section 10

Behind the Supply Curve: Profit, Production, and Costs

In Section 9 we examined the factors that affect consumer choice—the demand side of the supply and demand model. In this section we turn our attention to the factors that affect producer choice and the supply side of the supply and demand model. We'll begin with the concept of profit and examine profit maximization as the goal of a firm. We will then investigate the firm's production function, which shows the relationship between the inputs used for production and the output that is produced. Next we'll consider the costs that influence firms' decisions about supply. The final module in this section introduces the models of market structure used to understand how the supply side of the economy works.

Brand-X Pictures

Module 52 Defining Profit

What you will learn in this Module:

- The difference between explicit and implicit costs and their importance in decision making
- The different types of profit, including economic profit, accounting profit, and normal profit
- How to calculate profit

Understanding Profit

The primary goal of most firms is to maximize profit. Other goals, such as maximizing market share or protecting the environment, may also figure into a firm's mission. But economic models generally start with the assumption that firms attempt to maximize profit. So we will begin with an explanation of how economists define and calculate profit. In the next module we will look at how firms go about maximizing their profit.

In general, a firm's profit equals its *total revenue*—which is equal to the price of the output times the quantity sold, or $P \times Q$—minus the cost of all the inputs used to produce its output, its *total cost*. That is,

$$\text{Profit} = \text{Total Revenue} - \text{Total Cost}$$

However, there are different types of costs that may be used to calculate different types of profit. To start the discussion of how to calculate profit, we'll look at two different types of costs, *explicit costs* and *implicit costs*.

Explicit versus Implicit Costs

Suppose that, after graduating from high school, you have two options: to go to college or to take a job immediately. You would like to continue your education but are concerned about the cost.

But what exactly is the cost of attending college? Here is where it is important to remember the concept of opportunity cost: the cost of the time spent getting a degree is what you forgo by not taking a job for the years you go to college. The opportunity cost of additional education, like any cost, can be broken into two parts: the *explicit cost* and the *implicit cost.*

An **explicit cost** is a cost that requires an outlay of money. For example, the explicit cost of a year of college includes tuition. An **implicit cost,** though, does not involve an outlay of money; instead, it is measured by the value, in dollar terms, of the benefits that are forgone. For example, the implicit cost of a year spent in college includes the income you would have earned if you had taken a job instead.

A common mistake, both in economic analysis and in real business situations, is to ignore implicit costs and focus exclusively on explicit costs. But often the implicit cost

An **explicit cost** is a cost that involves actually laying out money. An **implicit cost** does not require an outlay of money; it is measured by the value, in dollar terms, of benefits that are forgone.

of an activity is quite substantial—indeed, sometimes it is much larger than the explicit cost.

Table 52.1 gives a breakdown of hypothetical explicit and implicit costs associated with spending a year in college instead of taking a job. The explicit cost consists of tuition, books, supplies, and a computer for doing assignments—all of which require you to spend money. The implicit cost is the salary you would have earned if you had taken a job instead. As you can see, the forgone salary is $35,000 and the explicit cost is $19,500, making the implicit cost more than the explicit cost in this example. So ignoring the implicit cost of an action can lead to a seriously misguided decision.

The **accounting profit** of a business is the business's total revenue minus the explicit cost and depreciation.

table **52.1**

Opportunity Cost of an Additional Year of School

Explicit cost		Implicit cost	
Tuition	$17,000	Forgone salary	$35,000
Books and supplies	1,000		
Computer	1,500		
Total explicit cost	**19,500**	**Total implicit cost**	**35,000**
Total opportunity cost = Total explicit cost + Total implicit cost = $54,500			

A slightly different way of looking at the implicit cost in this example can deepen our understanding of opportunity cost. The forgone salary is the cost of using your own resources—your time—in going to college rather than working. The use of your *time* for more education, despite the fact that you don't have to spend any money, is still costly to you. This illustrates an important aspect of opportunity cost: in considering the cost of an activity, you should include the cost of using any of your own resources for that activity. You can calculate the cost of using your own resources by determining what they would have earned in their next best alternative use.

Accounting Profit versus Economic Profit

As the example of going to college suggests, taking account of implicit as well as explicit costs can be very important when making decisions. This is true whether the decisions affect individuals, groups, governments, or businesses.

Consider the case of Babette's Cajun Café, a small restaurant in New Orleans. This year Babette brought in $100,000 in revenue. Out of that revenue, she paid her expenses: the cost of food ingredients and other supplies, the cost of wages for her employees, and the rent for her restaurant space. This year her expenses were $60,000. We assume that Babette owns her restaurant equipment—items such as appliances and furnishings. The question is: Is Babette's restaurant profitable?

Alamy

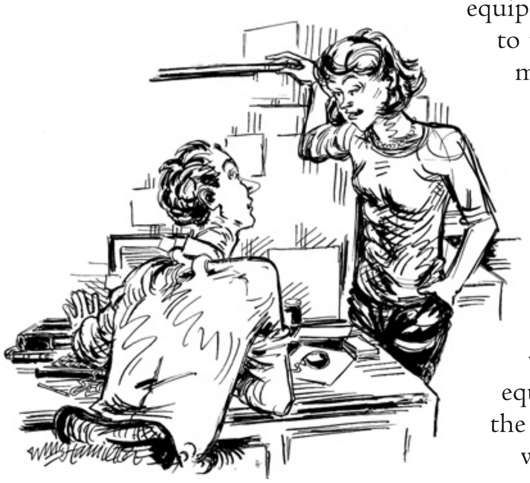
At first it might seem that the answer is obviously yes: she receives $100,000 from her customers and has expenses of only $60,000. Doesn't this mean that she has a profit of $40,000? Not according to her accountant, who reduces the number by $5,000 for the yearly *depreciation* (reduction in value) of the restaurant equipment. Depreciation occurs because equipment wears out over time. As a consequence, every few years Babette must replace her appliances and furnishings. The yearly depreciation amount reflects what an accountant estimates to be the reduction in the value of the machines due to wear and tear that year. This leaves $35,000, which is the business's **accounting profit.** That is, The **economic profit** of a business is the business's total revenue minus the opportunity cost of its resources. It is usually less than the accounting profit.

The **implicit cost of capital** is the opportunity cost of the capital used by a business—the income the owner could have realized from that capital if it had been used in its next best alternative way.

the accounting profit of a business is its total revenue minus its *explicit* cost and depreciation. The accounting profit is the number that Babette has to report on her income tax forms and that she would be obliged to report to anyone thinking of investing in her business.

Accounting profit is a very useful number, but suppose that Babette wants to decide whether to keep her restaurant open or do something else. To make this decision, she will need to calculate her **economic profit**—the total revenue she receives minus her *opportunity* cost, which includes implicit as well as explicit costs. In general, when economists use the simple term *profit,* they are referring to economic profit. (We adopt this simplification in this book.)

Why does Babette's economic profit differ from her accounting profit? Because she may have an implicit cost over and above the explicit cost her accountant has calculated. Businesses can face an implicit cost for two reasons. First, a business's capital—its equipment, buildings, tools, inventory, and financial assets—could have been put to use in some other way. If the business owns its capital, it does not pay any money for its use, but it pays an implicit cost because it does not use the capital in some other way. Second, the owner devotes time and energy to the business that could have been used elsewhere—a particularly important factor in small businesses, whose owners tend to put in many long hours.



"I've done the numbers, and I will marry you."

If Babette had rented her appliances and furnishings instead of owning them, her rent would have been an explicit cost. But because Babette owns her own equipment, she does not pay rent on them and her accountant deducts an estimate of their depreciation in the profit statement. However, this does not account for the opportunity cost of the equipment—what Babette forgoes by owning it. Suppose that instead of using the equipment in her own restaurant, the best alternative Babette has is to sell the equipment for $50,000 and put the money into a bank account where it would earn yearly interest of $3,000. This $3,000 is an implicit cost of running the business. The **implicit cost of capital** is the opportunity cost of the capital used by a business; it reflects the income that could have been earned if the capital had been used in its next best alternative way. It is just as much a true cost as if Babette had rented her equipment instead of owning it.

Finally, Babette should take into account the opportunity cost of her own time. Suppose that instead of running her own restaurant, she could earn $34,000 as a chef in someone else's restaurant. That $34,000 is also an implicit cost of her business.

Table 52.2, in the column titled Case 1, summarizes the accounting for Babette's Cajun Café, taking both explicit and implicit costs into account. It turns out, unfortunately, that

table **52.2**

Profit at Babette's Cajun Café

	Case 1	Case 2
Revenue	$100,000	$100,000
Explicit cost	–60,000	–60,000
Depreciation	–5,000	–5,000
Accounting profit	35,000	35,000
Implicit cost of business		
Income Babette could have earned on capital used in the next best way	–3,000	–3,000
Income Babette could have earned as a chef in someone else's restaurant	–34,000	–30,000
Economic profit	–2,000	+2,000

fyi

Farming in the Shadow of Suburbia

Beyond the sprawling suburbs, most of New England is covered by dense forest. But this is not the forest primeval: if you hike through the woods, you encounter many stone walls, relics of the region's agricultural past when stone walls enclosed fields and pastures. In 1880, more than half of New England's land was farmed; by 2009, the amount was down to 10%.

The remaining farms of New England are mainly located close to large metropolitan areas. There farmers get high prices for their produce from city dwellers who are willing to pay a premium for locally grown, extremely fresh fruits and vegetables.

But now even these farms are under economic pressure caused by a rise in the implicit cost of farming close to a metropolitan area. As metropolitan areas have expanded during the last two decades, farmers increasingly ask themselves whether they could do better by selling their land to property developers.


Michael S. Lewis/National Geographic/Getty Images

In 2009, the average value of an acre of farmland in the United States as a whole was $2,100; in Rhode Island, the most densely populated of the New England states, the average was $15,300. The Federal Reserve Bank of Boston has noted that "high land prices put intense pressure on the region's farms to generate incomes that are substantial enough to justify keeping the land in agriculture." The important point is that the pressure is intense even if the farmer owns the land because the land is a form of capital used to run the business.

Maintaining the land as a farm instead of selling it to a developer constitutes a large implicit cost of capital. A fact provided by the U.S. Department of Agriculture (USDA) helps us put a dollar figure on the portion of the implicit cost of capital due to development pressure for some Rhode Island farms. In 2004, a USDA program designed to prevent development of Rhode Island farmland by paying owners for the "development rights" to their land paid an average of $4,949 per acre for those rights alone. By 2009, the amount had risen to $15,357.

About two-thirds of New England's farms remaining in business earn very little money. They are maintained as "rural residences" by people with other sources of income—not so much because they are commercially viable, but more out of a personal commitment and the satisfaction these people derive from farm life. Although many businesses have important implicit costs, they can also have important benefits to their owners that go beyond the revenue earned.

although the business makes an accounting profit of $35,000, its economic profit is actually negative. This means that Babette would be better off financially if she closed the restaurant and devoted her time and capital to something else. If, however, some of Babette's cost should fall sufficiently, she could earn a positive economic profit. In that case, she would be better off financially if she continued to operate the restaurant. For instance, consider the column titled Case 2: here we assume that what Babette could earn as a chef employed by someone else has dropped to $30,000 (say, due to a soft labor market). In this case, her economic profit is positive: she is earning more than her explicit and implicit costs and she should keep her restaurant open.

In real life, discrepancies between accounting profit and economic profit are extremely common. As the FYI above explains, this is a message that has found a receptive audience among real-world businesses.

Normal Profit

In the example above, when Babette is earning an economic profit, her total revenue is higher than the sum of her implicit and explicit costs. This means that operating her restaurant makes Babette better off financially than she would be using her resources in any other activity. When Babette earns a negative economic profit (which can also be described as a *loss*), it means that Babette would be better off financially if she devoted her resources to her next best alternative. As this example illustrates, economic profits signal the best use of resources. A positive economic profit indicates that the current use is the best use of resources. A negative economic profit indicates that there is a better alternative use for resources.

An economic profit equal to zero is also known as a **normal profit.** It is an economic profit just high enough to keep a firm engaged in its current activity.

But what about an economic profit *equal to* zero? Most of us would generally think earning zero profit was a bad thing. After all, a firm's goal is to maximize profit—profit is what firms are after! However, an economic profit equal to zero is not bad at all. An economic profit of zero means that the firm could not do any better using its resources in any alternative activity. Another name for an economic profit of zero is a **normal profit.** A firm earning a normal profit is earning just enough to keep it using its resources in its current activity. After all, it can't do any better in any other activity!

Module 52 AP Review

Solutions appear at the back of the book.

Check Your Understanding

1. Karma and Don run a furniture-refinishing business from their home. Which of the following represent an explicit cost of the business and which represent an implicit cost?
   a. supplies such as paint stripper, varnish, polish, sandpaper, and so on
   b. basement space that has been converted into a workroom
   c. wages paid to a part-time helper
   d. a van that they inherited and use only for transporting furniture
   e. the job at a larger furniture restorer that Karma gave up in order to run the business
2. a. Suppose you are in business earning an accounting profit of $25,000. What is your economic profit if the implicit cost of your capital is $2,000 and the opportunity cost of your time is $23,000? Explain your answer.
   b. What does your answer to part a tell you about the advisability of devoting your time and capital to this business?

Tackle the Test: Multiple-Choice Questions

1. Which of the following is an example of an *implicit* cost of going out for lunch?
   a. the amount of the tip you leave the waiter
   b. the total bill you charge to your credit card
   c. the cost of gas to drive to the restaurant
   d. the value of the time you spent eating lunch
   e. all of the above
2. Which of the following is an *implicit* cost of attending college?
   a. tuition
   b. books
   c. laptop computer
   d. lab fees
   e. forgone salary
3. Which of the following is the best definition of accounting profit? Accounting profit equals total revenue minus depreciation and total
   a. explicit cost only.
   b. implicit cost only.
   c. explicit cost plus implicit cost.
   d. opportunity cost.
   e. explicit cost plus opportunity cost.
4. Which of the following is considered when calculating economic profit but not accounting profit?
   a. implicit cost
   b. explicit cost
   c. total revenue
   d. marginal cost
   e. All of the above are considered when calculating accounting profit.
5. You sell T-shirts at your school's football games. Each shirt costs $5 to make and sells for $10. Each game lasts two hours and you sell 100 shirts per game. You could always be earning $8 per hour at your other job. Which of the following is correct? Your accounting profit from selling shirts at a game is
   a. $1,000 and your economic profit is $500.
   b. $500 and your economic profit is $1,000.
   c. $500 and your economic profit is $484.
   d. $484 and your economic profit is $500.
   e. $500 and your economic profit is also $500.

Tackle the Test: Free-Response Questions

1. Your firm is selling 10,000 units of output at a price of $10 per unit. Your firm's total explicit cost is $70,000. Your firm's implicit cost of capital is $10,000, and your opportunity cost is $20,000.
   a. Calculate total revenue.
   b. Calculate total implicit cost.
   c. Calculate your accounting profit.
   d. Calculate your economic profit.
   e. What does the value of your economic profit calculated in part d tell you?

Answer (5 points)

1 point: Total revenue = $100,000

1 point: Total implicit cost = $30,000

1 point: Accounting profit = $30,000

1 point: Economic profit = $0

1 point: Because your firm earns normal profit, there is no better alternative use for your resources.

2. Sunny owns and operates Sunny's Sno Cone Stand. Use the data in the table provided to answer the questions below.

Sunny's Sno Cone Stand: January

Price of Sno Cone	$2
Sno Cones sold	2,000
Explicit cost	$400
Depreciation	$100
Implicit cost of capital	$200

   a. Calculate Sunny's Sno Cone Stand's total revenue for January.
   b. Calculate Sunny's Sno Cone Stand's accounting profit for January.
   c. What additional information would Sunny need in order to determine whether or not to continue operating the Sno Cone Stand?
   d. Explain how Sunny would determine whether or not to continue operating the business on the basis of these numbers.

What you will learn in this **Module:**

- The principle of marginal analysis
- How to determine the profit-maximizing level of output using the optimal output rule

Module 53 Profit Maximization

Maximizing Profit

In the previous module we learned about different types of profit, how to calculate profit, and how firms can use profit calculations to make decisions—for instance to determine whether to continue using resources for the same activity or not. In this module we ask the question: what quantity of output would maximize the producer's profit? First we will find the profit-maximizing quantity by calculating the total profit at each quantity for comparison. Then we will use marginal analysis to determine the *optimal output rule,* which turns out to be simple: as our discussion of marginal analysis in Module 1 suggested, a producer should produce up until marginal benefit equals marginal cost.

Consider Jennifer and Jason, who run an organic tomato farm. Suppose that the market price of organic tomatoes is $18 per bushel and that Jennifer and Jason can sell as many as they would like at that price. Then we can use the data in Table 53.1 to find their profit-maximizing level of output.

The first column shows the quantity of output in bushels, and the second column shows Jennifer and Jason's total revenue from their output: the market value of their output. Total revenue, *TR,* is equal to the market price multiplied by the quantity of output:

(53-1) $TR = P \times Q$

In this example, total revenue is equal to $18 per bushel times the quantity of output in bushels.

The third column of Table 53.1 shows Jennifer and Jason's total cost, *TC.* The fourth column shows their profit, equal to total revenue minus total cost:

(53-2) $\text{Profit} = TR - TC$

As indicated by the numbers in the table, profit is maximized at an output of five bushels, where profit is equal to $18. But we can gain more insight into the profit-maximizing choice of output by viewing it as a problem of marginal analysis, a task we'll dive into next.

table **53.1**

Profit for Jennifer and Jason's Farm When Market Price Is $18

Quantity of tomatoes *Q* (bushels)	Total revenue *TR*	Total cost *TC*	Profit *TR – TC*
0	$0	$14	–$14
1	18	30	–12
2	36	36	0
3	54	44	10
4	72	56	16
5	90	72	18
6	108	92	16
7	126	116	10

Using Marginal Analysis to Choose the Profit-Maximizing Quantity of Output

The **principle of marginal analysis** provides a clear message about when to stop doing anything: proceed until *marginal benefit* equals *marginal cost.* To apply this principle, consider the effect on a producer's profit of increasing output by one unit. The marginal benefit of that unit is the additional revenue generated by selling it; this measure has a name—it is called the **marginal revenue** of that output. The general formula for marginal revenue is:

$$\textbf{(53-3)}\quad \text{Marginal revenue} = \begin{array}{l}\text{Change in total}\\ \text{revenue generated}\\ \text{by one additional}\\ \text{unit of output}\end{array} = \frac{\text{Change in total revenue}}{\text{Change in quantity of output}}$$

or

$$MR = \Delta TR/\Delta Q$$

In this equation, the Greek uppercase delta (the triangular symbol) represents the change in a variable.

The application of the principle of marginal analysis to the producer's decision of how much to produce is called the **optimal output rule,** which states that profit is maximized by producing the quantity at which the marginal revenue of the last unit produced is equal to its marginal cost. As this rule suggests, we will see that Jennifer and Jason maximize their profit by equating marginal revenue and marginal cost.

Note that there may not be any particular quantity at which marginal revenue exactly equals marginal cost. In this case the producer should produce until one more unit would cause marginal benefit to fall below marginal cost. As a common simplification, we can think of marginal cost as rising steadily, rather than jumping from one level at one quantity to a different level at the next quantity. This ensures that marginal cost will equal marginal revenue at some quantity. We employ this simplified approach in what follows.

Consider Table 53.2 on the next page, which provides cost and revenue data for Jennifer and Jason's farm. The second column contains the farm's total cost of output.

