section 11

Market Structures: Perfect Competition and Monopoly

Section 10 explained how factors including the number of firms in the industry, the type of product sold, and the existence of barriers to entry determine the market power of firms. We learned about the four basic market structures—perfect competition, monopoly, oligopoly, and monopolistic competition. We can think about these structures as falling along a spectrum from perfect competition at one end to monopoly at the other, with monopolistic competition and oligopoly lying in between. To shed more light on the market structure spectrum, consider two very different markets introduced in previous sections: the market for organic tomatoes and the market for diamonds.

In the United States, a growing interest in healthy living has steadily increased the demand for products such as organically grown fruits and vegetables. Over the past decade, the markets for these products have been healthy as well, with an average growth rate of 20% per year. It costs a bit more to grow crops without chemical fertilizers and pesticides, but consumers are willing to pay higher prices for the benefits of fruits and vegetables grown the natural way. The farmers in each area who pioneered organic farming techniques had little competition and many prospered thanks to these higher prices.

But with profits as a lure for expanded production, the high prices were unlikely to persist. Over time, farmers already producing organically would increase their capacity, and conventional farmers would enter the organic food fray, increasing supply and driving down price. With a large and growing number of buyers and sellers, undifferentiated products, and few barriers to entry, the organic food market increasingly resembles a *perfectly competitive* market.

Bjorn Andren/Nordic Photos/Getty Images

In contrast, the market for diamonds is dominated by one supplier, De Beers. For generations, diamonds have been valued not just for their attractive appearance, but also for their rarity. But geologists will tell you that diamonds aren't all that rare. In fact, they are fairly common and only seem rare compared to other gem-quality stones. This is because De Beers *makes* them rare: the company controls most of the world's diamond mines and limits the quantity supplied to the market. This makes De Beers resemble a *monopolist,* the sole (or almost sole) producer of a good. Because De Beers controls so much of the world's diamond supply, other firms have considerable difficulty trying to enter the diamond market and increase the quantity of the gems available.

In this section we will study how markets like those for organic tomatoes and diamonds differ, and how these markets respond to market conditions. We will see how firms positioned at opposite ends of the spectrum of market power—from perfect competition to monopoly—make key decisions about output and prices. Then, in Section 12, we will complete our exploration of market structure with a closer look at oligopoly and monopolistic competition.

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

- How a price-taking firm determines its profit-maximizing quantity of output
- How to assess whether or not a competitive firm is profitable

Module 58 Introduction to Perfect Competition

Recall the example of the market for organic tomatoes from our discussions in Section 10. Jennifer and Jason run an organic tomato farm. But many other organic tomato farmers, such as Yves and Zoe, sell their output to the same grocery store chains. Since organic tomatoes are a standardized product, consumers don't care which farmer produces the organic tomatoes they buy. And because so many farmers sell organic tomatoes, no individual farmer has a large market share, which means that no individual farmer can have a measurable effect on market prices. These farmers are price-taking producers and their customers are price-taking consumers. The market for organic tomatoes meets the two necessary conditions for perfect competition: there are many producers each with a small market share, and the firms produce a standardized product. In this module, we build the model of perfect competition and use it to look at a representative firm in the market.

Production and Profits

Jennifer and Jason's tomato farm will maximize its profit by producing bushels of tomatoes up to the point at which marginal revenue equals marginal cost. We know this from the producer's *optimal output rule* introduced in Module 53—profit is maximized by producing the quantity at which the marginal revenue of the last unit produced is equal to its marginal cost. Always remember, $MR = MC$ at the optimal quantity of output. This will be true for any profit-maximizing firm in any market structure.

We can review how to apply the optimal output rule with the help of Table 58.1, which provides various short-run cost measures for Jennifer and Jason's farm. The second column contains the farm's variable cost, and the third column shows its total cost of output based on the assumption that the farm incurs a fixed cost of $14. The fourth column shows their marginal cost. Notice that, in this example, the marginal cost initially falls as output rises but then begins to increase, so that the marginal cost curve has the familiar "swoosh" shape.

The fifth column contains the farm's marginal revenue, which has an important feature: Jennifer and Jason's marginal revenue is constant at $18 for every output level.

table **58.1**

Short-Run Costs for Jennifer and Jason's Farm

Quantity of tomatoes *Q* (bushels)	Variable cost *VC*	Total cost *TC*	Marginal cost of bushel $MC = \Delta TC/\Delta Q$	Marginal revenue of bushel *MR*	Net gain of bushel = $MR - MC$
0	$0	$14			
			$16	$18	$2
1	16	30			
			6	18	12
2	22	36			
			8	18	10
3	30	44			
			12	18	6
4	42	56			
			16	18	2
5	58	72			
			20	18	−2
6	78	92			
			24	18	−6
7	102	116			

The sixth and final column shows the calculation of the net gain per bushel of tomatoes, which is equal to marginal revenue minus marginal cost—or, equivalently in this case, market price minus marginal cost. As you can see, it is positive for the 1st through 5th bushels; producing each of these bushels raises Jennifer and Jason's profit. For the 6th bushel, however, net gain is negative: producing it would decrease, not increase, profit. So 5 bushels are Jennifer and Jason's profit-maximizing output; it is the level of output at which marginal cost rises from a level below market price to a level above market price, passing through the market price of $18 along the way.


Anthony-Masterson/Digital Vision/Getty Images

This example illustrates an application of the optimal output rule to the particular case of a price-taking firm—the **price-taking firm's optimal output rule:** *price equals marginal cost at the price-taking firm's optimal quantity of output.* That is, a price-taking firm's profit is maximized by producing the quantity of output at which the market price is equal to the marginal cost of the last unit produced. Why? Because *in the case of a price-taking firm, marginal revenue is equal to the market price.* A price-taking firm cannot influence the market price by its actions. It always takes the market price as given because it cannot lower the market price by selling more or raise the market price by selling less. So, for a price-taking firm, the additional revenue generated by producing one more unit is always the market price. We will need to keep this fact in mind in future modules, in which we will learn that in the three other market structures, firms are *not* price takers. Therefore, marginal revenue is *not* equal to the market price.

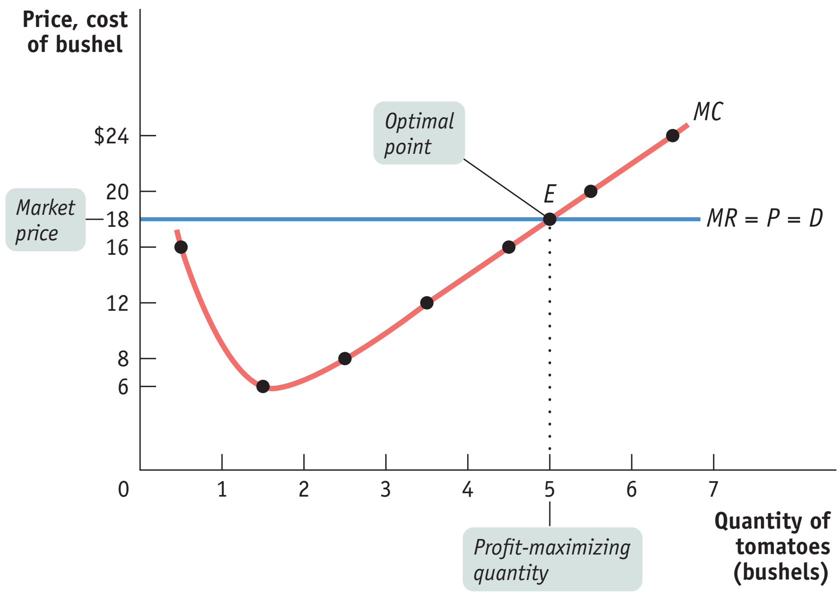
Figure 58.1 on the next page shows Jennifer and Jason's profit-maximizing quantity of output. The figure shows the marginal cost curve, *MC,* drawn from the data in the fourth column of Table 58.1. We plot the marginal cost of increasing output from 1 to 2 bushels halfway between 1 and 2, and likewise for each incremental change. The horizontal line at $18 is Jennifer and Jason's marginal revenue curve. Remember from Module 53 that whenever a firm is a price-taker, its marginal revenue curve is a horizontal line at the market price: it can sell as much as it likes at the market price. Regardless of whether it sells more or less, the market price is unaffected. In effect, the individual firm faces a horizontal, perfectly elastic demand curve for its output—an individual demand curve that is equivalent to its marginal revenue curve. In fact, the horizontal line with the height of the market price represents the perfectly competitive

The **price-taking firm's optimal output rule** says that a price-taking firm's profit is maximized by producing the quantity of output at which the market price is equal to the marginal cost of the last unit produced.

figure 58.1

The Price-Taking Firm's Profit-Maximizing Quantity of Output

At the profit-maximizing quantity of output, the market price 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 and represents the firm's demand curve. Here, the profit-maximizing point is at an output of 5 bushels of tomatoes, the output quantity at point *E*.



firm's demand, marginal revenue, and *average revenue*—the average amount of revenue taken in per unit—because price equals average revenue whenever every unit is sold for the same price. The marginal cost curve crosses the marginal revenue curve at point *E*. Sure enough, the quantity of output at *E* is 5 bushels.

