For more practice, do related problems 6.4, 6.5, 6.6, and 6.7 on pages 361 and 362 at the end of this chapter.

>> End Solved Problem 10-6

Over time, most firms in an industry will build factories or stores that are at least as large as the minimum efficient scale but not so large that diseconomies of scale occur. In the bookstore industry, stores will sell between 20,000 and 40,000 books per month. However, firms often do not know the exact shape of their long-run average cost curves. As a result, they may mistakenly build factories or stores that are either too large or too small.

Making the Connection | The Colossal River Rouge: Diseconomies of Scale at Ford Motor Company

When Henry Ford started the Ford Motor Company in 1903, automobile companies produced cars in small workshops, using highly skilled workers. Ford introduced two new ideas that allowed him to take advantage of economies of scale. First, Ford used identical—or, interchangeable—parts so that unskilled workers could assemble the cars. Second, instead of having groups of workers moving from one stationary automobile to the next, he had the workers remain stationary while the automobiles moved along an assembly line. Ford built a large factory at Highland Park, outside Detroit, where he used these ideas to produce the famous Model T at an average cost well below what his competitors could match using older production methods in smaller factories.



Is it possible for a factory to be too big?

Ford believed that he could produce automobiles at an even lower average cost by building a still larger plant along the River Rouge. Unfortunately, Ford's River Rouge plant was too large and suffered from diseconomies of scale. Ford's managers had great difficulty coordinating the production of automobiles in such a large plant. The following description of the River Rouge comes from a biography of Ford by Allan Nevins and Frank Ernest Hill:

> A total of 93 separate structures stood on the [River Rouge] site. . . . Railroad trackage covered 93 miles, conveyors 27 [miles]. About 75,000 men worked in the great plant. A force of 5000 did

> nothing but keep it clean, wearing out 5000 mops and 3000 brooms a month, and using 86 tons of soap on the floors, walls, and 330 acres of windows. The Rouge was an industrial city, immense, concentrated, packed with power. . . . By its very massiveness and complexity, it denied men at the top contact with and understanding of those beneath, and gave those beneath a sense of being lost in inexorable immensity and power.

Beginning in 1927, Ford produced the Model A—its only car model at that time—at the River Rouge plant. Ford failed to achieve economies of scale and actually *lost money* on each of the four Model A body styles.

Ford could not raise the price of the Model A to make it profitable because at a higher price, the car could not compete with similar models produced by competitors such as General Motors and Chrysler. He eventually reduced the cost of making the Model A by constructing smaller factories spread out across the country. These smaller factories produced the Model A at a lower average cost than was possible at the River Rouge plant.

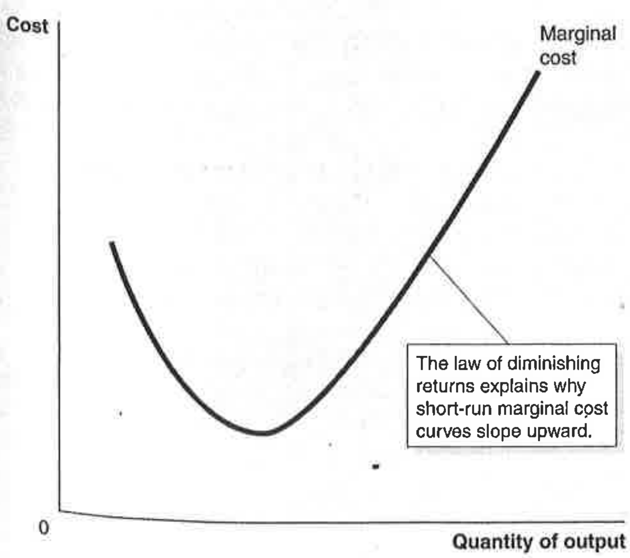
Source for quote: Allan Nevins and Frank Ernest Hill, *Ford: Expansion and Challenge, 1915–1933*, New York: Scribner, 1957, pp. 293, 295.

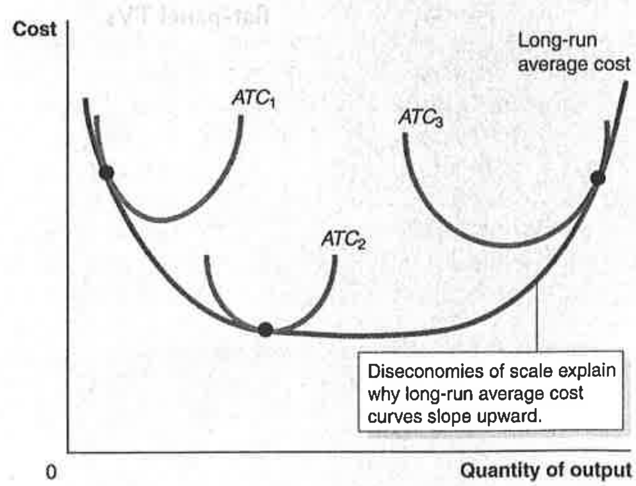
YOUR TURN: Test your understanding by doing related problem 6.8 on page 362 at the end of this chapter.

Don't Let This Happen to **YOU!**

DON'T CONFUSE DIMINISHING RETURNS WITH DISECONOMIES OF SCALE

The concepts of diminishing returns and diseconomies of scale may seem similar, but, in fact, they are unrelated. Diminishing returns applies only to the short run, when at least one of the firm's inputs, such as the quantity of machinery it uses, is fixed. The law of diminishing returns tells us that in the short run, hiring more workers will, at some point, result in less additional output. Diminishing returns explains why marginal cost curves eventually slope upward. Diseconomies of scale apply only in the long run, when the firm is free to vary all its inputs, can adopt new technology, and can vary the amount of machinery it uses and the size of its facility. Diseconomies of scale explain why long-run average cost curves eventually slope upward.





YOUR TURN: Test your understanding by doing related problem 6.10 on page 362 at the end of this chapter.