According to the **principle of marginal analysis,** every activity should continue until marginal benefit equals marginal cost.

Marginal revenue is the change in total revenue generated by an additional unit of output.

The **optimal output rule** says that profit is maximized by producing the quantity of output at which the marginal revenue of the last unit produced is equal to its marginal cost.

table 53.2

Short-Run Costs for Jennifer and Jason's Farm

Quantity of tomatoes *Q* (bushels)	Total cost *TC*	Marginal cost of bushel $MC = \Delta TC/\Delta Q$	Marginal revenue of bushel *MR*	Net gain of bushel = $MR - MC$
0	$14			
		$16	$18	$2
1	30			
		6	18	12
2	36			
		8	18	10
3	44			
		12	18	6
4	56			
		16	18	2
5	72			
		20	18	–2
6	92			
		24	18	–6
7	116			

The third column shows their marginal cost. Notice that, in this example, marginal cost initially falls as output rises but then begins to increase, so that the marginal cost curve has a "swoosh" shape. (Later it will become clear that this shape has important implications for short-run production decisions.)

The fourth column contains the farm's marginal revenue, which has an important feature: Jennifer and Jason's marginal revenue is assumed to be constant at $18 for every output level. The assumption holds true for a particular type of market—perfectly competitive markets—which we will study in Modules 58–60, but for now it is just to make the calculations easier. The fifth and final column shows the calculation of the net gain per bushel of tomatoes, which is equal to marginal revenue minus marginal cost. As you can see, it is positive for the first through fifth bushels; producing each of these bushels raises Jennifer and Jason's profit. For the sixth and seventh bushels, however, net gain is negative: producing them would decrease, not increase, profit. (You can verify this by reexamining Table 53.1.) So five bushels are Jennifer and Jason's profit-maximizing output; it is the level of output at which marginal cost is equal to the market price, $18.

Figure 53.1 shows that Jennifer and Jason's profit-maximizing quantity of output is, indeed, the number of bushels at which the marginal cost of production is equal to marginal revenue (which is equivalent to price in perfectly competitive markets). The figure shows the **marginal cost curve,** *MC,* drawn from the data in the third column of Table 53.2. We plot the marginal cost of increasing output from one to two bushels halfway between one and two, and so on. The horizontal line at $18 is Jennifer and Jason's **marginal revenue curve.** Note that marginal revenue stays the same regardless of how much Jennifer and Jason sell because we have assumed marginal revenue is constant.

Does this mean that the firm's production decision can be entirely summed up as "produce up to the point where the marginal cost of production is equal to the price"? No, not quite. Before applying the principle of marginal analysis to determine how much to produce, a potential producer must, as a first step, answer an "either–or" question: Should I produce at all? If the answer to that question is yes, the producer then proceeds to the second step—a "how much" decision: maximizing profit by choosing the quantity of output at which marginal cost is equal to price.

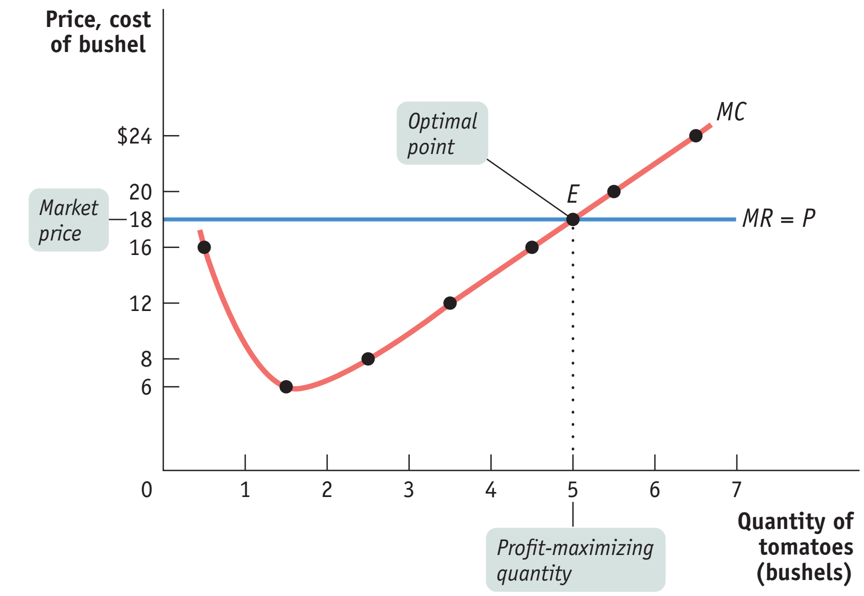
The **marginal cost curve** shows how the cost of producing one more unit depends on the quantity that has already been produced.

The **marginal revenue curve** shows how marginal revenue varies as output varies.

figure 53.1

The Firm's Profit-Maximizing Quantity of Output

At the profit-maximizing quantity of output, marginal revenue is equal to marginal cost. It is located at the point where the marginal cost curve crosses the marginal revenue curve, which is a horizontal line at the market price. Here, the profit-maximizing point is at an output of 5 bushels of tomatoes, the output quantity at point *E.*



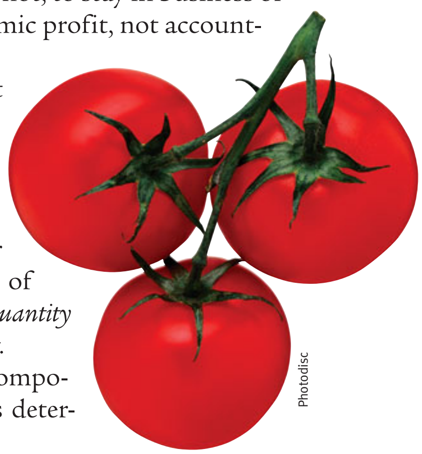
To understand why the first step in the production decision involves an "either–or" question, we need to ask how we determine whether it is profitable or unprofitable to produce at all.

When Is Production Profitable?

Recall that a firm's decision whether or not to stay in a given business depends on its *economic profit*—a measure based on the opportunity cost of resources used in the business. To put it a slightly different way: in the calculation of economic profit, a firm's total cost incorporates the implicit cost—the benefits forgone in the next best use of the firm's resources—as well as the explicit cost in the form of actual cash outlays. In contrast, *accounting profit* is profit calculated using only the explicit costs incurred by the firm. This means that economic profit incorporates the opportunity cost of resources owned by the firm and used in the production of output, while accounting profit does not. As in the example of Babette's Cajun Café, a firm may make positive accounting profit while making zero or even negative economic profit. It's important to understand clearly that a firm's decision to produce or not, to stay in business or to close down permanently, should be based on economic profit, not accounting profit.

So we will assume, as we always do, that the cost numbers given in Tables 53.1 and 53.2 include all costs, implicit as well as explicit, and that the profit numbers in Table 53.1 are economic profit. What determines whether Jennifer and Jason's farm earns a profit or generates a loss? The answer is that whether or not it is profitable depends on the market price of tomatoes—specifically, *whether selling the firm's optimal quantity of output at the market price results in at least a normal profit.*

In the next modules, we look in detail at the two components used to calculate profit; firm revenue (which is determined by the level of production) and firm cost.



Photodisc

Module 53 AP Review

Solutions appear at the back of the book.

Check Your Understanding

1. Suppose a firm can sell as many units of output as it wants for a price of $15 per unit and faces total costs as indicated in the table below. Use the optimal output rule to determine the profit-maximizing level of output for the firm.

Q	*TC*
0	$2
1	10
2	20
3	33
4	50
5	71

2. Use the data from Question 1 to graph the firm's *MC* and *MR* curves and show the profit-maximizing level of output.

Tackle the Test: Multiple-Choice Questions

Use the data in the table provided to answer questions 1–3.

Quantity *Q*	*Total Revenue* *TR*	*Total Cost* *TC*
0	$0	$14
1	18	30
2	36	36
3	54	44
4	72	56
5	90	72
6	108	92
7	126	116

1. What is the marginal revenue of the third unit of output?
   a. $8
   b. $14
   c. $18
   d. $44
   e. $54

2. What is the marginal cost of the first unit of output?
   a. $0
   b. $14
   c. $16
   d. $18
   e. $30

3. At what level of output is profit maximized?
   a. 0
   b. 1
   c. 3
   d. 5
   e. 7

4. A firm should continue to produce as long as its
   a. total revenue is less than its total costs.
   b. total revenue is greater than its total explicit costs.
   c. accounting profit is greater than its economic profit.
   d. accounting profit is not negative.
   e. economic profit is at least zero.

5. A firm earns a normal profit when its
   a. accounting profit equals 0.
   b. economic profit is positive.
   c. total revenue equals its total costs.
   d. accounting profit equals its economic profit.
   e. economic profit equals its total explicit and implicit costs.

Tackle the Test: Free-Response Questions

1. Use the data in the table provided.

Quantity *Q*	*Total Revenue* *TR*	*Total Cost* *TC*
0	$0	$7
1	18	23
2	36	29
3	54	37
4	72	49
5	90	65
6	108	87
7	126	112

   a. What is the marginal revenue of the fourth unit?
   b. Calculate profit at a quantity of two. Explain how you calculated the profit.
   c. What is the profit-maximizing level of output? Explain how to use the optimal output rule to determine the profit-maximizing level of output.

Answer (5 points)

1 point: $18

1 point: $7

1 point: $36 – $29 or TR – TC

1 point: 5 units

1 point: The optimal output rule states that profit is maximized when $MC = MR$. Here, MC never exactly equals MR. When this occurs, the firm should produce the largest quantity at which MR exceeds MC. At a quantity of 5, $MC = \$16$ and $MR = \$18$. For the sixth unit, $MC = \$22$ and $MR = \$18$, and because $MC > MR$, the sixth unit would add more to total cost than it would to total revenue, and it therefore should not be produced.

2. Use a graph to illustrate the typical shape of the two curves used to find a firm's profit-maximizing level of output on the basis of the optimal output rule. Assume all units of output can be sold for $5. Indicate the profit-maximizing level of output with a "Q^*" on the appropriate axis. (You don't have enough information to provide a specific numerical answer.)

What you will learn in this **Module:**

- The importance of the firm's production function, the relationship between the quantity of inputs and the quantity of output
- Why production is often subject to diminishing returns to inputs

Module 54 The Production Function

The Production Function

A *firm* produces goods or services for sale. To do this, it must transform inputs into output. The quantity of output a firm produces depends on the quantity of inputs; this relationship is known as the firm's **production function.** As we'll see, a firm's production function underlies its *cost curves.* As a first step, let's look at the characteristics of a hypothetical production function.

Inputs and Output

To understand the concept of a production function, let's consider a farm that we assume, for the sake of simplicity, produces only one output, wheat, and uses only two inputs, land and labor. This particular farm is owned by a couple named George and Martha. They hire workers to do the actual physical labor on the farm. Moreover, we will assume that all potential workers are of the same quality—they are all equally knowledgeable and capable of performing farmwork.

George and Martha's farm sits on 10 acres of land; no more acres are available to them, and they are currently unable to either increase or decrease the size of their farm by selling, buying, or leasing acreage. Land here is what economists call a **fixed input**—an input whose quantity is fixed for a period of time and cannot be varied. George and Martha are, however, free to decide how many workers to hire. The labor provided by these workers is called a **variable input**—an input whose quantity the firm can vary at any time.

In reality, whether or not the quantity of an input is really fixed depends on the time horizon. In the **long run**—that is, given that a long enough period of time has elapsed—firms can adjust the quantity of any input. So there are no fixed inputs in the long run. In contrast, the **short run** is defined as the time period during which at least one input is fixed. Later, we'll look more carefully at the distinction between the short run and the long run. But for now, we will restrict our attention to the short run and assume that at least one input (land) is fixed.

A **production function** is the relationship between the quantity of inputs a firm uses and the quantity of output it produces.

A **fixed input** is an input whose quantity is fixed for a period of time and cannot be varied.

A **variable input** is an input whose quantity the firm can vary at any time.

The **long run** is the time period in which all inputs can be varied.

The **short run** is the time period in which at least one input is fixed.

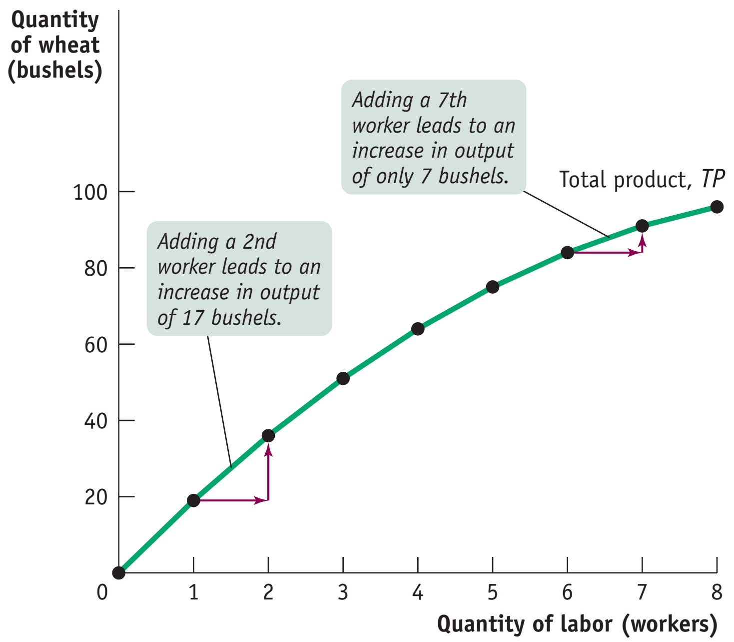
George and Martha know that the quantity of wheat they produce depends on the number of workers they hire. Using modern farming techniques, one worker can cultivate the 10-acre farm, albeit not very intensively. When an additional worker is added, the land is divided equally among all the workers: each worker has 5 acres to cultivate when 2 workers are employed, each cultivates $3\frac{1}{3}$ acres when 3 are employed, and so on. So as additional workers are employed, the 10 acres of land are cultivated more intensively and more bushels of wheat are produced. The relationship between the quantity of labor and the quantity of output, for a given amount of the fixed input, constitutes the farm's production function. The production function for George and Martha's farm, where land is the fixed input and labor is the variable input, is shown in the first two columns of the table in Figure 54.1; the diagram there shows the same information graphically. The curve in Figure 54.1 shows how the quantity of output depends on the quantity of the variable input for a given quantity of the fixed input; it is called the farm's **total product curve.** The physical quantity of output, bushels of wheat, is measured on the vertical axis; the quantity of the variable input, labor (that is, the number of workers employed), is measured on the horizontal axis. The total product curve here slopes upward, reflecting the fact that more bushels of wheat are produced as more workers are employed.

Although the total product curve in Figure 54.1 slopes upward along its entire length, the slope isn't constant: as you move up the curve to the right, it flattens out. To understand this changing slope, look at the third column of the table in Figure 54.1, which shows the *change in the quantity of output* generated by adding one more worker. That is, it shows the **marginal product** of labor, or *MPL*: the additional quantity of output from using one more unit of labor (one more worker).

The **total product curve** shows how the quantity of output depends on the quantity of the variable input, for a given quantity of the fixed input.

The **marginal product** of an input is the additional quantity of output produced by using one more unit of that input.

figure 54.1 Production Function and Total Product Curve for George and Martha's Farm



Quantity of labor *L* (workers)	Quantity of wheat *Q* (bushels)	Marginal product of labor $MPL = \Delta Q/\Delta L$ (bushels per worker)
0	0	
		19
1	19	
		17
2	36	
		15
3	51	
		13
4	64	
		11
5	75	
		9
6	84	
		7
7	91	
		5
8	96	

The table shows the production function, the relationship between the quantity of the variable input (labor, measured in number of workers) and the quantity of output (wheat, measured in bushels) for a given quantity of the fixed input. It also shows the marginal product of labor on George and Martha's farm.

The total product curve shows the production function graphically. It slopes upward because more wheat is produced as more workers are employed. It also becomes flatter because the marginal product of labor declines as more and more workers are employed.

In this example, we have data at intervals of 1 worker—that is, we have information on the quantity of output when there are 3 workers, 4 workers, and so on. Sometimes data aren't available in increments of 1 unit—for example, you might have information on the quantity of output only when there are 40 workers and when there are 50 workers. In this case, you can use the following equation to calculate the marginal product of labor:

$$\text{(54-1)} \quad \begin{matrix}\text{Marginal}\\ \text{product}\\ \text{of labor}\end{matrix} = \begin{matrix}\text{Change in quantity of}\\ \text{output produced by one}\\ \text{additional unit of labor}\end{matrix} = \frac{\text{Change in quantity of output}}{\text{Change in quantity of labor}}$$

or

$$MPL = \frac{\Delta Q}{\Delta L}$$

Recall that Δ, the Greek uppercase delta, represents the change in a variable. Now we can explain the significance of the slope of the total product curve: it is equal to the marginal product of labor. The slope of a line is equal to "rise" over "run." This implies that the slope of the total product curve is the change in the quantity of output (the "rise") divided by the change in the quantity of labor (the "run"). And this, as we can see from Equation 54-1, is simply the marginal product of labor. So in Figure 54.1, the fact that the marginal product of the first worker is 19 also means that the slope of the total product curve in going from 0 to 1 worker is 19. Similarly, the slope of the total product curve in going from 1 to 2 workers is the same as the marginal product of the second worker, 17, and so on.

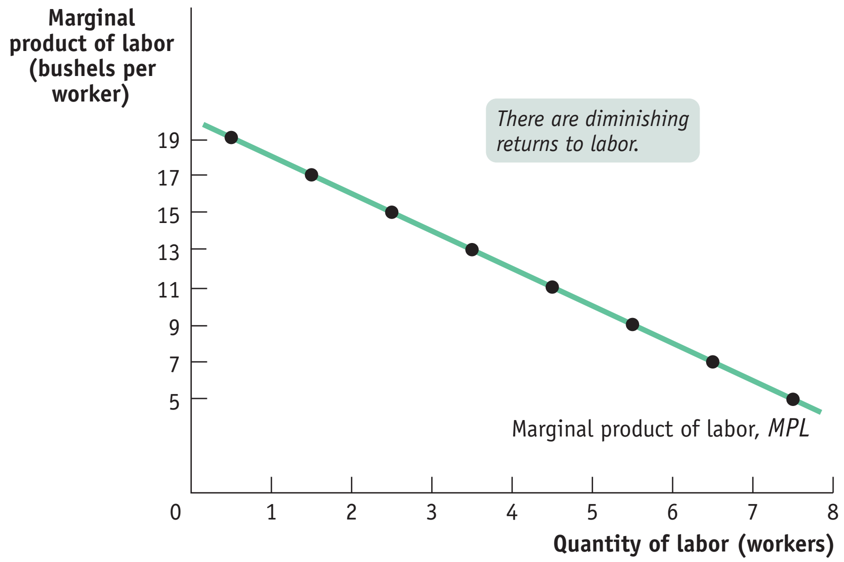
In this example, the marginal product of labor steadily declines as more workers are hired—that is, each successive worker adds less to output than the previous worker. So as employment increases, the total product curve gets flatter.

Figure 54.2 shows how the marginal product of labor depends on the number of workers employed on the farm. The marginal product of labor, *MPL,* is measured on the vertical axis in units of physical output—bushels of wheat—produced per additional worker, and the number of workers employed is measured on the horizontal axis. You can see from the table in Figure 54.1 that if 5 workers are employed instead of 4, output rises from 64 to 75 bushels; in this case the marginal product of labor is

figure 54.2

Marginal Product of Labor Curve for George and Martha's Farm

The marginal product of labor curve plots each worker's marginal product, the increase in the quantity of output generated by each additional worker. The change in the quantity of output is measured on the vertical axis and the number of workers employed on the horizontal axis. The first worker employed generates an increase in output of 19 bushels, the second worker generates an increase of 17 bushels, and so on. The curve slopes downward due to diminishing returns to labor.