Does this mean that the price-taking 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 it produce at all? If the answer to that question is yes, it 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.

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. In the next module we'll see that unprofitable firms shut down in the long run, but tolerate losses in the short run up to a certain point.

When Is Production Profitable?

Remember from Module 52 that firms make their production decisions with the goal of maximizing *economic profit*—a measure based on the opportunity cost of resources used by the firm. 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 all of the opportunity cost of resources owned by the firm and used in the production of output, while accounting profit does not. A firm may make positive accounting profit while making zero or even negative economic profit. It's important to understand that a firm's decisions of how much to produce, and whether or not to stay in business, should be based on economic profit, not accounting profit.

iStockphoto

So we will assume, as usual, that the cost numbers given in Table 58.1 include all costs, implicit as well as explicit. What determines whether Jennifer

table **58.2**

Short-Run Average Costs for Jennifer and Jason's Farm

Quantity of tomatoes *Q* (bushels)	Variable cost *VC*	Total cost *TC*	Short-run average variable cost of bushel *AVC* = *VC*/*Q*	Short-run average total cost of bushel *ATC* = *TC*/*Q*
1	$16.00	$30.00	$16.00	$30.00
2	22.00	36.00	11.00	18.00
3	30.00	44.00	10.00	14.67
4	42.00	56.00	10.50	14.00
5	58.00	72.00	11.60	14.40
6	78.00	92.00	13.00	15.33
7	102.00	116.00	14.57	16.57

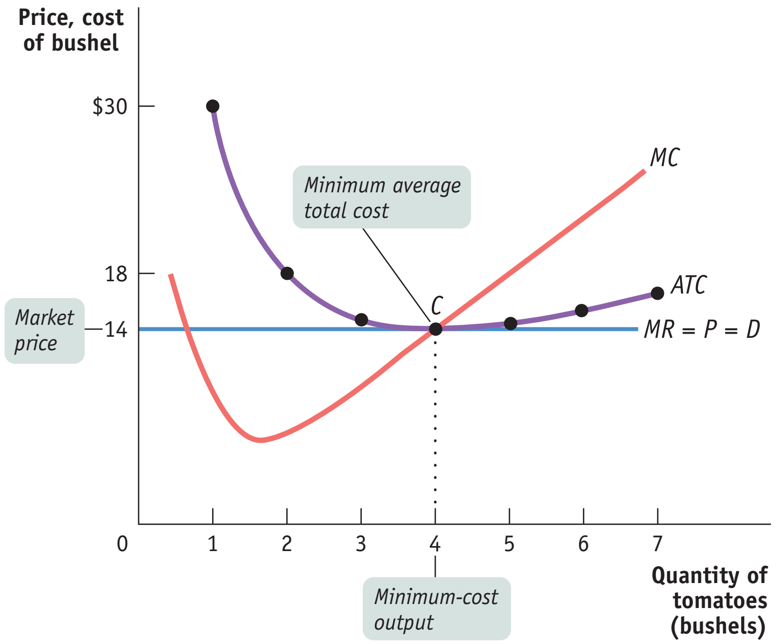
and Jason's farm earns a profit or generates a loss? This depends on the market price of tomatoes—specifically, *whether the market price is more or less than the farm's minimum average total cost.*

In Table 58.2 we calculate short-run average variable cost and short-run average total cost for Jennifer and Jason's farm. These are short-run values because we take fixed cost as given. (We'll turn to the effects of changing fixed cost shortly.) The short-run average total cost curve, *ATC,* is shown in Figure 58.2, along with the marginal cost curve, *MC,* from Figure 58.1. As you can see, average total cost is minimized at point *C,* corresponding to an output of 4 bushels—the *minimum-cost output*—and an average total cost of $14 per bushel.

figure **58.2**

Costs and Production in the Short Run

This figure shows the marginal cost curve, *MC,* and the short-run average total cost curve, *ATC.* When the market price is $14, output will be 4 bushels of tomatoes (the minimum-cost output), represented by point *C.* The price of $14 is equal to the firm's minimum average total cost, so at this price the firm breaks even.



To see how these curves can be used to decide whether production is profitable or unprofitable, recall that profit is equal to total revenue minus total cost, $TR - TC$. This means:

- If the firm produces a quantity at which $TR > TC$, the firm is profitable.
- If the firm produces a quantity at which $TR = TC$, the firm breaks even.
- If the firm produces a quantity at which $TR < TC$, the firm incurs a loss.

We can also express this idea in terms of revenue and cost per unit of output. If we divide profit by the number of units of output, Q, we obtain the following expression for profit per unit of output:

$$\text{(58-1)} \quad \text{Profit}/Q = TR/Q - TC/Q$$

TR/Q is average revenue, which is the market price. TC/Q is average total cost. So a firm is profitable if the market price for its product is more than the average total cost of the quantity the firm produces; a firm experiences losses if the market price is less than the average total cost of the quantity the firm produces. This means:

- If the firm produces a quantity at which $P > ATC$, the firm is profitable.
- If the firm produces a quantity at which $P = ATC$, the firm breaks even.
- If the firm produces a quantity at which $P < ATC$, the firm incurs a loss.

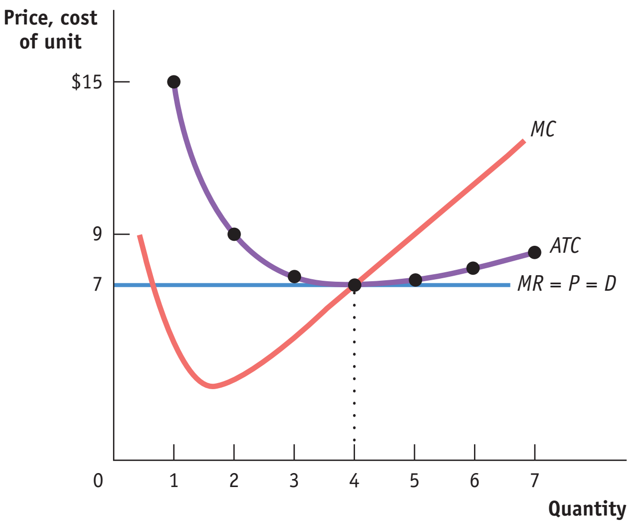
In summary, in the short run a firm will maximize profit by producing the quantity of output at which $MC = MR$. A perfectly competitive firm is a price-taker, so it can sell as many units of output as it would like at the market price. This means that for a perfectly competitive firm it is always true that $MR = P$. The firm is profitable, or breaks even, as long as the market price is greater than, or equal to, average total cost. In the next module, we develop the perfect competition model using graphs to analyze the firm's level of profit.

Module 58 AP Review

Solutions appear at the back of the book.

Check Your Understanding

1. Refer to the graph provided.



a. At what level of output does the firm maximize profit? Explain how you know.
b. At the profit-maximizing quantity of output, is the firm profitable, does it just break even, or does it earn a loss? Explain.

2. If a firm has a total cost of $500 at a quantity of 50 units, and it is at that quantity that average total cost is minimized for the firm, what is the lowest price that would allow the firm to break even (that is, earn a normal profit)? Explain.