11 bushels—the same number found in Figure 54.2. To indicate that 11 bushels is the marginal product when employment rises from 4 to 5, we place the point corresponding to that information halfway between 4 and 5 workers.

In this example the marginal product of labor falls as the number of workers increases. That is, there are *diminishing returns to labor* on George and Martha's farm. In general, there are **diminishing returns to an input** when an increase in the quantity of that input, holding the quantity of all other inputs fixed, reduces that input's marginal product. Due to diminishing returns to labor, the *MPL* curve is negatively sloped.

There are **diminishing returns to an input** when an increase in the quantity of that input, holding the levels of all other inputs fixed, leads to a decline in the marginal product of that input.

To grasp why diminishing returns can occur, think about what happens as George and Martha add more and more workers without increasing the number of acres. As the number of workers increases, the land is farmed more intensively and the number of bushels increases. But each additional worker is working with a smaller share of the 10 acres—the fixed input—than the previous worker. As a result, the additional worker cannot produce as much output as the previous worker. So it's not surprising that the marginal product of the additional worker falls.

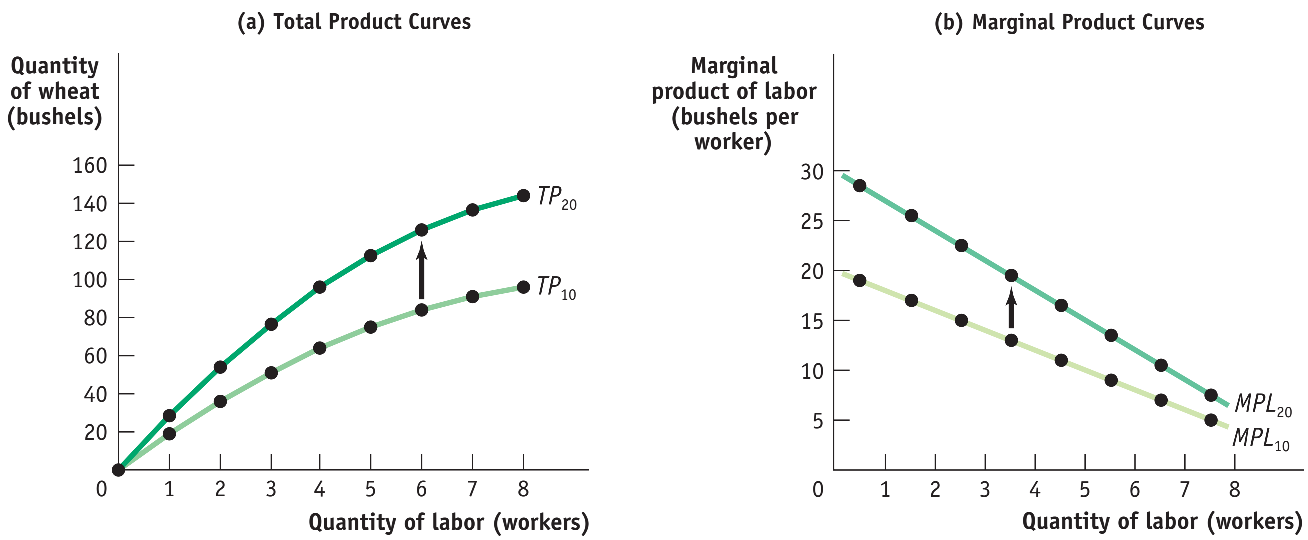
The crucial point to emphasize about diminishing returns is that, like many propositions in economics, it is an "other things equal" proposition: each successive unit of an input will raise production by less than the last *if the quantity of all other inputs is held fixed.*

What would happen if the levels of other inputs were allowed to change? You can see the answer illustrated in Figure 54.3. Panel (a) shows two total product curves, TP_{10} and TP_{20}. TP_{10} is the farm's total product curve when its total area is 10 acres (the same curve as in Figure 54.1). TP_{20} is the total product curve when the farm's area has increased to 20 acres. Except when 0 workers are employed, TP_{20} lies everywhere above TP_{10} because with more acres available, any given number of workers produces more output. Panel (b) shows the corresponding marginal product of labor curves.

Photodisc

With diminishing marginal returns to labor, as more and more workers are added to a fixed amount of land, each worker adds less to total output than the previous worker.

figure 54.3 **Total Product, Marginal Product, and the Fixed Input**



This figure shows how the quantity of output—illustrated by the total product curve—and marginal product depend on the level of the fixed input. Panel (a) shows two total product curves for George and Martha's farm, TP_{10} when their farm is 10 acres and TP_{20} when it is 20 acres. With more land, each worker can produce more wheat. So an increase in the fixed input shifts the total product curve up from TP_{10} to TP_{20}. This also implies that the marginal product of each worker is higher when the farm is 20 acres than when it is 10 acres. As a result, an increase in acreage also shifts the marginal product of labor curve up from MPL_{10} to MPL_{20}. Panel (b) shows the marginal product of labor curves. Note that both marginal product of labor curves still slope downward due to diminishing returns to labor.

fyi

Was Malthus Right?

In 1798, Thomas Malthus, an English pastor, authored the book *An Essay on the Principle of Population,* which introduced the principle of diminishing returns to an input. Malthus's writings were influential in his own time and continue to provoke heated argument to this day.

Malthus argued that as a country's population grew but its land area remained fixed, it would become increasingly difficult to grow enough food. Though more intensive cultivation of the land could increase yields, as the marginal product of labor declined, each successive farmer would add less to the total than the last.

From this argument, Malthus drew a powerful conclusion—that misery was the normal condition of humankind. In a country with a small population and abundant land (a description of the United States at the time), he argued, families would be large and the population would grow rapidly. Ultimately, the pressure of population on the land would reduce the condition of most people to a level at which starvation and disease held the population in check. (Arguments like this led the historian Thomas Carlyle to dub economics the "dismal science.")

Happily, over the long term, Malthus's predictions have turned out to be wrong. World population has increased from about 1 billion when Malthus wrote to more than 6.8 billion in 2010, but in most of the world people eat better now than ever before. So was Malthus completely wrong? And do his incorrect predictions refute the idea of diminishing returns? No, on both counts.

First, the Malthusian story is a pretty accurate description of 57 of the last 59 centuries: peasants in eighteenth-century France probably did not live much better than Egyptian peasants in the age of the pyramids. Yet diminishing returns does not mean that using more labor to grow food on a given amount of land will lead to a decline in the marginal product of labor—*if* there is also a radical improvement in farming technology. Fortunately, since the eighteenth century, technological progress has been so rapid that it has alleviated much of the limits imposed by diminishing returns. Diminishing returns implies that the marginal product declines when *all* other things—including technology—remain the same. So the happy fact that Malthus's predictions were wrong does not invalidate the concept of diminishing returns.

Typically, however, technological progress relaxes the limits imposed by diminishing returns only over the very long term. This was demonstrated in 2008 when bad weather, an ethanol-driven increase in the demand for corn, and a brisk rise in world income led to soaring world grain prices. As farmers scrambled to plant more acreage, they ran up against limits in the availability of inputs like land and fertilizer. Hopefully, we can prove Malthus wrong again before long.

MPL_{10} is the marginal product of labor curve given 10 acres to cultivate (the same curve as in Figure 54.2), and MPL_{20} is the marginal product of labor curve given 20 acres. Both curves slope downward because, in each case, the amount of land is fixed, albeit at different levels. But MPL_{20} lies everywhere above MPL_{10}, reflecting the fact that the marginal product of the same worker is higher when he or she has more of the fixed input to work with.

Figure 54.3 demonstrates a general result: the position of the total product curve depends on the quantities of other inputs. If you change the quantities of the other inputs, both the total product curve and the marginal product curve of the remaining input will shift. The importance of the "other things equal" assumption in discussing diminishing returns is illustrated in the FYI above.

Module 54 AP Review

Solutions appear at the back of the book.

Check Your Understanding

1. Bernie's ice-making company produces ice cubes using a 10-ton machine and electricity (along with water, which we will ignore as an input for simplicity). The quantity of output, measured in pounds of ice, is given in the accompanying table.
   a. What is the fixed input? What is the variable input?
   b. Construct a table showing the marginal product of the variable input. Does it show diminishing returns?
   c. Suppose a 50% increase in the size of the fixed input increases output by 100% for any given amount of the variable input. What is the fixed input now? Construct a table showing the quantity of output and the marginal product in this case.

Quantity of electricity (kilowatts)	*Quantity of ice (pounds)*
0	0
1	1,000
2	1,800
3	2,400
4	2,800

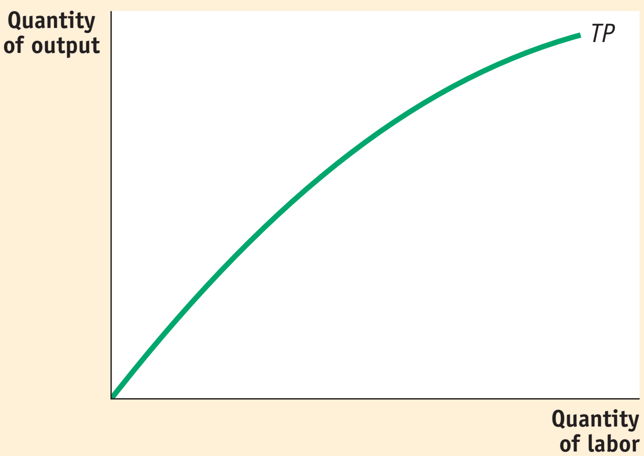
Tackle the Test: Multiple-Choice Questions

1. A production function shows the relationship between inputs and
   a. fixed costs.
   b. variable costs.
   c. total revenue.
   d. output.
   e. profit.
2. Which of the following defines the short run?
   a. less than a year
   b. when all inputs are fixed
   c. when no inputs are variable
   d. when only one input is variable
   e. when at least one input is fixed
3. The slope of the total product curve is also known as
   a. marginal product.
   b. marginal cost.
   c. average product.
   d. average revenue.
   e. profit.
4. Diminishing returns to an input ensures that as a firm continues to produce, the total product curve will have what kind of slope?
   a. negative decreasing
   b. positive decreasing
   c. negative increasing
   d. positive increasing
   e. positive constant
5. Historically, the limits imposed by diminishing returns have been alleviated by
   a. investment in capital.
   b. increases in the population.
   c. discovery of more land.
   d. Thomas Malthus.
   e. economic models.

Tackle the Test: Free-Response Questions

1. Draw a correctly labeled graph of a production function that exhibits diminishing returns to labor. Assume labor is the variable input and capital is the fixed input. Explain how your graph illustrates diminishing returns to labor.

Answer (4 points)



1 point: Graph with vertical axis labeled "Quantity of output" or "Q" and horizontal axis labeled "Quantity of labor" or "L"

1 point: Upward sloping curve labeled "Total product" or "TP"

1 point: The slope of the total product curve is positive and decreasing

1 point: Explanation that a positive and decreasing slope illustrates diminishing returns to labor because each additional unit of labor increases total product by less than the previous unit of labor

2. Use the data in the table below to graph the production function and the marginal product of labor. Do the data illustrate diminishing returns to labor? Explain.

Quantity of labor *L*	*Quantity of output* *Q*
0	0
1	19
2	36
3	51
4	64
5	75
6	84
7	91
8	96

What you will learn in this **Module:**

- The various types of cost a firm faces, including fixed cost, variable cost, and total cost
- How a firm's costs generate marginal cost curves and average cost curves

Module 55 Firm Costs

From the Production Function to Cost Curves

Now that we have learned about the firm's production function, we can use that knowledge to develop its cost curves. To see how a firm's production function is related to its cost curves, let's turn once again to George and Martha's farm. Once George and Martha know their production function, they know the relationship between inputs of labor and land and output of wheat. But if they want to maximize their profits, they need to translate this knowledge into information about the relationship between the quantity of output and cost. Let's see how they can do this.

To translate information about a firm's production function into information about its cost, we need to know how much the firm must pay for its inputs. We will assume that George and Martha face either an explicit or an implicit cost of $400 for the use of the land. As we learned previously, it is irrelevant whether George and Martha must rent the land for $400 from someone else or whether they own the land themselves and forgo earning $400 from renting it to someone else. Either way, they pay an opportunity cost of $400 by using the land to grow wheat. Moreover, since the land is a fixed input for which George and Martha pay $400 whether they grow one bushel of wheat or one hundred, its cost is a **fixed cost,** denoted by *FC*—a cost that does not depend on the quantity of output produced. In business, a fixed cost is often referred to as an "overhead cost."

We also assume that George and Martha must pay each worker $200. Using their production function, George and Martha know that the number of workers they must hire depends on the amount of wheat they intend to produce. So the cost of labor, which is equal to the number of workers multiplied by $200, is a **variable cost,** denoted by *VC*—a cost that depends on the quantity of output produced. Adding the fixed cost and the variable cost of a given quantity of output gives the **total cost,** or *TC*, of that quantity of output. We can express the relationship among fixed cost, variable cost, and total cost as an equation:

(55-1) Total cost = Fixed cost + Variable cost

or

$$TC = FC + VC$$

A **fixed cost** is a cost that does not depend on the quantity of output produced. It is the cost of the fixed input.

A **variable cost** is a cost that depends on the quantity of output produced. It is the cost of the variable input.

The **total cost** of producing a given quantity of output is the sum of the fixed cost and the variable cost of producing that quantity of output.

The table in Figure 55.1 shows how total cost is calculated for George and Martha's farm. The second column shows the number of workers employed, *L*. The third column shows the corresponding level of output, *Q*, taken from the table in Figure 54.1. The fourth column shows the variable cost, *VC*, equal to the number of workers multiplied by $200. The fifth column shows the fixed cost, *FC*, which is $400 regardless of the quantity of wheat produced. The sixth column shows the total cost of output, *TC*, which is the variable cost plus the fixed cost.

The first column labels each row of the table with a letter, from *A* to *I*. These labels will be helpful in understanding our next step: drawing the **total cost curve,** a curve that shows how total cost depends on the quantity of output.

George and Martha's total cost curve is shown in the diagram in Figure 55.1, where the horizontal axis measures the quantity of output in bushels of wheat and the vertical axis measures total cost in dollars. Each point on the curve corresponds to one row of the table in Figure 55.1. For example, point *A* shows the situation when 0 workers are employed: output is 0, and total cost is equal to fixed cost, $400. Similarly, point *B* shows the situation when 1 worker is employed: output is 19 bushels, and total cost is $600, equal to the sum of $400 in fixed cost and $200 in variable cost.

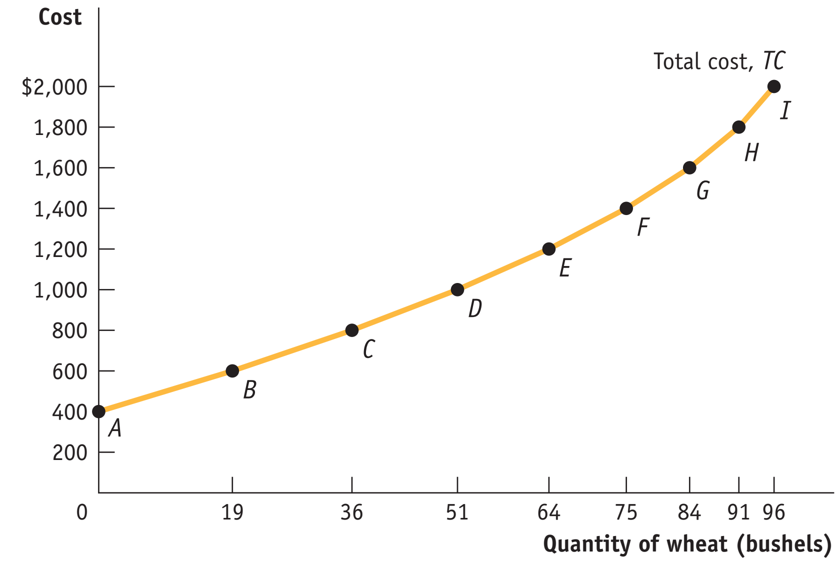
Like the total product curve, the total cost curve slopes upward: due to the increasing variable cost, the more output produced, the higher the farm's total cost.

The **total cost curve** shows how total cost depends on the quantity of output.

figure 55.1

Total Cost Curve for George and Martha's Farm

The table shows the variable cost, fixed cost, and total cost for various output quantities on George and Martha's 10-acre farm. The total cost curve shows how total cost (measured on the vertical axis) depends on the quantity of output (measured on the horizontal axis). The labeled points on the curve correspond to the rows of the table. The total cost curve slopes upward because the number of workers employed, and hence total cost, increases as the quantity of output increases. The curve gets steeper as output increases due to diminishing returns to labor.



Point on graph	Quantity of labor *L* (workers)	Quantity of wheat *Q* (bushels)	Variable cost *VC*	Fixed cost *FC*	Total cost *TC = FC + VC*
A	0	0	$0	$400	$400
B	1	19	200	400	600
C	2	36	400	400	800
D	3	51	600	400	1,000
E	4	64	800	400	1,200
F	5	75	1,000	400	1,400
G	6	84	1,200	400	1,600
H	7	91	1,400	400	1,800
I	8	96	1,600	400	2,000

But unlike the total product curve, which gets flatter as employment rises, the total cost curve gets *steeper*. That is, the slope of the total cost curve is greater as the amount of output produced increases. As we will soon see, the steepening of the total cost curve is also due to diminishing returns to the variable input. Before we can see why, we must first look at the relationships among several useful measures of cost.

Two Key Concepts: Marginal Cost and Average Cost

We've just learned how to derive a firm's total cost curve from its production function. Our next step is to take a deeper look at total cost by deriving two extremely useful measures: *marginal cost* and *average cost*. As we'll see, these two measures of the cost of production have a somewhat surprising relationship to each other. Moreover, they will prove to be vitally important in later modules, where we will use them to analyze the firm's output decision and the market supply curve.

Marginal Cost

Module 53 explained that marginal cost is the added cost of doing something one more time. In the context of production, marginal cost is the change in total cost generated by producing one more unit of output. We've already seen that marginal product is easiest to calculate if data on output are available in increments of one unit of input. Similarly, marginal cost is easiest to calculate if data on total cost are available in increments of one unit of output because the increase in total cost for each unit is clear. When the data come in less convenient increments, it's still possible to calculate marginal cost over each interval. But for the sake of simplicity, let's work with an example in which the data come in convenient one-unit increments.

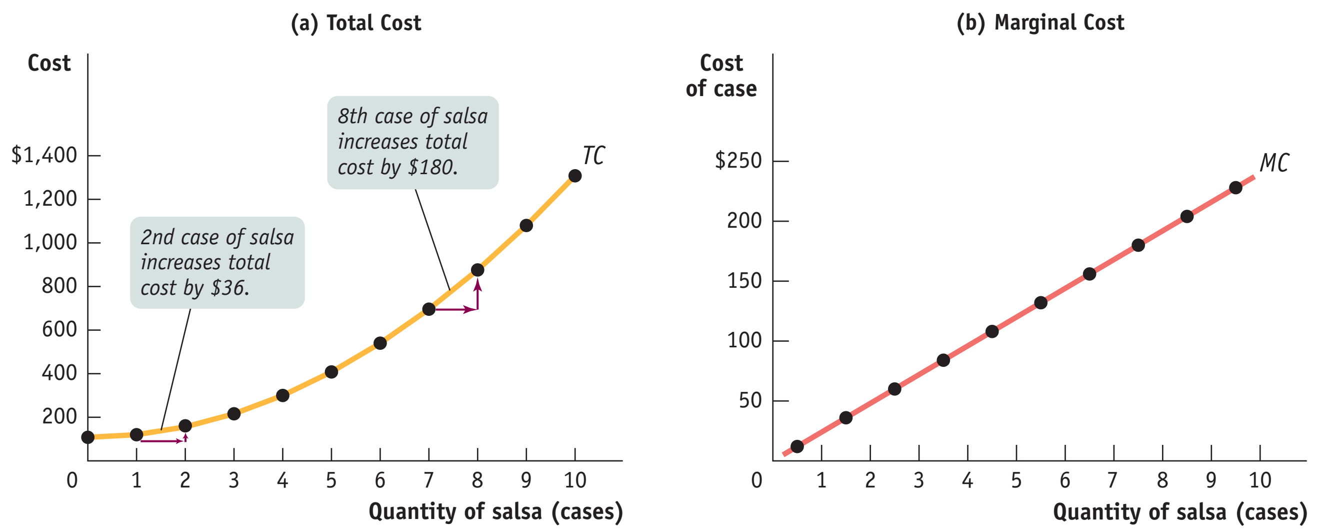
Selena's Gourmet Salsas produces bottled salsa; Table 55.1 shows how its costs per day depend on the number of cases of salsa it produces per day. The firm has a fixed

table 55.1

Costs at Selena's Gourmet Salsas

Quantity of salsa Q (cases)	Fixed cost FC	Variable cost VC	Total cost $TC = FC + VC$	Marginal cost of case $MC = \Delta TC/\Delta Q$
0	$108	$0	$108	
				$12
1	108	12	120	
				36
2	108	48	156	
				60
3	108	108	216	
				84
4	108	192	300	
				108
5	108	300	408	
				132
6	108	432	540	
				156
7	108	588	696	
				180
8	108	768	876	
				204
9	108	972	1,080	
				228
10	108	1,200	1,308	

figure 55.2 Total Cost and Marginal Cost Curves for Selena's Gourmet Salsas



Panel (a) shows the total cost curve from Table 55.1. Like the total cost curve in Figure 55.1, it slopes upward and gets steeper as we move up it to the right. Panel (b) shows the marginal cost curve. It also slopes upward, reflecting diminishing returns to the variable input.

cost of $108 per day, shown in the second column, which is the daily rental cost of its food-preparation equipment. The third column shows the variable cost, and the fourth column shows the total cost. Panel (a) of Figure 55.2 plots the total cost curve. Like the total cost curve for George and Martha's farm in Figure 55.1, this curve slopes upward, getting steeper as quantity increases.

The significance of the slope of the total cost curve is shown by the fifth column of Table 55.1, which indicates marginal cost—the additional cost of each additional unit. The general formula for marginal cost is:

$$\textbf{(55-2)}\quad \text{Marginal cost} = \begin{array}{l}\text{Change in total}\\ \text{cost generated by}\\ \text{one additional unit}\\ \text{of output}\end{array} = \frac{\text{Change in total cost}}{\text{Change in quantity of output}}$$

or

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

As in the case of marginal product, marginal cost is equal to "rise" (the increase in total cost) divided by "run" (the increase in the quantity of output). So just as marginal product is equal to the slope of the total product curve, marginal cost is equal to the slope of the total cost curve.

iStockphoto

Now we can understand why the total cost curve gets steeper as it increases from left to right: as you can see in Table 55.1, marginal cost at Selena's Gourmet Salsas rises as output increases. And because marginal cost equals the slope of the total cost curve, a higher marginal cost means a steeper slope. Panel (b) of Figure 55.2 shows the marginal cost curve corresponding to the data in Table 55.1. Notice that, as in Figure 53.1, we plot the marginal cost for increasing output from 0 to 1 case of salsa halfway between 0 and 1, the marginal cost for increasing output from 1 to 2 cases of salsa halfway between 1 and 2, and so on.

Average total cost, often referred to simply as **average cost,** is total cost divided by quantity of output produced.

Why does the marginal cost curve slope upward? Because there are diminishing returns to inputs in this example. As output increases, the marginal product of the variable input declines. This implies that more and more of the variable input must be used to produce each additional unit of output as the amount of output already produced rises. And since each unit of the variable input must be paid for, the additional cost per additional unit of output also rises.

Recall that the flattening of the total product curve is also due to diminishing returns: if the quantities of other inputs are fixed, the marginal product of an input falls as more of that input is used. The flattening of the total product curve as output increases and the steepening of the total cost curve as output increases are just flip-sides of the same phenomenon. That is, as output increases, the marginal cost of output also increases because the marginal product of the variable input decreases. Our next step is to introduce another measure of cost: *average cost.*

Average Cost

In addition to total cost and marginal cost, it's useful to calculate **average total cost,** often simply called **average cost.** The average total cost is total cost divided by the quantity of output produced; that is, it is equal to total cost per unit of output. If we let *ATC* denote average total cost, the equation looks like this:

$$\textbf{(55-3)} \quad ATC = \frac{\text{Total cost}}{\text{Quantity of output}} = \frac{TC}{Q}$$

Average total cost is important because it tells the producer how much the *average* or *typical* unit of output costs to produce. Marginal cost, meanwhile, tells the producer how much *one more* unit of output costs to produce. Although they may look very similar, these two measures of cost typically differ. And confusion between them is a major source of error in economics, both in the classroom and in real life. Table 55.2 uses data from Selena's Gourmet Salsas to calculate average total cost. For example, the total cost of producing 4 cases of salsa is $300, consisting of $108 in fixed cost and $192 in variable cost (from Table 55.1). So the average total cost of producing 4 cases of salsa is

table **55.2**

Average Costs for Selena's Gourmet Salsas

Quantity of salsa *Q* (cases)	Total cost *TC*	Average total cost of case *ATC = TC/Q*	Average fixed cost of case *AFC = FC/Q*	Average variable cost of case *AVC = VC/Q*
1	$120	$120.00	$108.00	$12.00
2	156	78.00	54.00	24.00
3	216	72.00	36.00	36.00
4	300	75.00	27.00	48.00
5	408	81.60	21.60	60.00
6	540	90.00	18.00	72.00
7	696	99.43	15.43	84.00
8	876	109.50	13.50	96.00
9	1,080	120.00	12.00	108.00
10	1,308	130.80	10.80	120.00

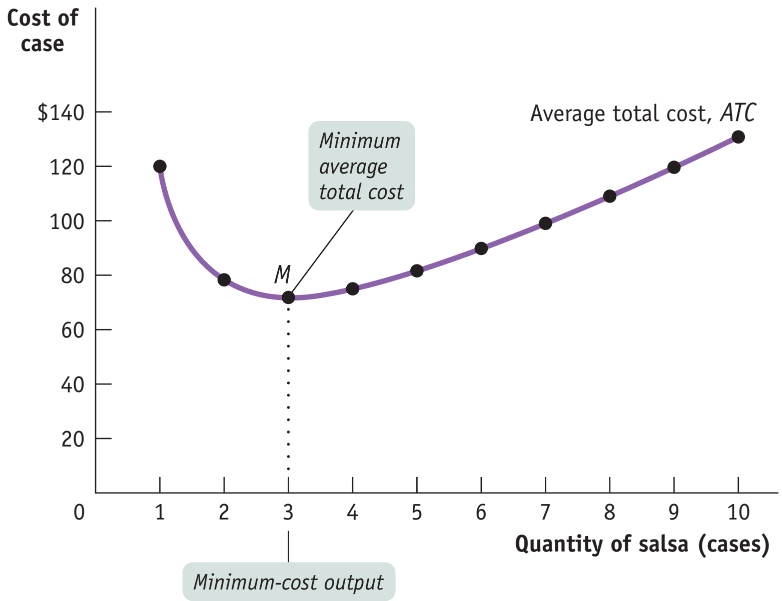
$300/4 = $75. You can see from Table 55.2 that as the quantity of output increases, average total cost first falls, then rises.

Figure 55.3 plots that data to yield the *average total cost curve,* which shows how average total cost depends on output. As before, cost in dollars is measured on the vertical axis and quantity of output is measured on the horizontal axis. The average total cost curve has a distinctive U shape that corresponds to how average total cost first falls and then rises as output increases. Economists believe that such **U-shaped average total cost curves** are the norm for firms in many industries.

figure 55.3

Average Total Cost Curve for Selena's Gourmet Salsas

The average total cost curve at Selena's Gourmet Salsas is U-shaped. At low levels of output, average total cost falls because the "spreading effect" of falling average fixed cost dominates the "diminishing returns effect" of rising average variable cost. At higher levels of output, the opposite is true and average total cost rises. At point *M,* corresponding to an output of three cases of salsa per day, average total cost is at its minimum level, the minimum average total cost.



To help our understanding of why the average total cost curve is U-shaped, Table 55.2 breaks average total cost into its two underlying components, *average fixed cost* and *average variable cost.* **Average fixed cost,** or *AFC,* is fixed cost divided by the quantity of output, also known as the fixed cost per unit of output. For example, if Selena's Gourmet Salsas produces 4 cases of salsa, average fixed cost is $108/4 = $27 per case. **Average variable cost,** or *AVC,* is variable cost divided by the quantity of output, also known as variable cost per unit of output. At an output of 4 cases, average variable cost is $192/4 = $48 per case. Writing these in the form of equations:

$$\textbf{(55-4)}\quad AFC = \frac{\text{Fixed cost}}{\text{Quantity of output}} = \frac{FC}{Q}$$

$$AVC = \frac{\text{Variable cost}}{\text{Quantity of output}} = \frac{VC}{Q}$$

Average total cost is the sum of average fixed cost and average variable cost; it has a U shape because these components move in opposite directions as output rises.

Average fixed cost falls as more output is produced because the numerator (the fixed cost) is a fixed number but the denominator (the quantity of output) increases as more is produced. Another way to think about this relationship is that, as more output is produced, the fixed cost is spread over more units of output; the end result is that the

A **U-shaped average total cost curve** falls at low levels of output and then rises at higher levels.

Average fixed cost is the fixed cost per unit of output.

Average variable cost is the variable cost per unit of output.


Photodisc

fixed cost *per unit of output*—the average fixed cost—falls. You can see this effect in the fourth column of Table 55.2: average fixed cost drops continuously as output increases. Average variable cost, however, rises as output increases. As we've seen, this reflects diminishing returns to the variable input: each additional unit of output adds more to variable cost than the previous unit because increasing amounts of the variable input are required to make another unit.

So increasing output has two opposing effects on average total cost—the "spreading effect" and the "diminishing returns effect":

- *The spreading effect.* The larger the output, the greater the quantity of output over which fixed cost is spread, leading to lower average fixed cost.
- *The diminishing returns effect.* The larger the output, the greater the amount of variable input required to produce additional units, leading to higher average variable cost.

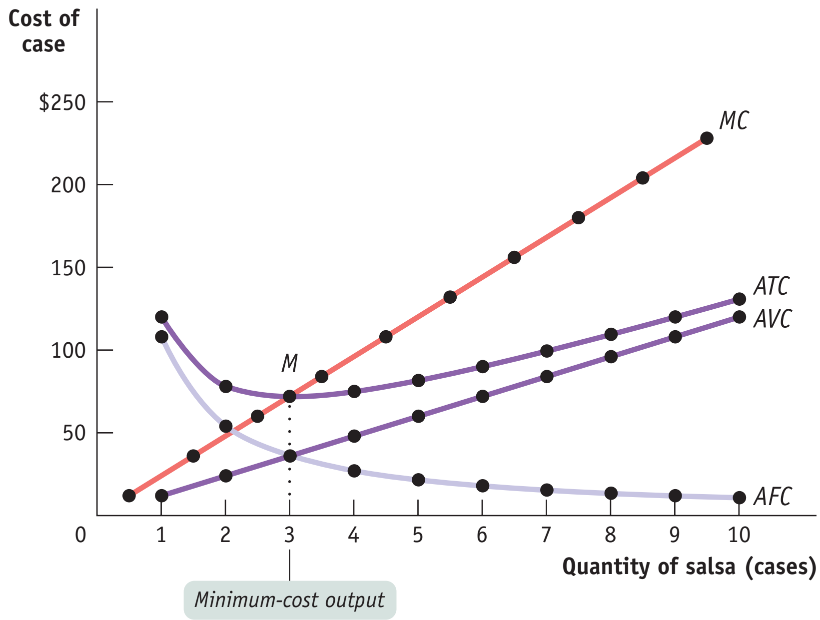
At low levels of output, the spreading effect is very powerful because even small increases in output cause large reductions in average fixed cost. So at low levels of output, the spreading effect dominates the diminishing returns effect and causes the average total cost curve to slope downward. But when output is large, average fixed cost is already quite small, so increasing output further has only a very small spreading effect. Diminishing returns, however, usually grow increasingly important as output rises. As a result, when output is large, the diminishing returns effect dominates the spreading effect, causing the average total cost curve to slope upward. At the bottom of the U-shaped average total cost curve, point *M* in Figure 55.3, the two effects exactly balance each other. At this point average total cost is at its minimum level, the minimum average total cost.

Figure 55.4 brings together in a single picture the four other cost curves that we have derived from the total cost curve for Selena's Gourmet Salsas: the marginal cost curve (*MC*), the average total cost curve (*ATC*), the average variable cost curve (*AVC*), and the average fixed cost curve (*AFC*). All are based on the information in Tables 55.1 and 55.2. As before, cost is measured on the vertical axis and the quantity of output is measured on the horizontal axis.

figure 55.4

Marginal Cost and Average Cost Curves for Selena's Gourmet Salsas

Here we have the family of cost curves for Selena's Gourmet Salsas: the marginal cost curve (*MC*), the average total cost curve (*ATC*), the average variable cost curve (*AVC*), and the average fixed cost curve (*AFC*). Note that the average total cost curve is U-shaped and the marginal cost curve crosses the average total cost curve at the bottom of the U, point *M*, corresponding to the minimum average total cost from Table 55.2 and Figure 55.3.



Let's take a moment to note some features of the various cost curves. First of all, marginal cost slopes upward—the result of diminishing returns that make an additional unit of output more costly to produce than the one before. Average variable cost also slopes upward—again, due to diminishing returns—but is flatter than the marginal cost curve. This is because the higher cost of an additional unit of output is averaged across all units, not just the additional unit, in the average variable cost measure. Meanwhile, average fixed cost slopes downward because of the spreading effect.

Finally, notice that the marginal cost curve intersects the average total cost curve from below, crossing it at its lowest point, point *M* in Figure 55.4. This last feature is our next subject of study.

The **minimum-cost output** is the quantity of output at which average total cost is lowest—it corresponds to the bottom of the U-shaped average total cost curve.

Minimum Average Total Cost

For a U-shaped average total cost curve, average total cost is at its minimum level at the bottom of the U. Economists call the quantity of output that corresponds to the minimum average total cost the **minimum-cost output.** In the case of Selena's Gourmet Salsas, the minimum-cost output is three cases of salsa per day.

In Figure 55.4, the bottom of the U is at the level of output at which the marginal cost curve crosses the average total cost curve from below. Is this an accident? No—it reflects general principles that are always true about a firm's marginal cost and average total cost curves:

- At the minimum-cost output, average total cost *is equal to* marginal cost.
- At output less than the minimum-cost output, marginal cost *is less than* average total cost and average total cost is falling.
- And at output greater than the minimum-cost output, marginal cost *is greater than* average total cost and average total cost is rising.

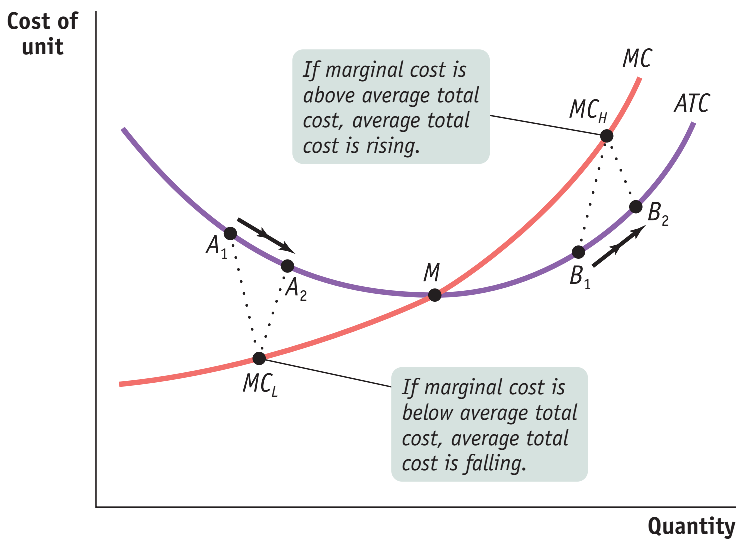
To understand these principles, think about how your grade in one course—say, a 3.0 in physics—affects your overall grade point average. If your GPA before receiving that grade was more than 3.0, the new grade lowers your average.

Similarly, if marginal cost—the cost of producing one more unit—is less than average total cost, producing that extra unit lowers average total cost. This is shown in Figure 55.5 by the movement from A_1 to A_2. In this case, the marginal cost of producing

figure 55.5

The Relationship Between the Average Total Cost and the Marginal Cost Curves

To see why the marginal cost curve (*MC*) must cut through the average total cost curve at the minimum average total cost (point *M*), corresponding to the minimum-cost output, we look at what happens if marginal cost is different from average total cost. If marginal cost is *less* than average total cost, an increase in output must reduce average total cost, as in the movement from A_1 to A_2. If marginal cost is *greater* than average total cost, an increase in output must increase average total cost, as in the movement from B_1 to B_2.



an additional unit of output is low, as indicated by the point MC_L on the marginal cost curve. When the cost of producing the next unit of output is less than average total cost, increasing production reduces average total cost. So any quantity of output at which marginal cost is less than average total cost must be on the downward-sloping segment of the U.

But if your grade in physics is more than the average of your previous grades, this new grade raises your GPA. Similarly, if marginal cost is greater than average total cost, producing that extra unit raises average total cost. This is illustrated by the movement from B_1 to B_2 in Figure 55.5, where the marginal cost, MC_H, is higher than average total cost. So any quantity of output at which marginal cost is greater than average total cost must be on the upward-sloping segment of the U.

Finally, if a new grade is exactly equal to your previous GPA, the additional grade neither raises nor lowers that average—it stays the same. This corresponds to point *M* in Figure 55.5: when marginal cost equals average total cost, we must be at the bottom of the U because only at that point is average total cost neither falling nor rising.

Does the Marginal Cost Curve Always Slope Upward?