Tackle the Test: Multiple-Choice Questions

1. A perfectly competitive firm will maximize profit at the quantity at which the firm's marginal revenue equals
   a. price.
   b. average revenue.
   c. total cost.
   d. marginal cost.
   e. demand.

2. Which of the following is correct for a perfectly competitive firm?
   I. The marginal revenue curve is the demand curve.
   II. The firm maximizes profit when price equals marginal cost.
   III. The market demand curve is horizontal.
   a. I only
   b. II only
   c. III only
   d. I and II only
   e. I, II, and III

3. A firm is profitable if
   a. $TR < TC$.
   b. $AR < ATC$.
   c. $MC < ATC$.
   d. $ATC < P$.
   e. $ATC > MC$.

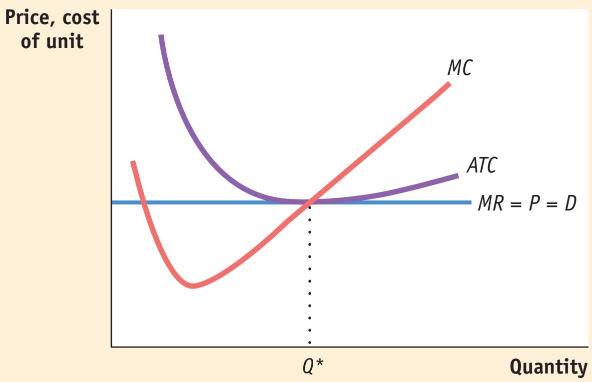
4. If a firm has a total cost of $200, its profit-maximizing level of output is 10 units, and it is breaking even (that is, earning a normal profit), what is the market price?
   a. $200
   b. $100
   c. $20
   d. $10
   e. $2

5. What is the firm's profit if the price of its product is $5 and it produces 500 units of output at a total cost of $1,000?
   a. $5,000
   b. $2,500
   c. $1,500
   d. −$1,500
   e. −$2,500

Tackle the Test: Free-Response Questions

1. Draw a correctly labeled graph showing a profit-maximizing perfectly competitive firm producing at its minimum average total cost.

Answer (6 points)



1 point: Vertical axis and horizontal axis labels are correct ("Price, cost of unit" on vertical axis; "Quantity" on horizontal axis).

1 point: The line representing demand, marginal revenue, and price is horizontal and correctly labeled.

1 point: Marginal cost is "swoosh" shaped or upward sloping and correctly labeled.

1 point: Average total cost is U-shaped and correctly labeled.

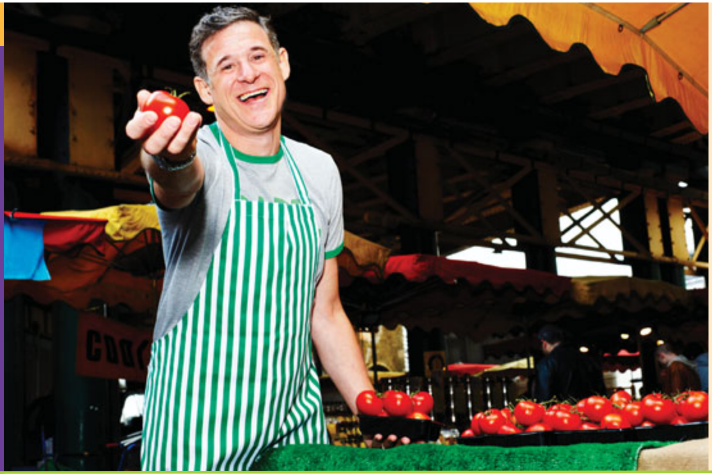
1 point: Quantity is found where $MC = MR$.

1 point: Average total cost reaches its minimum point at the profit-maximizing level of output.

2. Refer to the table provided. Price is equal to $14.
   a. Calculate the firm's marginal cost at each quantity.
   b. Determine the firm's profit-maximizing level of output.
   c. Calculate the firm's profit at the profit-maximizing level of output.

Short-Run Costs for Jennifer and Jason's Farm

Quantity of tomatoes Q (bushels)	*Variable cost VC*	*Total cost TC*
0	$0	$14
1	16	30
2	22	36
3	30	44
4	42	56
5	58	72
6	78	92
7	102	116



What you will learn in this Module:

- How to evaluate a perfectly competitive firm's situation using a graph
- How to determine a perfect competitor's profit or loss
- How a firm decides whether to produce or shut down in the short run

Module 59 Graphing Perfect Competition

We have just learned that for a perfectly competitive firm, a comparison of the market price to the firm's average total cost determines whether the firm is earning a profit, taking a loss, or breaking even with a normal profit of zero. Now we can evaluate the profitability of perfectly competitive firms in a variety of situations.

Interpreting Perfect Competition Graphs

Figure 59.1 illustrates how the market price determines whether a firm is profitable. It also shows how profits are depicted graphically. Each panel shows the marginal cost curve, *MC*, and the short-run average total cost curve, *ATC*. Average total cost is minimized at point *C*. Panel (a) shows the case in which the market price of tomatoes is $18 per bushel. Panel (b) shows the case in which the market price of tomatoes is lower, $10 per bushel.

In panel (a), we see that at a price of $18 per bushel the profit-maximizing quantity of output is 5 bushels, indicated by point *E*, where the marginal cost curve, *MC*, intersects the marginal revenue curve, *MR*—which for a price-taking firm is a horizontal line at the market price. At that quantity of output, average total cost is $14.40 per bushel, indicated by point *Z*. Since the price per bushel exceeds the average total cost per bushel, Jennifer and Jason's farm is profitable.

Jennifer and Jason's total profit when the market price is $18 is represented by the area of the shaded rectangle in panel (a). To see why, notice that total profit can be expressed in terms of profit per unit:

$$\textbf{(59-1)}\quad \text{Profit} = TR - TC = (TR/Q - TC/Q) \times Q$$

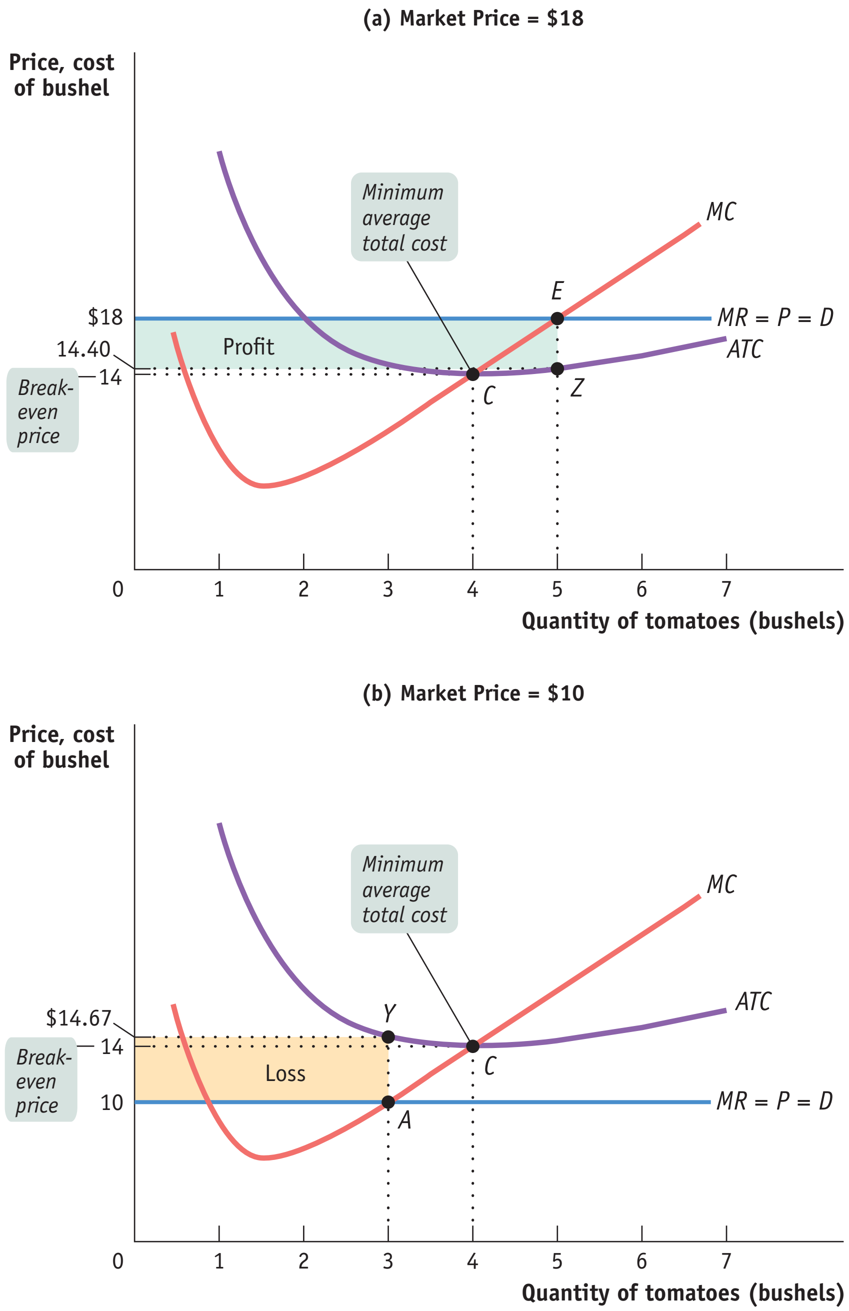
or, equivalently, because P is equal to TR/Q and ATC is equal to TC/Q,

$$\text{Profit} = (P - ATC) \times Q$$

figure 59.1

Profitability and the Market Price

In panel (a) the market price is \$18. The farm is profitable because price exceeds minimum average total cost, the break-even price, \$14. The farm's optimal output choice is indicated by point *E*, corresponding to an output of 5 bushels. The average total cost of producing 5 bushels is indicated by point *Z* on the *ATC* curve, corresponding to an amount of \$14.40. The vertical distance between *E* and *Z* corresponds to the farm's per-unit profit, \$18.00 – \$14.40 = \$3.60. Total profit is given by the area of the shaded rectangle, 5 × \$3.60 = \$18.00. In panel (b) the market price is \$10; the farm is unprofitable because the price falls below the minimum average total cost, \$14. The farm's optimal output choice when producing is indicated by point *A*, corresponding to an output of 3 bushels. The farm's per-unit loss, \$14.67 – \$10.00 = \$4.67, is represented by the vertical distance between *A* and *Y*. The farm's total loss is represented by the shaded rectangle, 3 × \$4.67 = \$14.00 (adjusted for rounding error).



The height of the shaded rectangle in panel (a) corresponds to the vertical distance between points *E* and *Z*. It is equal to *P* – *ATC* = \$18.00 – \$14.40 = \$3.60 per bushel. The shaded rectangle has a width equal to the output: *Q* = 5 bushels. So the area of that rectangle is equal to Jennifer and Jason's profit: 5 bushels × \$3.60 profit per bushel = \$18.

What about the situation illustrated in panel (b)? Here the market price of tomatoes is \$10 per bushel. Producing until price equals marginal cost leads to a profit-maximizing output of 3 bushels, indicated by point *A*. At this output, Jennifer and Jason have an average total cost of \$14.67 per bushel, indicated by point *Y*. At their

The **break-even price** of a price-taking firm is the market price at which it earns zero profits.

profit-maximizing output quantity—3 bushels—average total cost exceeds the market price. This means that Jennifer and Jason's farm generates a loss, not a profit.

How much do they lose by producing when the market price is \$10? On each bushel they lose $ATC - P = \$14.67 - \$10.00 = \$4.67$, an amount corresponding to the vertical distance between points *A* and *Y*. And they produce 3 bushels, which corresponds to the width of the shaded rectangle. So the total value of the losses is $\$4.67 \times 3 = \$14.00$ (adjusted for rounding error), an amount that corresponds to the area of the shaded rectangle in panel (b).

But how does a producer know, in general, whether or not its business will be profitable? It turns out that the crucial test lies in a comparison of the market price to the firm's *minimum average total cost*. On Jennifer and Jason's farm, average total cost reaches its minimum, \$14, at an output of 4 bushels, indicated by point *C*. Whenever the market price exceeds the minimum average total cost, there are output levels for which the average total cost is less than the market price. In other words, the producer can find a level of output at which the firm makes a profit. So Jennifer and Jason's farm will be profitable whenever the market price exceeds \$14. And they will achieve the highest possible profit by producing the quantity at which marginal cost equals price.

Conversely, if the market price is less than the minimum average total cost, there is no output level at which price exceeds average total cost. As a result, the firm will be unprofitable at any quantity of output. As we saw, at a price of \$10—an amount less than the minimum average total cost—Jennifer and Jason did indeed lose money. By producing the quantity at which marginal cost equaled price, Jennifer and Jason did the best they could, but the best they could do was a loss of \$14. Any other quantity would have increased the size of their loss.

The minimum average total cost of a price-taking firm is called its **break-even price,** the price at which it earns zero economic profit (which we now know as a *normal profit*). A firm will earn positive profit when the market price is above the break-even price, and it will suffer losses when the market price is below the break-even price. Jennifer and Jason's break-even price of \$14 is the price at point *C* in Figure 59.1.

So the rule for determining whether a firm is profitable depends on a comparison of the market price of the good to the firm's break-even price—its minimum average total cost:

- Whenever the market price exceeds the minimum average total cost, the producer is profitable.
- Whenever the market price equals the minimum average total cost, the producer breaks even.
- Whenever the market price is less than the minimum average total cost, the producer is unprofitable.

The Short-Run Production Decision

You might be tempted to say that if a firm is unprofitable because the market price is below its minimum average total cost, it shouldn't produce any output. In the short run, however, this conclusion isn't right. In the short run, sometimes the firm should produce even if price falls below minimum average total cost. The reason is that total cost includes *fixed cost*—cost that does not depend on the amount of output produced and can be altered only in the long run. In the short run, fixed cost must still be paid, regardless of whether or not a firm produces. For example, if Jennifer and Jason have rented a tractor for the year, they have to pay the rent on the tractor regardless of whether they produce any tomatoes. *Since it cannot be changed in the short run, their fixed cost is irrelevant to their decision about whether to produce or shut down in the short run.* Although fixed cost should play no role in the decision about whether to produce in the short run, another type of cost—variable cost—does matter. Part of the variable cost for Jennifer and Jason is the wage cost of

iStockphoto

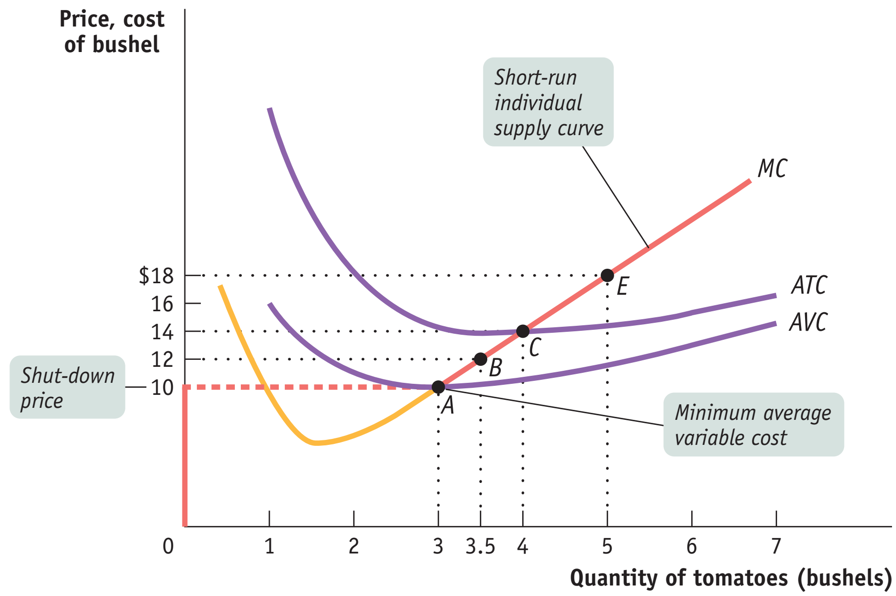
workers who must be hired to help with planting and harvesting. Variable cost can be eliminated by *not* producing, which makes it a critical consideration when determining whether or not to produce in the short run.