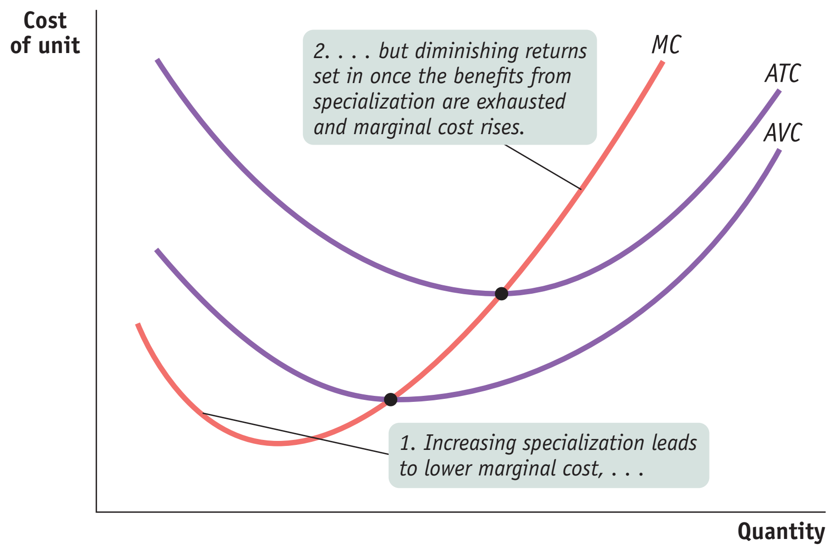
Up to this point, we have emphasized the importance of diminishing returns, which lead to a marginal product curve that always slopes downward and a marginal cost curve that always slopes upward. In practice, however, economists believe that marginal cost curves often slope *downward* as a firm increases its production from zero up to some low level, sloping upward only at higher levels of production: marginal cost curves look like the curve labeled *MC* in Figure 55.6.

This initial downward slope occurs because a firm often finds that, when it starts with only a very small number of workers, employing more workers and expanding output allows its workers to specialize in various tasks. This, in turn, lowers the firm's marginal cost as it expands output. For example, one individual producing salsa would have to perform all the tasks involved: selecting and preparing the ingredients, mixing the salsa, bottling and labeling it, packing it into cases, and so on. As more workers are employed, they can divide the tasks, with each worker specializing in one or a few aspects of salsa-making. This specialization leads to *increasing returns* to the hiring of additional workers and results in a marginal cost curve that initially slopes downward.

figure 55.6

More Realistic Cost Curves

A realistic marginal cost curve has a "swoosh" shape. Starting from a very low output level, marginal cost often falls as the firm increases output. That's because hiring additional workers allows greater specialization of their tasks and leads to increasing returns. Once specialization is achieved, however, diminishing returns to additional workers set in and marginal cost rises. The corresponding average variable cost curve is now U-shaped, like the average total cost curve.



But once there are enough workers to have completely exhausted the benefits of further specialization, diminishing returns to labor set in and the marginal cost curve changes direction and slopes upward. So typical marginal cost curves actually have the "swoosh" shape shown by *MC* in Figure 55.6. For the same reason, average variable cost curves typically look like *AVC* in Figure 55.6: they are U-shaped rather than strictly upward sloping.

However, as Figure 55.6 also shows, the key features we saw from the example of Selena's Gourmet Salsas remain true: the average total cost curve is U-shaped, and the marginal cost curve passes through the point of minimum average total cost.

Module 55 AP Review

Solutions appear at the back of the book.

Check Your Understanding

1. Alicia's Apple Pies is a roadside business. Alicia must pay $9.00 in rent each day. In addition, it costs her $1.00 to produce the first pie of the day, and each subsequent pie costs 50% more to produce than the one before. For example, the second pie costs $\$1.00 \times 1.5 = \$1.50$ to produce, and so on.
   a. Calculate Alicia's marginal cost, variable cost, average fixed cost, average variable cost, and average total cost as her daily pie output rises from 0 to 6. (*Hint:* The variable cost of two pies is just the marginal cost of the first pie, plus the marginal cost of the second, and so on.)
   b. Indicate the range of pies for which the spreading effect dominates and the range for which the diminishing returns effect dominates.
   c. What is Alicia's minimum-cost output? Explain why making one more pie lowers Alicia's average total cost when output is lower than the minimum-cost output. Similarly, explain why making one more pie raises Alicia's average total cost when output is greater than the minimum-cost output.

Tackle the Test: Multiple-Choice Questions

1. When a firm is producing zero output, total cost equals
   a. zero.
   b. variable cost.
   c. fixed cost.
   d. average total cost.
   e. marginal cost.
2. Which of the following statements is true?
   I. Marginal cost is the change in total cost generated by one additional unit of output.
   II. Marginal cost is the change in variable cost generated by one additional unit of output.
   III. The marginal cost curve must cross the minimum of the average total cost curve.
   a. I only
   b. II only
   c. III only
   d. I and II only
   e. I, II, and III
3. Which of the following is correct?
   a. AVC is the change in total cost generated by one additional unit of output.
   b. MC = TC/Q
   c. The average cost curve crosses at the minimum of the marginal cost curve.
   d. The AFC curve slopes upward.
   e. AVC = ATC – AFC
4. The slope of the total cost curve equals
   a. variable cost.
   b. average variable cost.
   c. average total cost.
   d. average fixed cost.
   e. marginal cost.
5.

Q	*VC*	*TC*
0	$0	$40
1	20	60
2	50	90
3	90	130
4	140	180
5	200	240

   On the basis of the data in the table above, what is the marginal cost of the third unit of output?
   a. 40
   b. 50
   c. 60
   d. 90
   e. 130

Tackle the Test: Free-Response Questions

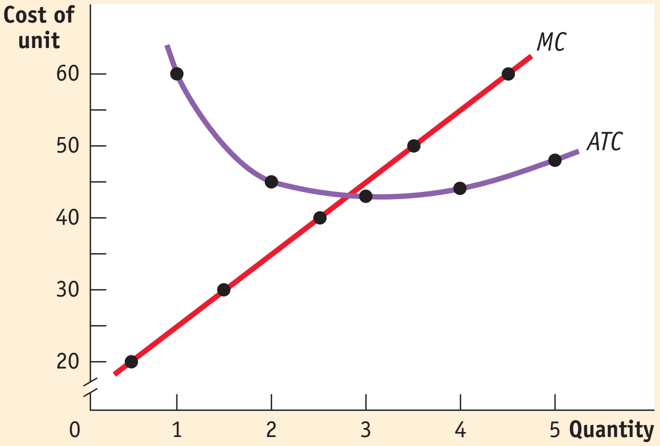
1. Use the information in the table below to answer the following questions.

Q	*VC*	*TC*
0	$0	$40
1	20	60
2	50	90
3	90	130
4	140	180
5	200	240

   a. What is the firm's level of fixed cost? Explain how you know.

   b. Draw one correctly labeled graph showing the firm's marginal and average total cost curves.

Answer (6 points)



1 point: *FC* = $40

1 point: We can identify the fixed cost as $40 because when the firm is not producing, it still incurs a cost of $40. This could only be the result of a fixed cost because variable cost is zero when output is zero.

1 point: Graph with correct labels ("Cost of unit" on vertical axis; "Quantity" on horizontal axis)

1 point: Upward sloping *MC* curve plotted according to data, labeled *"MC"*

1 point: U-shaped *ATC* curve plotted according to the provided data, labeled *"ATC"*

1 point: *MC* curve crossing at minimum of *ATC* curve (Note: We have simplified this graph by drawing smooth lines between discrete points. If we had drawn the *MC* curve as a step function instead, the *MC* curve would have crossed the *ATC* curve exactly at its minimum point.)

2. Draw a correctly labeled graph showing a firm with an upward sloping *MC* curve and typically shaped *ATC, AVC,* and *AFC* curves.

Module 56 Long-Run Costs and Economies of Scale

What you will learn in this Module:

- Why a firm's costs may differ between the short run and the long run
- How a firm can enjoy economies of scale

Up to this point, we have treated fixed cost as completely outside the control of a firm because we have focused on the short run. But all inputs are variable in the long run: this means that in the long run, even "fixed cost" may change. *In the long run, in other words, a firm's fixed cost becomes a variable it can choose.* For example, given time, Selena's Gourmet Salsas can acquire additional food-preparation equipment or dispose of some of its existing equipment. In this module, we will examine how a firm's costs behave in the short run and in the long run. We will also see that the firm will choose its fixed cost in the long run based on the level of output it expects to produce.

Short-Run versus Long-Run Costs

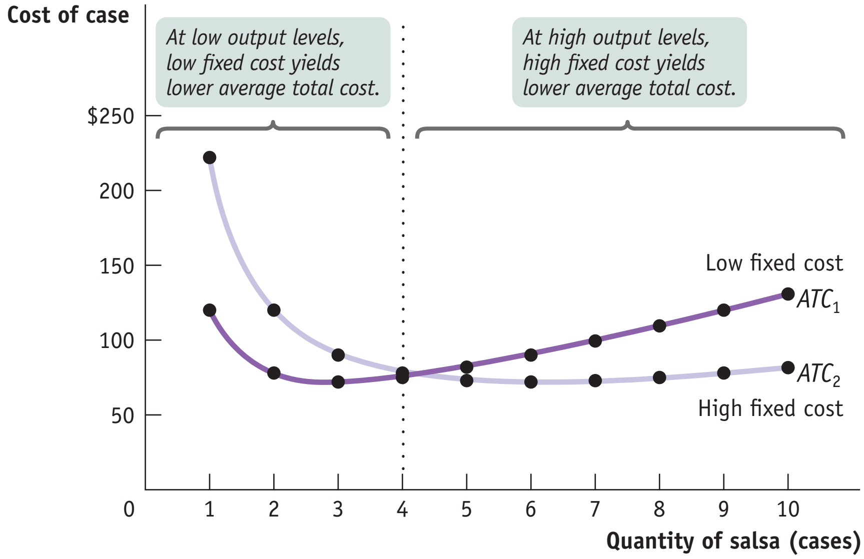
Let's begin by supposing that Selena's Gourmet Salsas is considering whether to acquire additional food-preparation equipment. Acquiring additional machinery will affect its total cost in two ways. First, the firm will have to either rent or buy the additional equipment; either way, that will mean a higher fixed cost in the short run. Second, if the workers have more equipment, they will be more productive: fewer workers will be needed to produce any given output, so variable cost for any given output level will be reduced.

The table in Figure 56.1 on the next page shows how acquiring an additional machine affects costs. In our original example, we assumed that Selena's Gourmet Salsas had a fixed cost of $108. The left half of the table shows variable cost as well as total cost and average total cost assuming a fixed cost of $108. The average total cost curve for this level of fixed cost is given by ATC_1 in Figure 56.1. Let's compare that to a situation in which the firm buys additional food-preparation equipment, doubling its fixed cost to $216 but reducing its variable cost at any given level of output. The right half of the table shows the firm's variable cost, total cost, and average total cost with this higher level of fixed cost. The average total cost curve corresponding to $216 in fixed cost is given by ATC_2 in Figure 56.1.

figure 56.1

Choosing the Level of Fixed Cost for Selena's Gourmet Salsas

There is a trade-off between higher fixed cost and lower variable cost for any given output level, and vice versa. ATC_1 is the average total cost curve corresponding to a fixed cost of $108; it leads to lower fixed cost and higher variable cost. ATC_2 is the average total cost curve corresponding to a higher fixed cost of $216 but lower variable cost. At low output levels, at 4 or fewer cases of salsa per day, ATC_1 lies below ATC_2: average total cost is lower with only $108 in fixed cost. But as output goes up, average total cost is lower with the higher amount of fixed cost, $216: at more than 4 cases of salsa per day, ATC_2 lies below ATC_1.



	Low fixed cost (FC = $108)			High fixed cost (FC = $216)		
Quantity of salsa (cases)	High variable cost	Total cost	Average total cost of case ATC_1	Low variable cost	Total cost	Average total cost of case ATC_2
1	$12	$120	$120.00	$6	$222	$222.00
2	48	156	78.00	24	240	120.00
3	108	216	72.00	54	270	90.00
4	192	300	75.00	96	312	78.00
5	300	408	81.60	150	366	73.20
6	432	540	90.00	216	432	72.00
7	588	696	99.43	294	510	72.86
8	768	876	109.50	384	600	75.00
9	972	1,080	120.00	486	702	78.00
10	1,200	1,308	130.80	600	816	81.60

From the figure you can see that when output is small, 4 cases of salsa per day or fewer, average total cost is smaller when Selena forgoes the additional equipment and maintains the lower fixed cost of $108: ATC_1 lies below ATC_2. For example, at 3 cases per day, average total cost is $72 without the additional machinery and $90 with the additional machinery. But as output increases beyond 4 cases per day, the firm's average total cost is lower if it acquires the additional equipment, raising its fixed cost to $216. For example, at 9 cases of salsa per day, average total cost is $120 when fixed cost is $108 but only $78 when fixed cost is $216.

Why does average total cost change like this when fixed cost increases? When output is low, the increase in fixed cost from the additional equipment outweighs the reduction in variable cost from higher worker productivity—that is, there are too few units of output over which to spread the additional fixed cost. So if Selena plans to produce 4 or fewer cases per day, she would be better off choosing the lower level of fixed cost, $108, to achieve a lower average total cost of production. When planned output is high, however, she should acquire the additional machinery.

In general, for each output level there is some choice of fixed cost that minimizes the firm's average total cost for that output level. So when the firm has a desired output level that it expects to maintain over time, it should choose the optimal fixed cost for that level—that is, the level of fixed cost that minimizes its average total cost.

Now that we are studying a situation in which fixed cost can change, we need to take *time* into account when discussing average total cost. All of the average total cost curves we have considered until now are defined for a given level of fixed cost—that is, they are defined for the short run, the period of time over which fixed cost doesn't vary. To reinforce that distinction, for the rest of this module we will refer to these average total cost curves as "short-run average total cost curves."

For most firms, it is realistic to assume that there are many possible choices of fixed cost, not just two. The implication: for such a firm, many possible short-run average total cost curves will exist, each corresponding to a different choice of fixed cost and so giving rise to what is called a firm's "family" of short-run average total cost curves.

At any given time, a firm will find itself on one of its short-run cost curves, the one corresponding to its current level of fixed cost; a change in output will cause it to move along that curve. If the firm expects that change in output level to be long-standing, then it is likely that the firm's current level of fixed cost is no longer optimal. Given sufficient time, it will want to adjust its fixed cost to a new level that minimizes average total cost for its new output level. For example, if Selena had been producing 2 cases of salsa per day with a fixed cost of $108 but found herself increasing her output to 8 cases per day for the foreseeable future, then in the long run she should purchase more equipment and increase her fixed cost to a level that minimizes average total cost at the 8-cases-per-day output level.

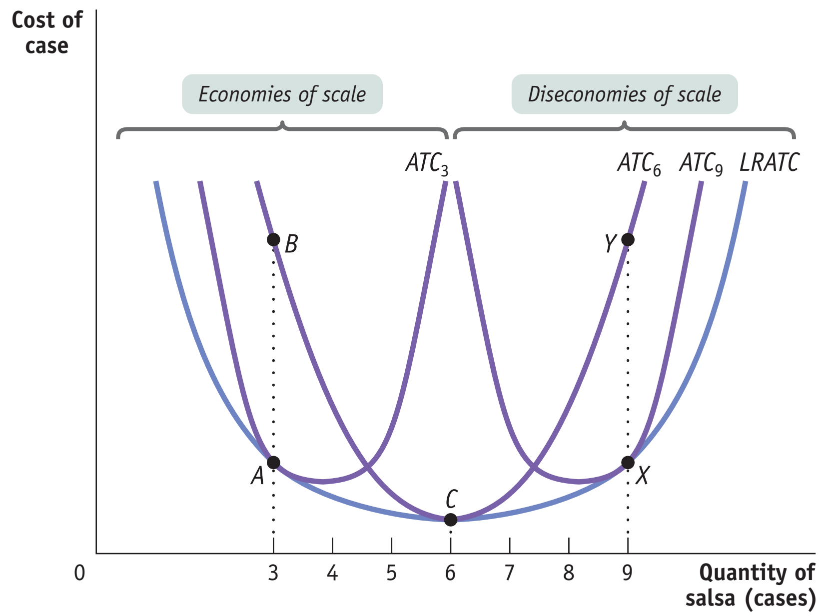
Suppose we do a thought experiment and calculate the lowest possible average total cost that can be achieved for each output level if the firm were to choose its fixed cost for each output level. Economists have given this thought experiment a name: the *long-run average total cost curve.* Specifically, the **long-run average total cost curve,** or *LRATC,* is the relationship between output and average total cost when fixed cost has been chosen to minimize average total cost *for each level of output.* If there are many possible choices of fixed cost, the long-run average total cost curve will have the familiar, smooth U shape, as shown by *LRATC* in Figure 56.2.

The **long-run average total cost curve** shows the relationship between output and average total cost when fixed cost has been chosen to minimize average total cost for each level of output.

figure 56.2

Short-Run and Long-Run Average Total Cost Curves

Short-run and long-run average total cost curves differ because a firm can choose its fixed cost in the long run. If Selena has chosen the level of fixed cost that minimizes short-run average total cost at an output of 6 cases, and actually produces 6 cases, then she will be at point *C* on *LRATC* and ATC_6. But if she produces only 3 cases, she will move to point *B*. If she expects to produce only 3 cases for a long time, in the long run she will reduce her fixed cost and move to point *A* on ATC_3. Likewise, if she produces 9 cases (putting her at point *Y*) and expects to continue this for a long time, she will increase her fixed cost in the long run and move to point *X*.



We can now draw the distinction between the short run and the long run more fully. In the long run, when a producer has had time to choose the fixed cost appropriate for its desired level of output, that producer will be at some point on the long-run average total cost curve. But if the output level is altered, the firm will no longer be on its long-run average total cost curve and will instead be moving along its current short-run average total cost curve. It will not be on its long-run average total cost curve again until it readjusts its fixed cost for its new output level.

Figure 56.2 illustrates this point. The curve ATC_3 shows short-run average total cost if Selena has chosen the level of fixed cost that minimizes average total cost at an output of 3 cases of salsa per day. This is confirmed by the fact that at 3 cases per day, ATC_3 touches *LRATC,* the long-run average total cost curve. Similarly, ATC_6 shows short-run average total cost if Selena has chosen the level of fixed cost that minimizes average total cost if her output is 6 cases per day. It touches *LRATC* at 6 cases per day. And ATC_9 shows short-run average total cost if Selena has chosen the level of fixed cost that minimizes average total cost if her output is 9 cases per day. It touches *LRATC* at 9 cases per day.

Suppose that Selena initially chose to be on ATC_6. If she actually produces 6 cases of salsa per day, her firm will be at point *C* on both its short-run and long-run average total cost curves. Suppose, however, that Selena ends up producing only 3 cases of salsa per day. In the short run, her average total cost is indicated by point *B* on ATC_6; it is no longer on *LRATC.* If Selena had known that she would be producing only 3 cases per day, she would have been better off choosing a lower level of fixed cost, the one corresponding to ATC_3, thereby achieving a lower average total cost. Then her firm would have found itself at point *A* on the long-run average total cost curve, which lies below point *B.*


Photodisc

Suppose, conversely, that Selena ends up producing 9 cases per day even though she initially chose to be on ATC_6. In the short run her average total cost is indicated by point *Y* on ATC_6. But she would be better off purchasing more equipment and incurring a higher fixed cost in order to reduce her variable cost and move to ATC_9. This would allow her to reach point *X* on the long-run average total cost curve, which lies below *Y.* The distinction between short-run and long-run average total costs is extremely important in making sense of how real firms operate over time. A company that has to increase output suddenly to meet a surge in demand will typically find that in the short run its average total cost rises sharply because it is hard to get extra production out of existing facilities. But given time to build new factories or add machinery, short-run average total cost falls.