Let's turn to Figure 59.2: it shows both the short-run average total cost curve, *ATC,* and the short-run average variable cost curve, *AVC,* drawn from the information in Table 59.1. Recall that the difference between the two curves—the vertical distance between them—represents average fixed cost, the fixed cost per unit of output, *FC/Q.* Because the marginal cost curve has a "swoosh" shape—falling at first before rising—the short-run average variable cost curve is U-shaped: the initial fall in marginal cost causes average variable cost to fall as well, and then the rise in marginal cost eventually pulls average variable cost up again. The short-run average variable cost curve reaches its minimum value of $10 at point *A,* at an output of 3 bushels.

figure 59.2

The Short-Run Individual Supply Curve

When the market price equals or exceeds Jennifer and Jason's *shut-down price* of $10, the minimum average variable cost indicated by point *A,* they will produce the output quantity at which marginal cost is equal to price. So at any price equal to or above the minimum average *variable* cost, the short-run individual supply curve is the firm's marginal cost curve; this corresponds to the upward-sloping segment of the individual supply curve. When market price falls below minimum average variable cost, the firm ceases operation in the short run. This corresponds to the vertical segment of the individual supply curve along the vertical axis.



The Shut-Down Price

We are now prepared to analyze the optimal production decision in the short run. We have two cases to consider:

- When the market price is below the minimum average *variable* cost
- When the market price is greater than or equal to the minimum average *variable* cost

When the market price is below the minimum average variable cost, the price the firm receives per unit is not covering its variable cost per unit. A firm in this situation should cease production immediately. Why? Because there is no level of output at which the firm's total revenue covers its variable cost—the cost it can avoid by not operating. In this case the firm maximizes its profit by not producing at all—by, in effect, minimizing its loss. It will still incur a fixed cost in the short run, but it will no longer incur any variable cost. This means that the minimum average variable cost determines the **shut-down price,** the price at which the firm ceases production in the short run.

When price is greater than minimum average variable cost, however, the firm should produce in the short run. In this case, the firm maximizes profit—or minimizes loss—by choosing the output level at which its marginal cost is equal to the market price. For example, if the market price of tomatoes is $18 per bushel, Jennifer and Jason should

A firm will cease production in the short run if the market price falls below the **shut-down price,** which is equal to minimum average variable cost.

The **short-run individual supply curve** shows how an individual firm's profit-maximizing level of output depends on the market price, taking fixed cost as given.

produce at point *E* in Figure 59.2, corresponding to an output of 5 bushels. Note that point *C* in Figure 59.2 corresponds to the farm's break-even price of $14 per bushel. Since *E* lies above *C,* Jennifer and Jason's farm will be profitable; they will generate a per-bushel profit of $18.00 – $14.40 = $3.60 when the market price is $18.

But what if the market price lies between the shut-down price and the break-even price—that is, between the minimum average *variable* cost and the minimum average *total* cost? In the case of Jennifer and Jason's farm, this corresponds to prices anywhere between $10 and $14—say, a market price of $12. At $12, Jennifer and Jason's farm is not profitable; since the market price is below the minimum average total cost, the farm is losing (on average) the difference between price and average total cost on every unit produced. Yet even though the market price isn't covering Jennifer and Jason's average total cost, it is covering their average variable cost and some—but not all—of the average fixed cost. If a firm in this situation shuts down, it will incur no variable cost but it will incur the *full* fixed cost. As a result, shutting down will generate an even greater loss than continuing to operate.

This means that whenever price falls between minimum average total cost and minimum average variable cost, the firm is better off producing some output in the short run. The reason is that by producing, it can cover its variable cost and at least some of its fixed cost, even though it is incurring a loss. In this case, the firm maximizes profit—that is, minimizes loss—by choosing the quantity of output at which its marginal cost is equal to the market price. So if Jennifer and Jason face a market price of $12 per bushel, their profit-maximizing output is given by point *B* in Figure 59.2, corresponding to an output of 3.5 bushels.

It's worth noting that the decision to produce when the firm is covering its variable cost but not all of its fixed cost is similar to the decision to ignore a *sunk cost,* a concept we studied previously. You may recall that a sunk cost is a cost that has already been incurred and cannot be recouped; and because it cannot be changed, it should have no effect on any current decision. In the short-run production decision, fixed cost is, in effect, like a sunk cost—it has been spent, and it can't be recovered in the short run. This comparison also illustrates why variable cost does indeed matter in the short run: it can be avoided by not producing.

And what happens if the market price is exactly equal to the shut-down price, the minimum average variable cost? In this instance, the firm is indifferent between producing 3 units or 0 units. As we'll see shortly, this is an important point when looking at the behavior of an industry as a whole. For the sake of clarity, we'll assume that the firm, although indifferent, does indeed produce output when price is equal to the shut-down price.

Putting everything together, we can now draw the **short-run individual supply curve** of Jennifer and Jason's farm, the red line in Figure 59.2; it shows how the profit-maximizing quantity of output in the short run depends on the price. As you can see, the curve is in two segments. The upward-sloping red segment starting at point *A* shows the short-run profit-maximizing output when market price is equal to or above the shut-down price of $10 per bushel. As long as the market price is equal to or above the shut-down price, Jennifer and Jason will produce the quantity of output at which marginal cost is equal to the market price. So at market prices equal to or above the shut-down price, the firm's short-run supply curve corresponds to its marginal cost curve. But at any market price below the minimum average variable cost, in this case, $10 per bushel—the firm shuts down and output drops to zero in the short run. This corresponds to the vertical segment of the curve that lies on top of the vertical axis.


Muntz/Stone/Getty Images

Do firms sometimes shut down temporarily without going out of business? Yes. In fact, in some industries temporary shutdowns are routine. The most common examples are industries in which demand is highly seasonal, like outdoor amusement parks in climates with cold winters. Such parks would have to offer very low prices to entice customers during the colder months—prices so low that the owners would not cover their variable cost (principally wages and electricity). The wiser choice economically is to shut down until warm weather brings enough customers who are willing to pay a higher price.

Changing Fixed Cost

Although fixed cost cannot be altered in the short run, in the long run firms can acquire or get rid of machines, buildings, and so on. In the long run the level of fixed cost is a matter of choice, and a firm will choose the level of fixed cost that minimizes the average total cost for its desired output level. Now we will focus on an even bigger question facing a firm when choosing its fixed cost: whether to incur *any* fixed cost at all by continuing to operate.

In the long run, a firm can always eliminate fixed cost by selling off its plant and equipment. If it does so, of course, it can't produce any output—it has exited the industry. In contrast, a new firm can take on some fixed cost by acquiring machines and other resources, which puts it in a position to produce—it can enter the industry. In most perfectly competitive industries the set of firms, although fixed in the short run, changes in the long run as some firms enter or exit the industry.

Consider Jennifer and Jason's farm once again. In order to simplify our analysis, we will sidestep the issue of choosing among several possible levels of fixed cost. Instead, we will assume that if they operate at all, Jennifer and Jason have only one possible choice of fixed cost: $14. Alternatively, they can choose a fixed cost of zero if they exit the industry. It is changes in fixed cost that cause short-run average total cost curves to differ from long-run total cost curves, so with this assumption, Jennifer and Jason's short-run and long-run average total cost curves are one and the same.

Suppose that the market price of organic tomatoes is consistently less than the break-even price of $14 over an extended period of time. In that case, Jennifer and Jason never fully cover their total cost: their business runs at a persistent loss. In the long run, then, they can do better by closing their business and leaving the industry. In other words, *in the long run* firms will exit an industry if the market price is consistently less than their break-even price—their minimum average total cost.

Conversely, suppose that the price of organic tomatoes is consistently above the break-even price, $14, for an extended period of time. Because their farm is profitable, Jennifer and Jason will remain in the industry and continue producing. But things won't stop there. The organic tomato industry meets the criterion of *free entry*: there are many potential organic tomato producers because the necessary inputs are easy to obtain. And the cost curves of those potential producers are likely to be similar to those of Jennifer and Jason, since the technology used by other producers is likely to be very similar to that used by Jennifer and Jason. If the price is high enough to generate profits for existing producers, it will also attract some of these potential producers into the industry. So *in the long run* a price in excess of $14 should lead to entry: new producers will come into the organic tomato industry.

As we will see shortly, exit and entry lead to an important distinction between the *short-run industry supply curve* and the *long-run industry supply curve.*

Summing Up: The Perfectly Competitive Firm's Profitability and Production Conditions

In this module we've studied what's behind the supply curve for a perfectly competitive, price-taking firm. A perfectly competitive firm maximizes profit, or minimizes loss, by producing the quantity that equates price and marginal cost. The exception is if price is below minimum average variable cost in the short run, or below minimum average total cost in the long run, in which case the firm is better off shutting down.

table **59.1**

Summary of the Perfectly Competitive Firm's Profitability and Production Conditions

Profitability condition (minimum *ATC* = break-even price)	Result
$P >$ minimum *ATC*	Firm profitable. Entry into industry in the long run.
$P =$ minimum *ATC*	Firm breaks even. No entry into or exit from industry in the long run.
$P <$ minimum *ATC*	Firm unprofitable. Exit from industry in the long run.
Production condition (minimum *AVC* = shut-down price)	**Result**
$P >$ minimum *AVC*	Firm produces in the short run. If $P <$ minimum *ATC,* firm covers variable cost and some but not all of fixed cost. If $P >$ minimum *ATC,* firm covers all variable cost and fixed cost.
$P =$ minimum *AVC*	Firm indifferent between producing in the short run or not. Just covers variable cost.
$P <$ minimum *AVC*	Firm shuts down in the short run. Does not cover variable cost.

Table 59.1 summarizes the perfectly competitive firm's profitability and production conditions. It also relates them to entry into and exit from the industry in the long run. Now that we understand how a perfectly competitive *firm* makes its decisions, we can go on to look at the supply curve for a perfectly competitive *market* and the long-run equilibrium in perfect competition.

fyi

Prices Are Up . . . but So Are Costs

In 2005 Congress passed the Energy Policy Act, mandating that by the year 2012, 7.5 billion gallons of alternative fuel—mostly corn-based ethanol—be added to the American fuel supply with the goal of reducing gasoline consumption. The unsurprising result of this mandate: the demand for corn skyrocketed, along with its price. In spring 2007, the price of corn was 50% higher than it had been a year earlier.

This development caught the eye of American farmers like Ronnie Gerik of Aquilla, Texas, who, in response to surging corn prices, reduced the size of his cotton crop and increased his corn acreage by 40%. He was not alone; within a year, the amount of U.S. acreage planted in corn increased by 15%.

Although this sounds like a sure way to make a profit, Gerik was actually taking a big gamble: even though the price of corn increased, so did the cost of the raw materials needed to grow it—by 20%. Consider the cost of just two inputs: fertilizer and fuel. Corn requires more fertilizer than other crops and, with more farmers planting corn, the increased demand for fertilizer led to a price increase. Corn also has to be transported farther away from the farm than cotton; at the same time that Gerik began shifting to greater corn production, diesel fuel became very expensive. Moreover, corn is much more sensitive to the amount of rainfall than a crop like cotton. So farmers who plant corn in drought-prone places like Texas are increasing their risk of loss. Gerik had to incorporate into his calculations his best guess of what a dry spell would cost him.

Despite all of this, what Gerik did made complete economic sense. By planting more corn, he was moving up his individual short-run supply curve for corn production. And because his individual supply curve is his marginal cost curve, his costs also went up because he had to use more inputs—inputs that had become more expensive to obtain.

Courtesy of Ronnie Gerik.

Although Gerik was taking a big gamble when he cut the size of his cotton crop to plant more corn, his decision made good economic sense.

So the moral of this story is that farmers will increase their corn acreage until the marginal cost of producing corn is approximately equal to the market price of corn—which shouldn't come as a surprise because corn production satisfies all the requirements of a perfectly competitive industry.

Module 59 AP Review

Solutions appear at the back of the book.

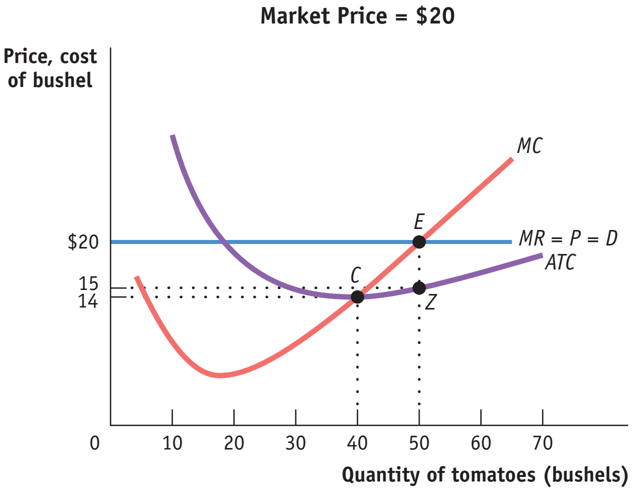
Check Your Understanding

1. Draw a short-run diagram showing a U-shaped average total cost curve, a U-shaped average variable cost curve, and a "swoosh"-shaped marginal cost curve. On it, indicate the range of prices for which the following actions are optimal.
   a. The firm shuts down immediately.
   b. The firm operates in the short run despite sustaining a loss.
   c. The firm operates while making a profit.

2. The state of Maine has a very active lobster industry, which harvests lobsters during the summer months. During the rest of the year, lobsters can be obtained by restaurants from producers in other parts of the world, but at a much higher price. Maine is also full of "lobster shacks," roadside restaurants serving lobster dishes that are open only during the summer. Supposing that the market demand for lobster dishes remains the same throughout the year, explain why it is optimal for lobster shacks to operate only during the summer.

Tackle the Test: Multiple-Choice Questions

For questions 1–3, refer to the graph provided.



1. The firm's total revenue is equal to
   a. $14.
   b. $20.
   c. $560.
   d. $750.
   e. $1,000.

2. The firm's total cost is equal to
   a. $14.
   b. $15.
   c. $560.
   d. $750.
   e. $1,000.

3. The firm is earning a
   a. profit equal to $5.
   b. profit equal to $250.
   c. loss equal to $15.
   d. loss equal to $750.
   e. loss equal to $250.

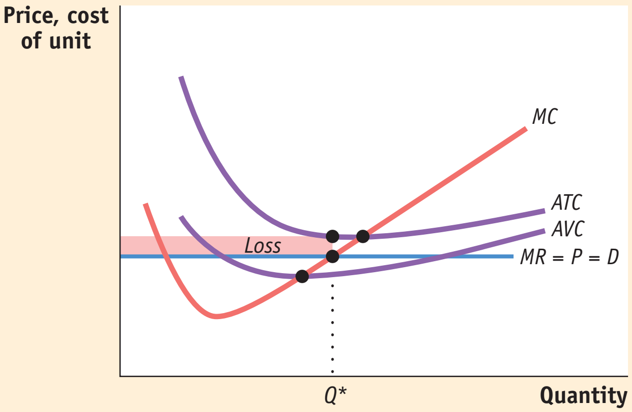
4. A firm should continue to produce in the short run as long as price is at least equal to
   a. *MR*.
   b. *MC*.
   c. minimum *ATC*.
   d. minimum *AVC*.
   e. *AFC*.

5. At prices that motivate the firm to produce at all, the short-run supply curve for a perfect competitor corresponds to which curve?
   a. the *ATC* curve
   b. the *AVC* curve
   c. the *MC* curve
   d. the *AFC* curve
   e. the *MR* curve

Tackle the Test: Free-Response Questions

1. Draw a correctly labeled graph showing a perfectly competitive firm producing and incurring a loss in the short run.

Answer (10 points)



1 point: Vertical axis is labeled "Price, cost of unit" or "Dollars per unit"; horizontal axis labeled "Quantity" or "Q."