Returns to Scale

What determines the shape of the long-run average total cost curve? It is the influence of *scale,* the size of a firm's operations, on its long-run average total cost of production. Firms that experience *scale effects* in production find that their long-run average total cost changes substantially depending on the quantity of output they produce. There are **economies of scale** when long-run average total cost declines as output increases. As you can see in Figure 56.2, Selena's Gourmet Salsas experiences economies of scale over output levels ranging from 0 up to 6 cases of salsa per day—the output levels over which the long-run average total cost curve is declining. Economies of scale can result from **increasing returns to scale,** which exist when output increases more than in proportion to an increase in all inputs. For example, if Selena could double all of her inputs and make more than twice as much salsa, she would be experiencing increasing returns to scale. With twice the inputs (and costs) and more than twice the salsa, she would be enjoying decreasing long-run average total costs, and thus economies of scale. Increasing returns to scale therefore imply economies of scale, although economies of scale exist whenever long-run average total cost is falling, whether or not all inputs are increasing by the same proportion.

There are **economies of scale** when long-run average total cost declines as output increases.

There are **increasing returns to scale** when output increases more than in proportion to an increase in all inputs. For example, with increasing returns to scale, doubling all inputs would cause output to more than double.

There are **diseconomies of scale** when long-run average total cost increases as output increases.

In contrast, there are **diseconomies of scale** when long-run average total cost increases as output increases. For Selena's Gourmet Salsas, decreasing returns to scale occur at output levels greater than 6 cases, the output levels over which its long-run average total cost curve is rising. Diseconomies of scale can result from **decreasing returns to scale,** which exist when output increases less than in proportion to an increase in all inputs—doubling the inputs results in less than double the output. When output increases directly in proportion to an increase in all inputs—doubling the inputs results in double the output—the firm is experiencing **constant returns to scale.**

What explains these scale effects in production? The answer ultimately lies in the firm's technology of production. Economies of scale often arise from the increased *specialization* that larger output levels allow—a larger scale of operation means that individual workers can limit themselves to more specialized tasks, becoming more skilled and efficient at doing them. Another source of economies of scale is a very large initial setup cost; in some industries—such as auto manufacturing, electricity generating, and petroleum refining—it is necessary to pay a high fixed cost in the form of plant and equipment before producing any output. A third source of economies of scale, found in certain high-tech industries such as software development, is *network externalities,* a topic covered in a later module. As we'll see when we study monopoly, increasing returns have very important implications for how firms and industries behave and interact.

Diseconomies of scale—the opposite scenario—typically arise in large firms due to problems of coordination and communication: as a firm grows in size, it becomes ever more difficult and therefore costly to communicate and to organize activities. Although economies of scale induce firms to grow larger, diseconomies of scale tend to limit their size.

There are **decreasing returns to scale** when output increases less than in proportion to an increase in all inputs.

There are **constant returns to scale** when output increases directly in proportion to an increase in all inputs.

A **sunk cost** is a cost that has already been incurred and is nonrecoverable. A sunk cost should be ignored in a decision about future actions.

Sunk Costs

To complete our discussion of costs, we need to include the concept of sunk costs. When making decisions, knowing what to ignore is important. Although we have devoted much attention to costs that are important to take into account when making a decision, some costs should be ignored when doing so. This section presents the kind of costs that people should ignore when making decisions—what economists call *sunk costs*—and explains why they should be ignored.

To gain some intuition, consider the following scenario. You own a car that is a few years old, and you have just replaced the brake pads at a cost of $250. But then you find out that the entire brake system is defective and also must be replaced. This will cost you an additional $1,500. Alternatively, you could sell the car and buy another of comparable quality, but with no brake defects, by spending an additional $1,600. What should you do: fix your old car, or sell it and buy another?

Some might say that you should take the latter option. After all, this line of reasoning goes, if you repair your car, you will end up having spent $1,750: $1,500 for the brake system and $250 for the brake pads. If you were instead to sell your old car and buy another, you would spend only $1,600.

But this reasoning, although it sounds plausible, is wrong. It ignores the fact that you have *already* spent $250 on brake pads, and that $250 is *nonrecoverable.* That is, having already been spent, the $250 cannot be recouped. Therefore, it should be ignored and should have no effect on your decision whether to repair your car and keep it or not. From a rational viewpoint, the real cost at this time of repairing and keeping your car is $1,500, not $1,750. So the correct decision is to repair your car and keep it rather than spend $1,600 on a new car.

© Greatstock Photographic Library / Alamy

In this example, the $250 that has already been spent and cannot be recovered is what economists call a **sunk cost.** Sunk costs should be ignored in making decisions because they have no influence on future costs and benefits. It's like the old saying, "There's no use crying over spilled milk": once something can't be recovered, it is irrelevant in making decisions about what to do in the future. This applies equally to individuals, firms, and governments—regardless of how much has been spent on a project in the past, if the future costs exceed the future benefits, the project should not continue.

It is often psychologically hard to ignore sunk costs. And if, in fact, you haven't yet incurred the costs, then you should take them into consideration. That is, if you had known

fyi

There's No Business Like Snow Business

Anyone who has lived both in a snowy city, like Chicago, and in a city that only occasionally experiences significant snowfall, like Washington, D.C., is aware of the differences in total cost that arise from making different choices about fixed cost.

In Washington, even a minor snowfall—say, an inch or two overnight—is enough to create chaos during the next morning's commute. The same snowfall in Chicago has hardly any effect at all. The reason is not that Washingtonians are wimps and Chicagoans are made of sterner stuff; it is that Washington, where it rarely snows, has only a fraction as many snowplows and other snow-clearing equipment as cities where heavy snow is a fact of life.

In this sense Washington and Chicago are like two producers who expect to produce different levels of output, where the "output" is snow removal. Washington, which rarely has significant snow, has chosen a low level of fixed cost in the form of snow-clearing equipment. This makes sense under normal circumstances but leaves the city unprepared when major snow does fall. Chicago, which knows that it will face lots of snow, chooses to accept the higher fixed cost that leaves it in a position to respond effectively.

AP/Wide World Photos

A lesson in returns to scale: cities with higher average annual snowfall maintain larger snowplow fleets.

at the beginning that it would cost $1,750 to repair your car, then the right choice *at that time* would have been to buy a new car for $1,600. But once you have already paid the $250 for brake pads, you should no longer include it in your decision making about your next actions. It may be hard to "let bygones be bygones," but it is the right way to make a decision.

Summing Up Costs: The Short and Long of It

If a firm is to make the best decisions about how much to produce, it has to understand how its costs relate to the quantity of output it chooses to produce. Table 56.1 provides a quick summary of the concepts and measures of cost you have learned about.

table 56.1

Concepts and Measures of Cost

	Measurement	Definition	Mathematical term
Short run	Fixed cost	Cost that does not depend on the quantity of output produced	*FC*
	Average fixed cost	Fixed cost per unit of output	$AFC = FC/Q$
Short run and long run	Variable cost	Cost that depends on the quantity of output produced	*VC*
	Average variable cost	Variable cost per unit of output	$AVC = VC/Q$
	Total cost	The sum of fixed cost (short run) and variable cost	$TC = FC$ (short run) $+ VC$
	Average total cost (average cost)	Total cost per unit of output	$ATC = TC/Q$
	Marginal cost	The change in total cost generated by producing one more unit of output	$MC = \Delta TC/\Delta Q$
Long run	Long-run average total cost	Average total cost when fixed cost has been chosen to minimize average total cost for each level of output	*LRATC*

Module 56 AP Review

Solutions appear at the back of the book.

Check Your Understanding

1. The accompanying table shows three possible combinations of fixed cost and average variable cost. Average variable cost is constant in this example. (It does not vary with the quantity of output produced.)

Choice	*Fixed cost*	*Average variable cost*
1	$8,000	$1.00
2	12,000	0.75
3	24,000	0.25

   a. For each of the three choices, calculate the average total cost of producing 12,000, 22,000, and 30,000 units. For each of these quantities, which choice results in the lowest average total cost?
   b. Suppose that the firm, which has historically produced 12,000 units, experiences a sharp, permanent increase in demand that leads it to produce 22,000 units. Explain how its average total cost will change in the short run and in the long run.
   c. Explain what the firm should do instead if it believes the change in demand is temporary.

2. In each of the following cases, explain whether the firm is likely to experience economies of scale or diseconomies of scale and why.
   a. an interior design firm in which design projects are based on the expertise of the firm's owner
   b. a diamond-mining company

Tackle the Test: Multiple-Choice Questions

1. In the long run,
   a. all inputs are variable.
   b. all inputs are fixed.
   c. some inputs are variable and others are fixed.
   d. a firm will go out of business.
   e. firms increase in size.

2. Which of the following is always considered the long run?
   a. 1 month
   b. 1 year
   c. 5 years
   d. 10 years
   e. none of the above

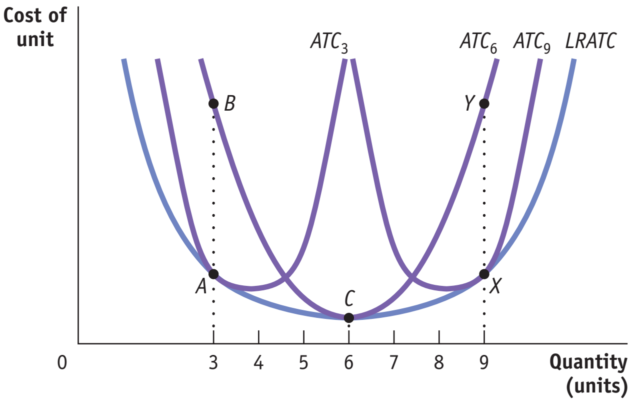
3. Which of the following statements is generally correct?
   I. The long-run average total cost curve is U-shaped.
   II. The short-run average total cost curve is U-shaped.
   III. Firms tend to experience economies of scale at low levels of production and diseconomies of scale at high levels of production.
   a. I only
   b. II only
   c. III only
   d. I and II only
   e. I, II, and III

4. When making decisions, which of the following costs should be ignored?
   a. average costs
   b. total costs
   c. marginal costs
   d. sunk costs
   e. None—no costs should be ignored.

5. Economies of scale will allow which of the following types of cities to lower their average total cost of clearing snow by investing in larger snow plow fleets? Cities with
   a. more people.
   b. more existing snow plows.
   c. less snowfall.
   d. larger budgets.
   e. more snowfall.

Tackle the Test: Free-Response Questions

1. Refer to the graph provided to answer the following questions.



a. The same level of fixed cost that puts the firm at point B when the quantity is 3 minimizes short-run average total cost for what output level?

b. At an output level of 3, is the firm experiencing economies or diseconomies of scale? Explain.

c. In the long run, if the firm expects to produce an output of 9, the firm will produce on which short-run average total cost curve and at which point on the graph?

Answer (5 points)

1 point: 6

1 point: economies of scale

1 point: Because at an output of 3, the LRATC is decreasing.

1 point: In the long run the firm will produce on ATC_9.

1 point: In the long run the firm will produce at point X.

2. Draw a correctly labeled graph showing a short-run average total cost curve and the corresponding long-run average total cost curve. On your graph, identify the areas of economies and diseconomies of scale.

Module 57 Introduction to Market Structure

What you will learn in this Module:

- The meaning and dimensions of market structure
- The four principal types of market structure—perfect competition, monopoly, oligopoly, and monopolistic competition

You may have noticed that this section is titled "Behind the Supply Curve," but we have yet to mention any supply curve. The reason is that to discuss the supply curve in a market, we need to identify the type of market we are looking at. In this module we will learn about the basic characteristics of the four major types of markets in the economy.

Types of Market Structure

The real world holds a mind-boggling array of different markets. Patterns of firm behavior vary as widely as the markets themselves: in some markets firms are extremely competitive; in others, they seem somehow to coordinate their actions to limit competition; and some markets are monopolies in which there is no competition at all. In order to develop principles and make predictions about markets and firm behavior, economists have developed four primary models of market structure: *perfect competition, monopoly, oligopoly,* and *monopolistic competition.*

This system of market structure is based on two dimensions:

- the number of firms in the market (one, few, or many)
- whether the goods offered are identical or *differentiated*

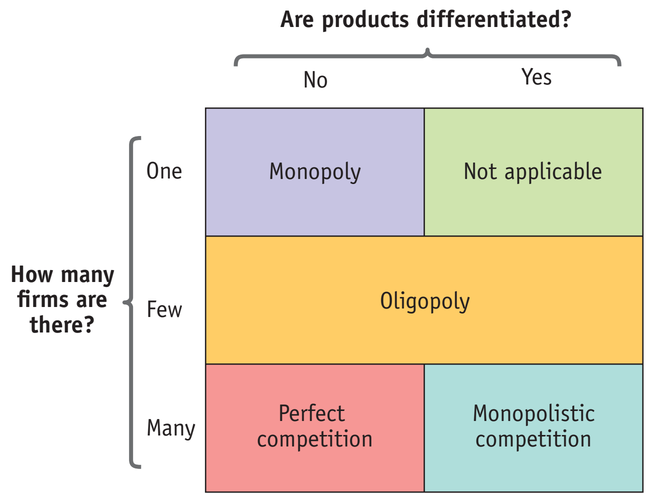
Differentiated goods are goods that are different but considered at least somewhat substitutable by consumers (think Coke versus Pepsi).

Figure 57.1 on the next page provides a simple visual summary of the types of market structure classified according to the two dimensions. In *perfect competition* many firms each sell an identical product. In *monopoly,* a single firm sells a single, undifferentiated product. In *oligopoly,* a few firms—more than one but not a large number—sell products that may be either identical or differentiated. And in *monopolistic competition,* many firms each sell a differentiated product (think of producers of economics textbooks).

figure 57.1

Types of Market Structure

The behavior of any given firm and the market it occupies are analyzed using one of four models of market structure—monopoly, oligopoly, perfect competition, or monopolistic competition. This system for categorizing market structure is based on two dimensions: (1) whether products are differentiated or identical and (2) the number of firms in the industry—one, a few, or many.



Perfect Competition

Suppose that Yves and Zoe are neighboring farmers, both of whom grow organic tomatoes. Both sell their output to the same grocery store chains that carry organic foods; so, in a real sense, Yves and Zoe compete with each other.

Does this mean that Yves should try to stop Zoe from growing tomatoes or that Yves and Zoe should form an agreement to grow fewer? Almost certainly not: there are hundreds or thousands of organic tomato farmers (let's not forget Jennifer and Jason from Module 53!), and Yves and Zoe are competing with all those other growers as well as with each other. Because so many farmers sell organic tomatoes, if any one of them produced more or fewer, there would be no measurable effect on market prices.

When people talk about business competition, they often imagine a situation in which two or three rival firms are struggling for advantage. But economists know that when a business focuses on a few main competitors, it's actually a sign that competition is fairly limited. As the example of organic tomatoes suggests, when the number of competitors is large, it doesn't even make sense to identify rivals and engage in aggressive competition because each firm is too small within the scope of the market to make a significant difference.

We can put it another way: Yves and Zoe are *price-takers*. A firm is a **price-taker** when its actions cannot affect the market price of the good or service it sells. As a result, a price-taking firm takes the market price as given. When there is enough competition—when competition is what economists call "perfect"—then every firm is a price-taker. There is a similar definition for consumers: a **price-taking consumer** is a consumer who cannot influence the market price of the good or service by his or her actions. That is, the market price is unaffected by how much or how little of the good the consumer buys.

Defining Perfect Competition

In a **perfectly competitive market,** all market participants, both consumers and producers, are price-takers. That is, neither consumption decisions by individual consumers nor production decisions by individual producers affect the market price of the good.

The supply and demand model is a model of a perfectly competitive market. It depends fundamentally on the assumption that no individual buyer or seller of a good,

A **price-taking firm** is a firm whose actions have no effect on the market price of the good or service it sells.

A **price-taking consumer** is a consumer whose actions have no effect on the market price of the good or service he or she buys.

A **perfectly competitive market** is a market in which all market participants are price-takers.

such as coffee beans or organic tomatoes, believes that it is possible to individually affect the price at which he or she can buy or sell the good. For a firm, being a price-taker means that the demand curve is a horizontal line at the market price. If the firm charged more than the market price, buyers would go to any of the many alternative sellers of the same product. And it is unnecessary to charge a lower price because, as an insignificantly small part of the perfectly competitive market, the firm can sell all that it wants at the market price.

As a general rule, consumers are indeed price-takers. Instances in which consumers are able to affect the prices they pay are rare. It is, however, quite common for producers to have a significant ability to affect the prices they receive, a phenomenon we'll address later. So the model of perfect competition is appropriate for some but not all markets. An industry in which firms are price-takers is called a **perfectly competitive industry.** Clearly, some industries aren't perfectly competitive; in later modules we'll focus on industries that don't fit the perfectly competitive model.

Under what circumstances will all firms be price-takers? As we'll discover next, there are two necessary conditions for a perfectly competitive industry and a third condition is often present as well.

A **perfectly competitive industry** is an industry in which firms are price-takers.

A firm's **market share** is the fraction of the total industry output accounted for by that firm's output.

A good is a **standardized product,** also known as a **commodity,** when consumers regard the products of different firms as the same good.

Two Necessary Conditions for Perfect Competition

The markets for major grains, such as wheat and corn, are perfectly competitive: individual wheat and corn farmers, as well as individual buyers of wheat and corn, take market prices as given. In contrast, the markets for some of the food items made from these grains—in particular, breakfast cereals—are by no means perfectly competitive. There is intense competition among cereal brands, but not *perfect* competition. To understand the difference between the market for wheat and the market for shredded wheat cereal is to understand the two necessary conditions for perfect competition.

Scott Bauer/ARS/USDA

First, for an industry to be perfectly competitive, it must contain many firms, none of whom have a large **market share.** A firm's market share is the fraction of the total industry output accounted for by that firm's output. The distribution of market share constitutes a major difference between the grain industry and the breakfast cereal industry. There are thousands of wheat farmers, none of whom account for more than a tiny fraction of total wheat sales. The breakfast cereal industry, however, is dominated by four firms: Kellogg's, General Mills, Post, and Quaker Foods. Kellogg's alone accounts for about one-third of all cereal sales. Kellogg's executives know that if they try to sell more corn flakes, they are likely to drive down the market price of corn flakes. That is, they know that their actions influence market prices—due to their tremendous size, changes in their production will significantly affect the overall quantity supplied. It makes sense to assume that firms are price-takers only when they are numerous and relatively small.

Second, an industry can be perfectly competitive only if consumers regard the products of all firms as equivalent. This clearly isn't true in the breakfast cereal market: consumers don't consider Cap'n Crunch to be a good substitute for Wheaties. As a result, the maker of Wheaties has some ability to increase its price without fear that it will lose all its customers to the maker of Cap'n Crunch. Contrast this with the case of a **standardized product,** sometimes known as a **commodity,** which is a product that consumers regard as the same good even when it comes from different firms. Because wheat is a standardized product, consumers regard the output of one wheat producer as a perfect substitute for that of another producer. Consequently, one farmer cannot increase the price for his or her wheat without losing all sales to other wheat farmers. So the second necessary condition for a perfectly competitive industry is that the industry output is a standardized product. (See the FYI that follows.)

fyi

What's a Standardized Product?

A perfectly competitive industry must produce a standardized product. But is it enough for the products of different firms actually to be the same? No: people must also *think* that they are the same. And producers often go to great lengths to convince consumers that they have a distinctive, or *differentiated,* product, even when they don't.

Consider, for example, champagne—not the super-expensive premium champagnes, but the more ordinary stuff. Most people cannot tell the difference between champagne actually produced in the Champagne region of France, where the product originated, and similar products from Spain or California. But the French government has sought and obtained legal protection for the winemakers of Champagne, ensuring that around the world only bubbly wine from that region can be called champagne. If it's from someplace else, all the seller can do is say that it was produced using the *méthode Champenoise.* This creates a differentiation in the minds of consumers and lets the champagne producers of Champagne charge higher prices.

Similarly, Korean producers of *kimchi,* the spicy fermented cabbage that is the Korean national side dish, are doing their best to convince consumers that the same product packaged by Japanese firms is just not the real thing. The purpose is, of course, to ensure higher prices for Korean *kimchi.*

So is an industry perfectly competitive if it sells products that are indistinguishable except in name but that consumers, for whatever reason, don't think are standardized? No. When it comes to defining the nature of competition, the consumer is always right.

An industry has **free entry and exit** when new firms can easily enter into the industry and existing firms can easily leave the industry.

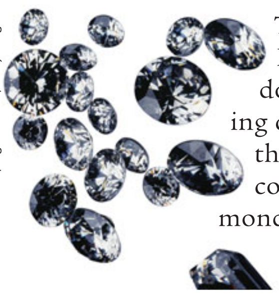
Free Entry and Exit

All perfectly competitive industries have many firms with small market shares, producing a standardized product. Most perfectly competitive industries are also characterized by one more feature: it is easy for new firms to enter the industry or for firms that are currently in the industry to leave. That is, no obstacles in the form of government regulations or limited access to key resources prevent new firms from entering the market. And no additional costs are associated with shutting down a company and leaving the industry. Economists refer to the arrival of new firms into an industry as *entry;* they refer to the departure of firms from an industry as *exit.* When there are no obstacles to entry into or exit from an industry, we say that the industry has **free entry and exit.**

Free entry and exit is not strictly necessary for perfect competition. However, it ensures that the number of firms in an industry can adjust to changing market conditions. And, in particular, it ensures that firms in an industry cannot act to keep other firms out.

To sum up, then, perfect competition depends on two necessary conditions. First, the industry must contain many firms, each having a small market share. Second, the industry must produce a standardized product. In addition, perfectly competitive industries are normally characterized by free entry and exit.

Monopoly


Paul Katz/Digital Vision/Getty Images

The De Beers monopoly of South Africa was created in the 1880s by Cecil Rhodes, a British businessman. By 1880, mines in South Africa already dominated the world's supply of diamonds. There were, however, many mining companies, all competing with each other. During the 1880s Rhodes bought the great majority of those mines and consolidated them into a single company, De Beers. By 1889, De Beers controlled almost all of the world's diamond production.

De Beers, in other words, became a *monopolist*. But what does it mean to be a monopolist? And what do monopolists do?

Defining Monopoly

As we mentioned earlier, the supply and demand model of a market is not universally valid. Instead, it's a model of perfect competition, which is only one of several types of market structure. A market will be perfectly competitive only if there are many firms, all of which produce the same good. Monopoly is the most extreme departure from perfect competition.

A **monopolist** is a firm that is the only producer of a good that has no close substitutes. An industry controlled by a monopolist is known as a **monopoly.**

In practice, true monopolies are hard to find in the modern American economy, partly because of legal obstacles. A contemporary entrepreneur who tried to consolidate all the firms in an industry the way Rhodes did would soon find himself in court, accused of breaking *antitrust* laws, which are intended to prevent monopolies from emerging. Monopolies do, however, play an important role in some sectors of the economy.

A **monopolist** is the only producer of a good that has no close substitutes. An industry controlled by a monopolist is known as a **monopoly.**

To earn economic profits, a monopolist must be protected by a **barrier to entry**—something that prevents other firms from entering the industry.

A **natural monopoly** exists when economies of scale provide a large cost advantage to a single firm that produces all of an industry's output.

Why Do Monopolies Exist?

A monopolist making profits will not go unnoticed by others. (Recall that this is "economic profit," revenue over and above the opportunity costs of the firm's resources.) But won't other firms crash the party, grab a piece of the action, and drive down prices and profits in the long run? If possible, yes, they will. For a profitable monopoly to persist, something must keep others from going into the same business; that "something" is known as a **barrier to entry.** There are four principal types of barriers to entry: control of a scarce resource or input, economies of scale, technological superiority, and government-created barriers.

Control of a Scarce Resource or Input A monopolist that controls a resource or input crucial to an industry can prevent other firms from entering its market. Cecil Rhodes made De Beers into a monopolist by establishing control over the mines that produced the great bulk of the world's diamonds.

Economies of Scale Many Americans have natural gas piped into their homes for cooking and heating. Invariably, the local gas company is a monopolist. But why don't rival companies compete to provide gas?

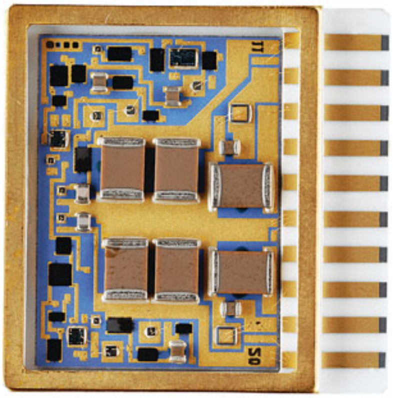
In the early nineteenth century, when the gas industry was just starting up, companies did compete for local customers. But this competition didn't last long; soon local gas companies became monopolists in almost every town because of the large fixed cost of providing a town with gas lines. The cost of laying gas lines didn't depend on how much gas a company sold, so a firm with a larger volume of sales had a cost advantage: because it was able to spread the fixed cost over a larger volume, it had a lower average total cost than smaller firms.

The natural gas industry is one in which average total cost falls as output increases, resulting in economies of scale and encouraging firms to grow larger. In an industry characterized by economies of scale, larger firms are more profitable and drive out smaller ones. For the same reason, established firms have a cost advantage over any potential entrant—a potent barrier to entry. So economies of scale can both give rise to and sustain a monopoly.

A monopoly created and sustained by economies of scale is called a **natural monopoly.** The defining characteristic of a natural monopoly is that it possesses economies of scale over the range of output that is relevant for the industry. The source of this condition is large fixed costs: when large fixed costs are required to operate, a given quantity of output is produced at lower average total cost by one large firm than by two or more smaller firms.

The most visible natural monopolies in the modern economy are local utilities—water, gas, electricity, local land-line phone service, and, in most locations, cable television. As we'll see later, natural monopolies pose a special challenge to public policy.

Technological Superiority A firm that maintains a consistent technological advantage over potential competitors can establish itself as a monopolist. For example, from the 1970s through the 1990s, the chip manufacturer Intel was able to maintain a consistent advantage over potential competitors in both the design and production of microprocessors, the chips that run computers. But technological superiority is typically not a barrier to entry over the longer term: over time competitors will invest in upgrading their technology to match that of the technology leader. In fact, in the last few years Intel found its technological superiority eroded by a competitor, Advanced Micro Devices (also known as AMD), which now produces chips approximately as fast and as powerful as Intel chips.



Photodisc

We should note, however, that in certain high-tech industries, technological superiority is not a guarantee of success against competitors. Some high-tech industries are characterized by *network externalities,* a condition that arises when the value of a good to a consumer rises as the number of other people who also use the good rises. In these industries, the firm possessing the largest network—the largest number of consumers currently using its product—has an advantage over its competitors in attracting new customers, an advantage that may allow it to become a monopolist. Microsoft is often cited as an example of a company with a technologically inferior product—its computer operating system—that grew into a monopolist through the phenomenon of network externalities.

Government-Created Barriers In 1998 the pharmaceutical company Merck introduced Propecia, a drug effective against baldness. Despite the fact that Propecia was very profitable and other drug companies had the know-how to produce it, no other firms challenged Merck's monopoly. That's because the U.S. government had given Merck the sole legal right to produce the drug in the United States. Propecia is an example of a monopoly protected by government-created barriers.

The most important legally created monopolies today arise from *patents* and *copyrights.* A **patent** gives an inventor the sole right to make, use, or sell that invention for a period that in most countries lasts between 16 and 20 years. Patents are given to the creators of new products, such as drugs or mechanical devices. Similarly, a **copyright** gives the creator of a literary or artistic work the sole right to profit from that work, usually for a period equal to the creator's lifetime plus 70 years.

The justification for patents and copyrights is a matter of incentives. If inventors were not protected by patents, they would gain little reward from their efforts: as soon as a valuable invention was made public, others would copy it and sell products based on it. And if inventors could not expect to profit from their inventions, then there would be no incentive to incur the costs of invention in the first place. Likewise for the creators of literary or artistic works. So the law allows a monopoly to exist temporarily by granting property rights that encourage invention and creation. Patents and copyrights are temporary because the law strikes a compromise. The higher price for the good that holds while the legal protection is in effect compensates inventors for the cost of invention; conversely, the lower price that results once the legal protection lapses benefits consumers.

Because the lifetime of the temporary monopoly cannot be tailored to specific cases, this system is imperfect and leads to some missed opportunities. In some cases there can be significant welfare issues. For example, the violation of American drug patents by pharmaceutical companies in poor countries has been a major source of controversy, pitting the needs of poor patients who cannot afford to pay retail drug prices against the interests of drug manufacturers who have incurred high research costs to discover these drugs. To solve this problem, some American drug companies and poor countries have negotiated deals in which the patents are honored but the American companies sell their drugs at deeply discounted prices. (This is an example of *price discrimination,* which we'll learn more about later.)

A **patent** gives an inventor a temporary monopoly in the use or sale of an invention.

A **copyright** gives the creator of a literary or artistic work the sole right to profit from that work.

Oligopoly

An industry with only a few firms is known as an **oligopoly;** a producer in such an industry is known as an **oligopolist.**

Oligopolists compete with each other for sales. But oligopolists aren't like producers in a perfectly competitive industry, who take the market as given. Oligopolists know their decisions about how much to produce will affect the market price. That is, like monopolists, oligopolists have some *market power.* Economists refer to a situation in which firms compete but also possess market power—which enables them to affect market prices—as **imperfect competition.** There are two important forms of imperfect competition: oligopoly and *monopolistic competition.* Of these, oligopoly is probably the more important in practice.

Many familiar goods and services are supplied by only a few competing sellers, which means the industries in question are oligopolies. For example, most air routes are served by only two or three airlines: in recent years, regularly scheduled shuttle service between New York and either Boston or Washington, D.C., has been provided only by Delta and US Airways. Three firms—Chiquita, Dole, and Del Monte, which own huge banana plantations in Central America—control 65% of world banana exports. Most cola beverages are sold by Coca-Cola and Pepsi. This list could go on for many pages.


Photodisc

It's important to realize that an oligopoly isn't necessarily made up of large firms. What matters isn't size per se; the question is how many competitors there are. When a small town has only two grocery stores, grocery service there is just as much an oligopoly as air shuttle service between New York and Washington.

Why are oligopolies so prevalent? Essentially, an oligopoly is the result of the same factors that sometimes produce a monopoly, but in somewhat weaker form. Probably the most important source of oligopolies is the existence of economies of scale, which give bigger firms a cost advantage over smaller ones. When these effects are very strong, as we have seen, they lead to a monopoly; when they are not that strong, they lead to an industry with a small number of firms. For example, larger grocery stores typically have lower costs than smaller stores. But the advantages of large scale taper off once grocery stores are reasonably large, which is why two or three stores often survive in small towns.

Is It an Oligopoly or Not?

In practice, it is not always easy to determine an industry's market structure just by looking at the number of sellers. Many oligopolistic industries contain a number of small "niche" firms, which don't really compete with the major players. For example, the U.S. airline industry includes a number of regional airlines such as New Mexico Airlines, which flies propeller planes between Albuquerque and Carlsbad, New Mexico; if you count these carriers, the U.S. airline industry contains nearly one hundred firms, which doesn't sound like competition among a small group. But there are only a handful of national competitors like American and United, and on many routes, as we've seen, there are only two or three competitors.

To get a better picture of market structure, economists often use two measures of market power: **concentration ratios** and the **Herfindahl–Hirschman Index.** Concentration ratios measure the percentage of industry sales accounted for by the "X" largest firms, where "X" can equal any number of firms. For example, the four-firm concentration ratio is the percentage of sales accounted for by the four largest firms and the eight-firm concentration ratio is the percentage of industry sales accounted for by the eight largest firms. Let's say that the largest four firms account for 25%, 20%, 15%, and10% of industry sales, then the four-firm concentration ratios would equal 70 (25+20+15+10). And if the next largest four firms in that industry account for 9%, 8%, 6%, and 2% of sales, the eight-firm concentration ratio would equal 95 (70 +9+8+6+2). The

An **oligopoly** is an industry with only a small number of firms. A producer in such an industry is known as an **oligopolist.**

When no one firm has a monopoly, but producers nonetheless realize that they can affect market prices, an industry is characterized by **imperfect competition.**

Concentration ratios measure the percentage of industry sales accounted for by the "X" largest firms, for example the four-firm concentration ratio or the eight-firm concentration ratio.

Herfindahl–Hirschman Index, or HHI, is the square of each firm's share of market sales summed over the industry. It gives a picture of the industry market structure.


Courtesy of Henry M. Trotter

four- and eight-firm concentration ratios are the most commonly used. A higher concentration ratio signals a market is more concentrated and thus is more likely to be an oligopoly.

Another measure of market concentration is the Herfindahl-Hirschman index, or HHI. The HHI for an industry is the square of each firm's share of market sales summed over the firms in the industry. Unlike concentration ratios, the HHI takes into account the distribution of market sales among the top firms by squaring each firm's market share, thereby giving more weight to larger firms. For example, if an industry contains only 3 firms and their market shares are 60%, 25%, and 15%, then the HHI for the industry is:

$$\text{HHI} = 60^2 + 25^2 + 15^2 = 4{,}450$$

By squaring each market share, the HHI calculation produces numbers that are much larger when a larger share of an industry output is dominated by fewer firms. This is confirmed by the data in Table 57.1. Here, the indices for industries dominated by a small number of firms, like the personal computer operating systems industry or the wide-body aircraft industry, are many times larger than the index for the retail grocery industry, which has numerous firms of approximately equal size.

table **57.1**

The HHI for Some Oligopolistic Industries

Industry	HHI	Largest firms
PC operating systems	9,182	Microsoft, Linux
Wide-body aircraft	5,098	Boeing, Airbus
Diamond mining	2,338	De Beers, Alrosa, Rio Tinto
Automobiles	1,432	GM, Ford, Chrysler, Toyota, Honda, Nissan, VW
Movie distributors	1,096	Buena Vista, Sony Pictures, 20th Century Fox, Warner Bros., Universal, Paramount, Lionsgate
Internet service providers	750	SBC, Comcast, AOL, Verizon, Road Runner, Earthlink, Charter, Qwest
Retail grocers	321	Walmart, Kroger, Sears, Target, Costco, Walgreens, Ahold, Albertsons

Sources: Canadian Government; Diamond Facts 2006; www.w3counter.com; Planet retail; Autodata; Reuters; ISP Planet; Swivel. Data cover 2006–2007.

Monopolistic Competition

Leo manages the Wonderful Wok stand in the food court of a big shopping mall. He offers the only Chinese food there, but there are more than a dozen alternatives, from Bodacious Burgers to Pizza Paradise. When deciding what to charge for a meal, Leo knows that he must take those alternatives into account: even people who normally prefer stir-fry won't order a $15 lunch from Leo when they can get a burger, fries, and drink for $4.

But Leo also knows that he won't lose all his business even if his lunches cost a bit more than the alternatives. Chinese food isn't the same thing as burgers or pizza. Some people will really be in the mood for Chinese that day, and they will buy from Leo even if they could have dined more cheaply on burgers. Of course, the reverse is also true: even if Chinese is a bit cheaper, some people will choose burgers instead. In other words, Leo does have some market power: he has *some* ability to set his own price.

So how would you describe Leo's situation? He definitely isn't a price-taker, so he isn't in a situation of perfect competition. But you wouldn't exactly call him a monopolist, either. Although he's the only seller of Chinese food in that food court, he does face competition from other food vendors.

Yet it would also be wrong to call him an oligopolist. Oligopoly, remember, involves competition among a small number of interdependent firms in an industry protected by some—albeit limited—barriers to entry and whose profits are highly interdependent. Because their profits are highly interdependent, oligopolists have an incentive to collude, tacitly or explicitly. But in Leo's case there are *lots* of vendors in the shopping mall, too many to make tacit collusion feasible.

Economists describe Leo's situation as one of **monopolistic competition.** Monopolistic competition is particularly common in service industries such as the restaurant and gas station industries, but it also exists in some manufacturing industries. It involves three conditions:

- a large number of competing firms,
- differentiated products, and
- free entry into and exit from the industry in the long run.

Monopolistic competition is a market structure in which there are many competing firms in an industry, each firm sells a differentiated product, and there is free entry into and exit from the industry in the long run.

In a monopolistically competitive industry, each producer has some ability to set the price of her differentiated product. But exactly how high she can set it is limited by the competition she faces from other existing and potential firms that produce close, but not identical, products.

Defining Monopolistic Competition Large Numbers In a monopolistically competitive industry there are many firms. Such an industry does not look either like a monopoly, where the firm faces no competition, or like an oligopoly, where each firm has only a few rivals. Instead, each seller has many competitors. For example, there are many vendors in a big food court, many gas stations along a major highway, and many hotels at a popular beach resort.

Differentiated Products In a monopolistically competitive industry, each firm has a product that consumers view as somewhat distinct from the products of competing firms. Such product differentiation can come in the form of different styles or types, different locations, or different levels of quality. At the same time, though, consumers see these competing products as close substitutes. If Leo's food court contained 15 vendors selling exactly the same kind and quality of food, there would be perfect competition: any seller who tried to charge a higher price would have no customers. But suppose that Wonderful Wok is the only Chinese food vendor, Bodacious Burgers is the only hamburger stand, and so on. The result of this differentiation is that each vendor has some ability to set his or her own price: each firm has some—albeit limited—market power.

Free Entry and Exit in the Long Run In monopolistically competitive industries, new firms, with their own distinct products, can enter the industry freely in the long run. For example, other food vendors would open outlets in the food court if they thought it would be profitable to do so. In addition, firms will exit the industry if they find they are not covering their costs in the long run.

Monopolistic competition, then, differs from the three market structures we have examined so far. It's not the same as perfect competition: firms have some power to set prices. It's not pure monopoly: firms face some competition. And it's not the same as oligopoly: there are many firms and free entry, which eliminates the potential for collusion that is so important in oligopoly. As we'll see in modules 66 and 67, competition among the sellers of differentiated products is the key to understanding how monopolistic competition works.

Now that we have introduced the idea of market structure and presented the four principal models of market structure, we can proceed in the next two sections to use the cost curves we have developed to build each of the four market structure models. These models will allow us to explain and predict firm behavior (e.g., price and quantity determination) and analyze individual markets.

Module 57 AP Review

Solutions appear at the back of the book.

Check Your Understanding

1. In each of the following situations, what type of market structure do you think the industry represents?
   a. There are three producers of aluminum in the world, a good sold in many places.
   b. There are thousands of farms that produce indistinguishable soybeans to thousands of buyers.
   c. Many designers sell high-fashion clothes. Each designer has a distinctive style and a somewhat loyal clientele.
   d. A small town in the middle of Alaska has one bicycle shop.