1 point: Demand curve is horizontal and labeled with some combination of "*P*," "*MR*," or "*D*."

1 point: *MC* is labeled and slopes upward in the shape of a swoosh.

1 point: Profit-maximizing quantity is labeled (for example, as "Q^*") on the horizontal axis where *MC* = *MR*.

1 point: *ATC* is labeled and U-shaped.

1 point: *ATC* is above price at the profit-maximizing output.

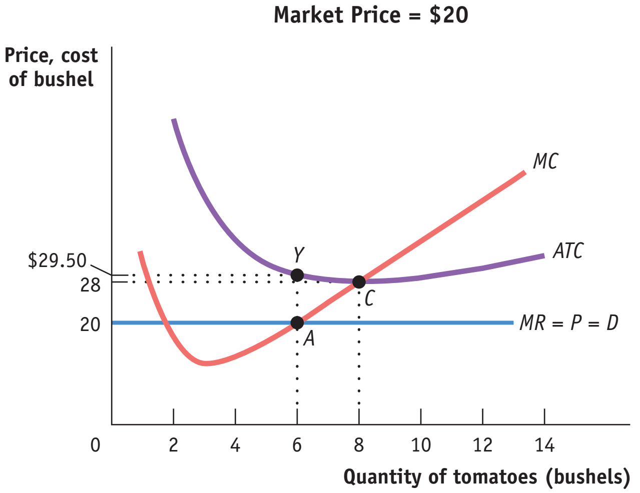
1 point: *MC* crosses *ATC* at the lowest point on *ATC*.

1 point: *AVC* is labeled and U-shaped.

1 point: *AVC* is below price at the profit-maximizing output.

1 point: Loss rectangle is correctly located and identified.

2. Refer to the graph provided.



a. Assuming it is appropriate for the firm to produce in the short run, what is the firm's profit-maximizing level of output?

b. Calculate the firm's total revenue.

c. Calculate the firm's total cost.

d. Calculate the firm's profit or loss.

e. If *AVC* were $22 at the profit-maximizing level of output, would the firm produce in the short run? Explain why or why not.

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

- Why industry behavior differs between the short run and the long run
- What determines the industry supply curve in both the short run and the long run

Module 60 Long-Run Outcomes in Perfect Competition

Up to this point we have been discussing the perfectly competitive firm's short-run situation—whether to produce or not, and if so, whether the firm earns a positive profit, breaks even with a normal profit, or takes a loss. In this module, we look at the long-run situation in a perfectly competitive market. We will see that perfect competition leads to some interesting and desirable market outcomes. Later, we will contrast these outcomes with the outcomes in monopolistic and imperfectly competitive markets.

The Industry Supply Curve

Why will an increase in the demand for organic tomatoes lead to a large price increase at first but a much smaller increase in the long run? The answer lies in the behavior of the **industry supply curve**—the relationship between the price and the total output of an industry as a whole. The industry supply curve is what we referred to in earlier modules as the supply curve or the market supply curve. But here we take some extra care to distinguish between the *individual supply curve* of a single firm and the supply curve of the industry as a whole.

As you might guess from the previous module, the industry supply curve must be analyzed in somewhat different ways for the short run and the long run. Let's start with the short run.

The Short-Run Industry Supply Curve

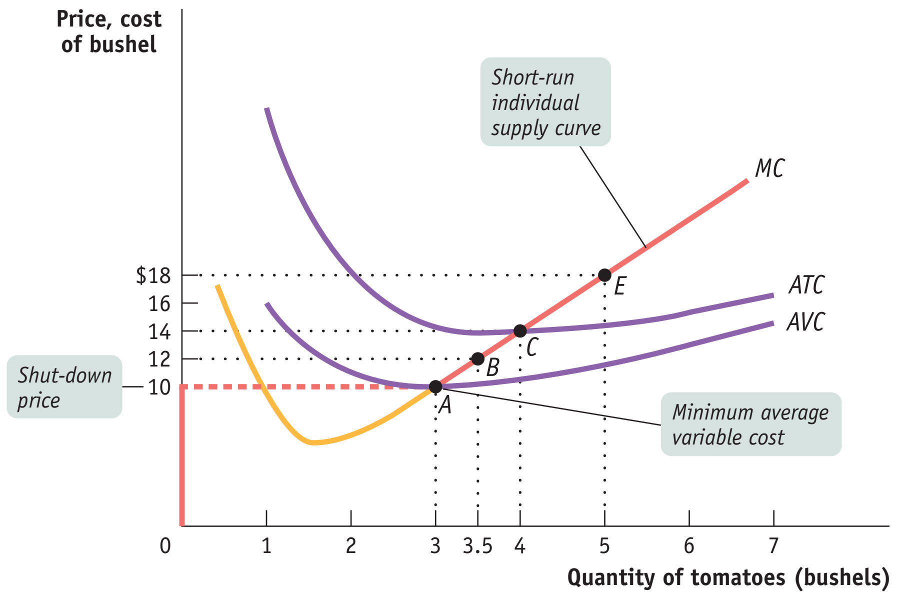
Recall that in the short run the number of firms in an industry is fixed—there is no entry or exit. And you may also remember that the industry supply curve is the horizontal sum of the individual supply curves of all firms—you find it by summing the total output across all suppliers at every given price. We will do that exercise here under the assumption that all the firms are alike—an assumption that makes the derivation particularly simple. So let's assume that there are 100 organic tomato farms, each with the same costs as Jennifer and Jason's farm. Each of these 100 farms will have an individual short-run supply curve like the one in Figure 59.2 from the previous module, which is reprinted on the next page for your convenience.

The **industry supply curve** shows the relationship between the price of a good and the total output of the industry as a whole.

figure 59.2

The Short-Run Individual Supply Curve

When the market price equals or exceeds Jennifer and Jason's *shut-down price* of $10, the minimum average variable cost indicated by point *A*, they will produce the output quantity at which marginal cost is equal to price. So at any price equal to or above the minimum average *variable* cost, the short-run individual supply curve is the firm's marginal cost curve; this corresponds to the upward-sloping segment of the individual supply curve. When market price falls below minimum average variable cost, the firm ceases operation in the short run. This corresponds to the vertical segment of the individual supply curve along the vertical axis.



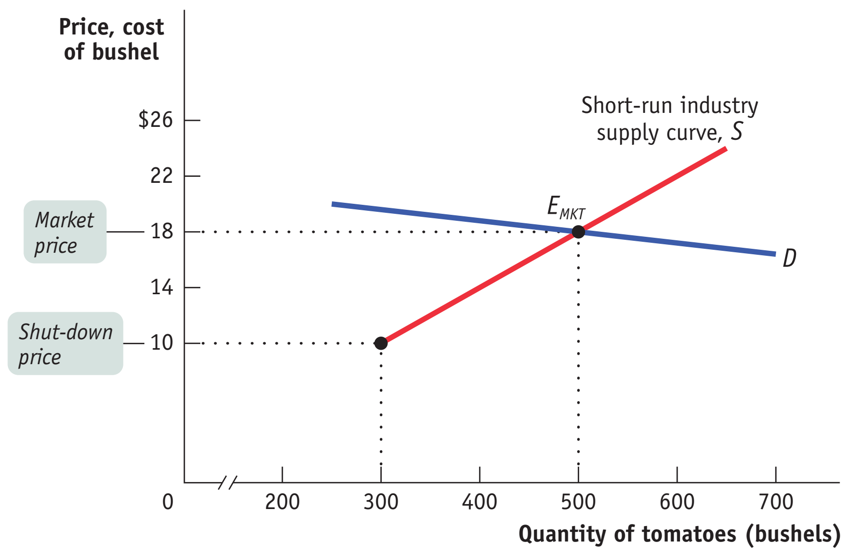
The **short-run industry supply curve** shows how the quantity supplied by an industry depends on the market price, given a fixed number of firms.