Tackle the Test: Multiple-Choice Questions

1. Which of the following is true for a perfectly competitive industry?
   I. There are many firms, each with a large market share.
   II. The firms in the industry produce a standardized product.
   III. There are barriers to entry and exit.
   a. I only
   b. II only
   c. III only
   d. I and II only
   e. I, II, and III
2. Which of the following is true for a monopoly?
   I. There is only one firm.
   II. The firm produces a product with many close substitutes.
   III. The industry has free entry and exit.
   a. I only
   b. II only
   c. III only
   d. I and II only
   e. I, II, and III
3. Which of the following is true for an oligopoly?
   I. There are a few firms, each with a large market share.
   II. The firms in the industry are interdependent.
   III. The industry experiences diseconomies of scale.
   a. I only
   b. II only
   c. III only
   d. I and II only
   e. I, II, and III
4. Which of the following is true for a monopolistically competitive industry?
   I. There are many firms, each with a small market share.
   II. The firms in the industry produce a standardized product.
   III. Firms are price-takers.
   a. I only
   b. II only
   c. III only
   d. I and II only
   e. I, II, and III
5. Which of the following is an example of differentiated products?
   a. Coke and Pepsi
   b. automobiles and bicycles
   c. trucks and gasoline
   d. stocks and bonds
   e. gold and silver

Tackle the Test: Free-Response Questions

1. For each of the following characteristics, indicate which market structure(s) exhibit that characteristic.
   a. many sellers
   b. price-takers
   c. barriers to entry
   d. differentiated product

Answer (7 points)

a. 1 point: perfect competition

1 point: monopolistic competition

b. 1 point: perfect competition

c. 1 point: monopoly

1 point: oligopoly

d. 1 point: oligopoly

1 point: monopolistic competition

2. a. Draw a correctly labeled graph of a perfectly competitive firm's demand curve if the market price is \$10.
   b. What does the firm's marginal revenue equal any time it sells one more unit of its output?

Section 10 Review

Summary

1. The cost of using a resource for a particular activity is the opportunity cost of that resource. Some opportunity costs are **explicit costs;** they involve a direct payment of cash. Other opportunity costs, however, are **implicit costs;** they involve no outlay of money but represent the inflows of cash that are forgone. Both explicit and implicit costs should be taken into account when making decisions. Firms use capital and their owners' time, so firms should base decisions on **economic profit,** which takes into account implicit costs such as the opportunity cost of the owners' time and the **implicit cost of capital. Accounting profit,** which firms calculate for the purposes of taxes and public reporting, is often considerably larger than economic profit because it includes only explicit costs and depreciation, not implicit costs. Finally, **normal profit** is a term used to describe an economic profit equal to zero—a profit just high enough to justify the use of resources in an activity.
2. A producer chooses output according to the **optimal output rule:** produce the quantity at which marginal revenue equals marginal cost. The **marginal revenue** for each unit of output is shown by the **marginal revenue curve.** More generally, the **principle of marginal analysis** suggests that every activity should continue until marginal benefit equals marginal cost.
3. The relationship between inputs and output is represented by a firm's **production function.** In the **short run,** the quantity of a **fixed input** cannot be varied but the quantity of a **variable input,** by definition, can. In the **long run,** the quantities of all inputs can be varied. For a given amount of the fixed input, the **total product curve** shows how the quantity of output changes as the quantity of the variable input changes. The **marginal product** of an input is the increase in output that results from using one more unit of that input.
4. There are **diminishing returns to an input** when its marginal product declines as more of the input is used, holding the quantity of all other inputs fixed.
5. **Total cost,** represented by the **total cost curve,** is equal to the sum of **fixed cost,** which does not depend on output, and **variable cost,** which does depend on output. Due to diminishing returns, marginal cost, the increase in total cost generated by producing one more unit of output, normally increases as output increases.
6. **Average total cost** (also known as **average cost**) is the total cost divided by the quantity of output. Economists believe that **U-shaped average total cost curves** are typical because average total cost consists of two parts: **average fixed cost,** which falls when output increases (the spreading effect), and **average variable cost,** which rises with output (the diminishing returns effect).
7. When average total cost is U-shaped, the bottom of the U is the level of output at which average total cost is minimized, the point of **minimum-cost output.** This is also the point at which the **marginal cost curve** crosses the average total cost curve from below. Due to gains from specialization, the marginal cost curve may slope downward initially before sloping upward, giving it a "swoosh" shape.
8. In the long run, a firm can change its fixed input and its level of fixed cost. By accepting higher fixed cost, a firm can lower its variable cost for any given output level, and vice versa. The **long-run average total cost curve** shows the relationship between output and average total cost when fixed cost has been chosen to minimize average total cost at each level of output. A firm moves along its short-run average total cost curve as it changes the quantity of output, and it returns to a point on both its short-run and long-run average total cost curves once it has adjusted fixed cost to its new output level.
9. As output increases, there are **economies of scale** if long-run average total cost decreases and **diseconomies of scale** if long-run average total cost increases. As all inputs are increased by the same proportion, there are **increasing returns to scale** if output increases by a larger proportion than the inputs; **decreasing returns to scale** if output increases by a smaller proportion; and **constant returns to scale** if output increases by the same proportion.
10. **Sunk costs** are expenditures that have already been made and cannot be recovered. Sunk costs should be ignored in making decisions about future actions because what is important is a comparison of future costs and future benefits.
11. There are four main types of market structure based on the number of firms in the industry and product differentiation: perfect competition, monopoly, oligopoly, and monopolistic competition.
12. A **monopolist** is a producer who is the sole supplier of a good without close substitutes. An industry controlled by a monopolist is a **monopoly.**
13. To persist, a monopoly must be protected by a **barrier to entry.** This can take the form of control of a natural resource or input, increasing returns to scale that give rise to a **natural monopoly,** technological superiority, or government rules that prevent entry by other firms, such as **patents** or **copyrights.**

14. In a **perfectly competitive market** all firms are **price-taking firms** and all consumers are **price-taking consumers**—no one's actions can influence the market price. Consumers are normally price-takers, but firms often are not. In a **perfectly competitive industry,** every firm in the industry is a price-taker.

15. There are two necessary conditions for a perfectly competitive industry: there are many firms, none of which has a large **market share,** and the industry produces a **standardized product** or **commodity**—goods that consumers regard as equivalent. A third condition is often satisfied as well: **free entry and exit** into and from the industry.

16. Many industries are **oligopolies:** there are only a few sellers. Oligopolies exist for more or less the same reasons that monopolies exist, but in weaker form. They are characterized by **imperfect competition:** firms compete but possess some market power.

17. **Monopolistic competition** is a market structure in which there are many competing firms, each producing a differentiated product, and there is free entry and exit in the long run. Product differentiation takes three main forms: by style or type, by location, and by quality. The extent of imperfect competition can be measured by the **concentration ratio,** or the **Herfindahl-Hirschman Index.**

Key Terms

Explicit cost, p. 530
Implicit cost, p. 530
Accounting profit, p. 531
Economic profit, p. 532
Implicit cost of capital, p. 532
Normal profit, p. 534
Principle of marginal analysis, p. 537
Marginal revenue, p. 537
Optimal output rule, p. 537
Marginal cost curve, p. 538
Marginal revenue curve, p. 538
Production function, p. 542
Fixed input, p. 542
Variable input, p. 542
Long run, p. 542
Short run, p. 542
Total product curve, p. 543
Marginal product, p. 543
Diminishing returns to an input, p. 545
Fixed cost, p. 548
Variable cost, p. 548
Total cost, p. 548
Total cost curve, p. 549
Average total cost, p. 552
Average cost, p. 552
U-shaped average total cost curve, p. 553
Average fixed cost, p. 553
Average variable cost, p. 553
Minimum-cost output, p. 555
Long-run average total cost curve, p. 561
Economies of scale, p. 562
Increasing returns to scale, p. 562
Diseconomies of scale, p. 562
Decreasing returns to scale, p. 563
Constant returns to scale, p. 563
Sunk cost, p. 563
Price-taking firm, p. 568
Price-taking consumer, p. 568
Perfectly competitive market, p. 568
Perfectly competitive industry, p. 569
Market share, p. 569
Standardized product, p. 569
Commodity, p. 569
Free entry and exit, p. 570
Monopolist, p. 571
Monopoly, p. 571
Barrier to entry, p. 571
Natural monopoly, p. 571
Patent, p. 572
Copyright, p. 572
Oligopoly, p. 573
Oligopolist, p. 573
Imperfect competition, p. 573
Concentration ratios, p. 573
Herfindahl–Hirschman Index, p. 573
Monopolistic competition, p. 575

Problems

1. Hiro owns and operates a small business that provides economic consulting services. During the year he spends $55,000 on traveling to clients and other expenses, and the computer that he owns depreciates by $2,000. If he didn't use the computer, he could sell it and earn yearly interest of $100 on the money created through this sale. Hiro's total revenue for the year is $100,000. Instead of working as a consultant for the year, he could teach economics at a small local college and make a salary of $50,000.

a. What is Hiro's accounting profit?

b. What is Hiro's economic profit?

c. Should Hiro continue working as a consultant, or should he teach economics instead?

2. Jackie owns and operates a Web-design business. Her computing equipment depreciates by $5,000 per year. She runs the business out of a room in her home. If she didn't use the room as her business office, she could rent it out for $2,000 per year. Jackie knows that if she didn't run her own business, she could return to her previous job at a large software company that would pay her a salary of $60,000 per year. Jackie has no other expenses.

a. How much total revenue does Jackie need to make in order to break even in the eyes of her accountant? That is, how much total revenue would give Jackie an accounting profit of just zero?

b. How much total revenue does Jackie need to make in order for her to want to remain self-employed? That is, how much total revenue would give Jackie an economic profit of just zero?

3. You own and operate a bike store. Each year, you receive revenue of $200,000 from your bike sales, and it costs you $100,000 to obtain the bikes. In addition, you pay $20,000 for electricity, taxes, and other expenses per year. Instead of running the bike store, you could become an accountant and receive a yearly salary of $40,000. A large clothing retail chain wants to expand and offers to rent the store from you for $50,000 per year. How do you explain to your friends that despite making a profit, it is too costly for you to continue running your store?

4. Suppose you have just paid a nonrefundable fee of $1,000 for your meal plan for this academic term. This allows you to eat dinner in the cafeteria every evening.

a. You are offered a part-time job in a restaurant where you can eat for free each evening. Your parents say that you should eat dinner in the cafeteria anyway, since you have already paid for those meals. Are your parents right? Explain why or why not.

b. You are offered a part-time job in a different restaurant where, rather than being able to eat for free, you receive only a large discount on your meals. Each meal there will cost you $2; if you eat there each evening this semester, it will add up to $200. Your roommate says that you should eat in the restaurant since it costs less than the $1,000 that you paid for the meal plan. Is your roommate right? Explain why or why not.

5. You have bought a $10 ticket in advance for the college soccer game, a ticket that cannot be resold. You know that going to the soccer game will give you a benefit equal to $20. After you have bought the ticket, you hear that there will be a professional baseball post-season game at the same time. Tickets to the baseball game cost $20, and you know that going to the baseball game will give you a benefit equal to $35. You tell your friends the following: "If I had known about the baseball game before buying the ticket to the soccer game, I would have gone to the baseball game instead. But now that I already have the ticket to the soccer game, it's better for me to just go to the soccer game." Are you making the correct decision? Justify your answer by calculating the benefits and costs of your decision.

6. You are the manager of a gym, and you have to decide how many customers to admit each hour. Assume that each customer stays exactly one hour. Customers are costly to admit because they inflict wear and tear on the exercise equipment. Moreover, each additional customer generates more wear and tear than the customer before. As a result, the gym faces increasing marginal cost. The accompanying table shows the marginal cost associated with each number of customers per hour.

Quantity of customers per hour	Marginal cost of customer
0	
	$14.00
1	
	14.50
2	
	15.00
3	
	15.50
4	
	16.00
5	
	16.50
6	
	17.00
7	

a. Suppose that each customer pays $15.25 for a one-hour workout. Use the principle of marginal analysis to find the optimal number of customers that you should admit per hour.

b. You increase the price of a one-hour workout to $16.25. What is the optimal number of customers per hour that you should admit now?

7. Georgia and Lauren are economics students who go to a karate class together. Both have to choose how many classes to go to per week. Each class costs $20. The accompanying table shows Georgia's and Lauren's estimates of the marginal benefit that each of them gets from each class per week.

Quantity of Classes	Lauren's marginal benefit of each class	Georgia's marginal benefit of each class
0		
	$23	$28
1		
	19	22
2		
	14	15
3		
	8	7
4		

a. Use marginal analysis to find Lauren's optimal number of karate classes per week. Explain your answer.

b. Use marginal analysis to find Georgia's optimal number of karate classes per week. Explain your answer.

8. Changes in the prices of key commodities can have a significant impact on a company's bottom line. According to a September 27, 2007, article in the *Wall Street Journal,* "Now, with oil, gas and electricity prices soaring, companies are beginning to realize that saving energy can translate into dramatically lower costs." Another *Wall Street Journal* article, dated September 9, 2007, states, "Higher grain prices are taking an increasing financial toll." Energy is an input into virtually all types of production; corn is an input into the production of beef, chicken, high-fructose corn syrup, and ethanol (the gasoline substitute fuel).

a. Explain how the cost of energy can be both a fixed cost and a variable cost for a company.

b. Suppose energy is a fixed cost and energy prices rise. What happens to the company's average total cost curve? What happens to its marginal cost curve? Illustrate your answer with a diagram.

c. Explain why the cost of corn is a variable cost but not a fixed cost for an ethanol producer.

d. When the cost of corn goes up, what happens to the average total cost curve of an ethanol producer? What happens to its marginal cost curve? Illustrate your answer with a diagram.

9. Marty's Frozen Yogurt is a small shop that sells cups of frozen yogurt in a university town. Marty owns three frozen-yogurt machines. His other inputs are refrigerators, frozen-yogurt mix, cups, sprinkle toppings, and, of course, workers. He estimates that his daily production function when he varies the number of workers employed (and at the same time, of course, yogurt mix, cups, and so on) is as shown in the accompanying table.

Quantity of labor (workers)	Quantity of frozen yogurt (cups)
0	0
1	110
2	200
3	270
4	300
5	320
6	330

a. What are the fixed inputs and variable inputs in the production of cups of frozen yogurt?

b. Draw the total product curve. Put the quantity of labor on the horizontal axis and the quantity of frozen yogurt on the vertical axis.

c. What is the marginal product of the first worker? The second worker? The third worker? Why does marginal product decline as the number of workers increases?

10. The production function for Marty's Frozen Yogurt is given in Problem 9. Marty pays each of his workers $80 per day. The cost of his other variable inputs is $0.50 per cup of yogurt. His fixed cost is $100 per day.

a. What is Marty's variable cost and total cost when he produces 110 cups of yogurt? 200 cups? Calculate variable and total cost for every level of output given in Problem 9.

b. Draw Marty's variable cost curve. On the same diagram, draw his total cost curve.

c. What is the marginal cost per cup for the first 110 cups of yogurt? For the next 90 cups? Calculate the marginal cost for all remaining levels of output.

11. The production function for Marty's Frozen Yogurt is given in Problem 9. The costs are given in Problem 10.

a. For each of the given levels of output, calculate the average fixed cost (*AFC*), average variable cost (*AVC*), and average total cost (*ATC*) per cup of frozen yogurt.

b. On one diagram, draw the *AFC, AVC,* and *ATC* curves.

c. What principle explains why the *AFC* declines as output increases? What principle explains why the *AVC* increases as output increases? Explain your answers.

d. How many cups of frozen yogurt are produced when average total cost is minimized?

12. The accompanying table shows a car manufacturer's total cost of producing cars.

Quantity of cars	*TC*
0	$500,000
1	540,000
2	560,000
3	570,000
4	590,000
5	620,000
6	660,000
7	720,000
8	800,000
9	920,000
10	1,100,000

a. What is this manufacturer's fixed cost?

b. For each level of output, calculate the variable cost (*VC*). For each level of output except zero, calculate the average variable cost (*AVC*), average total cost (*ATC*), and average fixed cost (*AFC*). What is the minimum-cost output?

c. For each level of output, calculate this manufacturer's marginal cost (*MC*).

d. On one diagram, draw the manufacturer's *AVC, ATC,* and *MC* curves.

13. Labor costs represent a large percentage of total costs for many firms. According to a September 1, 2007, *Wall Street Journal* article, U.S. labor costs were up 0.9% during the preceding three months and 0.8% over the three months preceding those.

a. When labor costs increase, what happens to average total cost and marginal cost? Consider a case in which labor costs are only variable costs and a case in which they are both variable and fixed costs.

An increase in labor productivity means each worker can produce more output. Recent data on productivity show that labor productivity in the U.S. nonfarm business sector grew 2% for each of the years 2005, 2006, and 2007. Annual growth in labor productivity averaged 1.5% from the mid-1970s to mid-1990s, 2.6% in the past decade, and 4% for a couple of years in the early 2000s.

b. When productivity growth is positive, what happens to the total product curve and the marginal product of labor curve? Illustrate your answer with a diagram.

c. When productivity growth is positive, what happens to the marginal cost curve and the average total cost curve? Illustrate your answer with a diagram.

d. If labor costs are rising over time on average, why would a company want to adopt equipment and methods that increase labor productivity?

14. Magnificent Blooms is a florist specializing in floral arrangements for weddings, graduations, and other events. The firm has a fixed cost associated with space and equipment of $100 per day. Each worker is paid $50 per day. The daily production function for Magnificent Blooms is shown in the accompanying table.

Quantity of labor (workers)	Quantity of floral arrangements
0	0
1	5
2	9
3	12
4	14
5	15

a. Calculate the marginal product of each worker. What principle explains why the marginal product per worker declines as the number of workers employed increases?

b. Calculate the marginal cost of each level of output. What principle explains why the marginal cost per floral arrangement increases as the number of arrangements increases?

15. You have the information shown in the accompanying table about a firm's costs. Complete the missing data.

Quantity	*TC*	*MC*	*ATC*	*AVC*
0	$20		—	—
		$20		
1	?		?	?
		10		
2	?		?	?
		16		
3	?		?	?
		20		
4	?		?	?
		24		
5	?		?	?

16. Evaluate each of the following statements. If a statement is true, explain why; if it is false, identify the mistake and try to correct it.

a. A decreasing marginal product tells us that marginal cost must be rising.

b. An increase in fixed cost increases the minimum-cost output.

c. An increase in fixed cost increases marginal cost.

d. When marginal cost is above average total cost, average total cost must be falling.

17. Mark and Jeff operate a small company that produces souvenir footballs. Their fixed cost is $2,000 per month. They can hire workers for $1,000 per worker per month. Their monthly production function for footballs is as given in the accompanying table.

Quantity of labor (workers)	Quantity of footballs
0	0
1	300
2	800
3	1,200
4	1,400
5	1,500

a. For each quantity of labor, calculate average variable cost (*AVC*), average fixed cost (*AFC*), average total cost (*ATC*), and marginal cost (*MC*).

b. On one diagram, draw the *AVC*, *ATC*, and *MC* curves.

c. At what level of output is Mark and Jeff's average total cost minimized?

18. You produce widgets. Currently you produce 4 widgets at a total cost of $40.

a. What is your average total cost?

b. Suppose you could produce one more (the fifth) widget at a marginal cost of $5. If you do produce that fifth widget, what will your average total cost be? Has your average total cost increased or decreased? Why?

c. Suppose instead that you could produce one more (the fifth) widget at a marginal cost of $20. If you do produce that fifth widget, what will your average total cost be? Has your average total cost increased or decreased? Why?