At a price below $10, no farms will produce. At a price of more than $10, each farm will produce the quantity of output at which its marginal cost is equal to the market price. As you can see from Figure 59.2, this will lead each farm to produce 4 bushels if the price is $14 per bushel, 5 bushels if the price is $18, and so on. So if there are 100 organic tomato farms and the price of organic tomatoes is $18 per bushel, the industry as a whole will produce 500 bushels, corresponding to 100 farms × 5 bushels per farm. The result is the **short-run industry supply curve,** shown as *S* in Figure 60.1. This curve shows the quantity that producers will supply at each price, *taking the number of farms as given.*

figure 60.1

The Short-Run Market Equilibrium

The short-run industry supply curve, *S,* is the industry supply curve taking the number of producers—here, 100—as given. It is generated by adding together the individual supply curves of the 100 producers. Below the shut-down price of $10, no producer wants to produce in the short run. Above $10, the short-run industry supply curve slopes upward, as each producer increases output as price increases. It intersects the demand curve, *D,* at point E_{MKT}, the point of short-run market equilibrium, corresponding to a market price of $18 and a quantity of 500 bushels.



The market demand curve, labeled D in Figure 60.1, crosses the short-run industry supply curve at E_{MKT}, corresponding to a price of \$18 and a quantity of 500 bushels. Point E_{MKT} is a **short-run market equilibrium:** the quantity supplied equals the quantity demanded, taking the number of farms as given. But the long run may look quite different because in the long run farms may enter or exit the industry.

There is a **short-run market equilibrium** when the quantity supplied equals the quantity demanded, taking the number of producers as given.

The Long-Run Industry Supply Curve

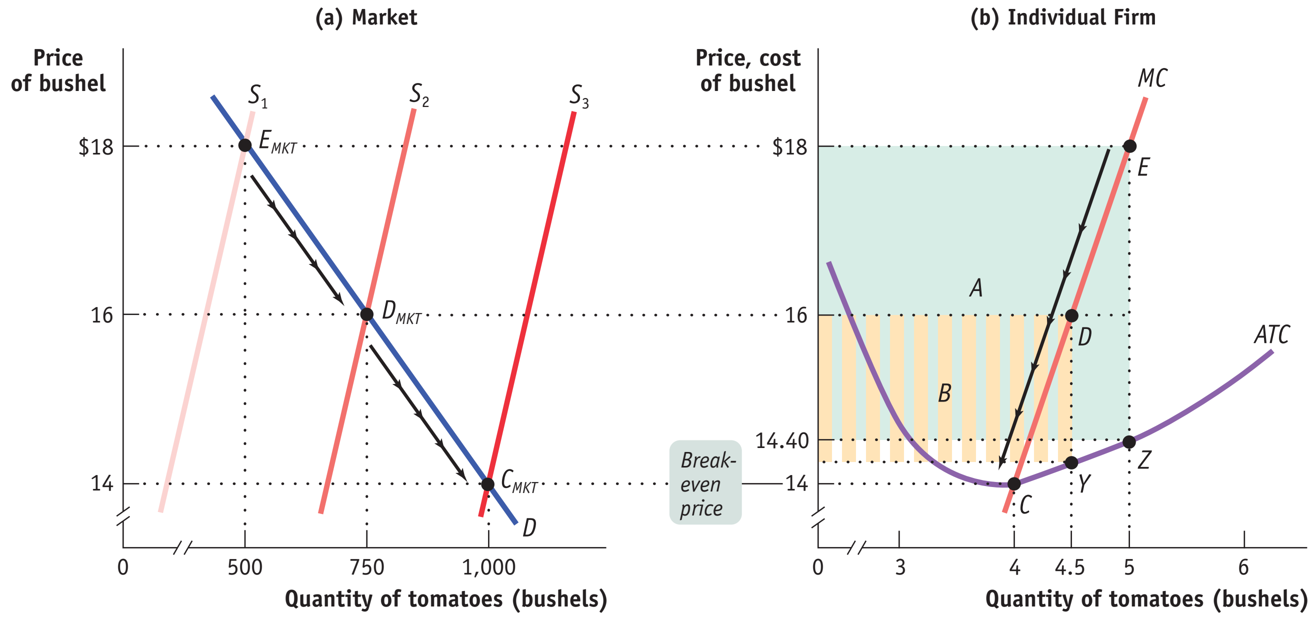
Suppose that in addition to the 100 farms currently in the organic tomato business, there are many other potential organic tomato farms. Suppose also that each of these potential farms would have the same cost curves as existing farms, like the one owned by Jennifer and Jason, upon entering the industry.

When will additional farms enter the industry? Whenever existing farms are making a profit—that is, whenever the market price is above the break-even price of \$14 per bushel, the minimum average total cost of production. For example, at a price of \$18 per bushel, new farms will enter the industry.

What will happen as additional farms enter the industry? Clearly, the quantity supplied at any given price will increase. The short-run industry supply curve will shift to the right. This will, in turn, alter the market equilibrium and result in a lower market price. Existing farms will respond to the lower market price by reducing their output, but the total industry output will increase because of the larger number of farms in the industry.

Figure 60.2 illustrates the effects of this chain of events on an existing farm and on the market; panel (a) shows how the market responds to entry, and panel (b) shows

figure 60.2 The Long-Run Market Equilibrium



Point E_{MKT} of panel (a) shows the initial short-run market equilibrium. Each of the 100 existing producers makes an economic profit, illustrated in panel (b) by the green rectangle labeled A, the profit of an existing firm. Profits induce entry by additional producers, shifting the short-run industry supply curve outward from S_1 to S_2 in panel (a), resulting in a new short-run equilibrium at point D_{MKT}, at a lower market price of \$16 and higher industry output. Existing firms reduce output and profit falls to the area given by the striped rectangle labeled B in panel (b). Entry continues to shift out the short-run industry supply curve, as price falls and industry output increases yet again. Entry ceases at point C_{MKT} on supply curve S_3 in panel (a). Here market price is equal to the break-even price; existing producers make zero economic profits and there is no incentive for entry or exit. Therefore C_{MKT} is also a long-run market equilibrium.

A market is in **long-run market equilibrium** when the quantity supplied equals the quantity demanded, given that sufficient time has elapsed for entry into and exit from the industry to occur.

how an individual existing farm responds to entry. (Note that these two graphs have been rescaled in comparison to Figures 59.2 and 60.1 to better illustrate how profit changes in response to price.) In panel (a), S_1 is the initial short-run industry supply curve, based on the existence of 100 producers. The initial short-run market equilibrium is at E_{MKT}, with an equilibrium market price of \$18 and a quantity of 500 bushels. At this price existing farms are profitable, which is reflected in panel (b): an existing farm makes a total profit represented by the green shaded rectangle labeled A when the market price is \$18.

These profits will induce new producers to enter the industry, shifting the short-run industry supply curve to the right. For example, the short-run industry supply curve when the number of farms has increased to 167 is S_2. Corresponding to this supply curve is a new short-run market equilibrium labeled D_{MKT}, with a market price of \$16 and a quantity of 750 bushels. At \$16, each farm produces 4.5 bushels, so that industry output is $167 \times 4.5 = 750$ bushels (rounded). From panel (b) you can see the effect of the entry of 67 new farms on an existing farm: the fall in price causes it to reduce its output, and its profit falls to the area represented by the striped rectangle labeled *B*.

Although diminished, the profit of existing farms at D_{MKT} means that entry will continue and the number of farms will continue to rise. If the number of farms rises to 250, the short-run industry supply curve shifts out again to S_3, and the market equilibrium is at C_{MKT}, with a quantity supplied and demanded of 1,000 bushels and a market price of \$14 per bushel.

Like E_{MKT} and D_{MKT}, C_{MKT} is a short-run equilibrium. But it is also something more. Because the price of \$14 is each farm's break-even price, an existing producer makes zero economic profit—neither a profit nor a loss, earning only the opportunity cost of the resources used in production—when producing its profit-maximizing output of 4 bushels. At this price there is no incentive either for potential producers to enter or for existing producers to exit the industry. So C_{MKT} corresponds to a **long-run market equilibrium**—a situation in which the quantity supplied equals the quantity demanded, given that sufficient time has elapsed for producers to either enter or exit the industry. In a long-run market equilibrium, all existing and potential producers have fully adjusted to their optimal long-run choices; as a result, no producer has an incentive to either enter or exit the industry.

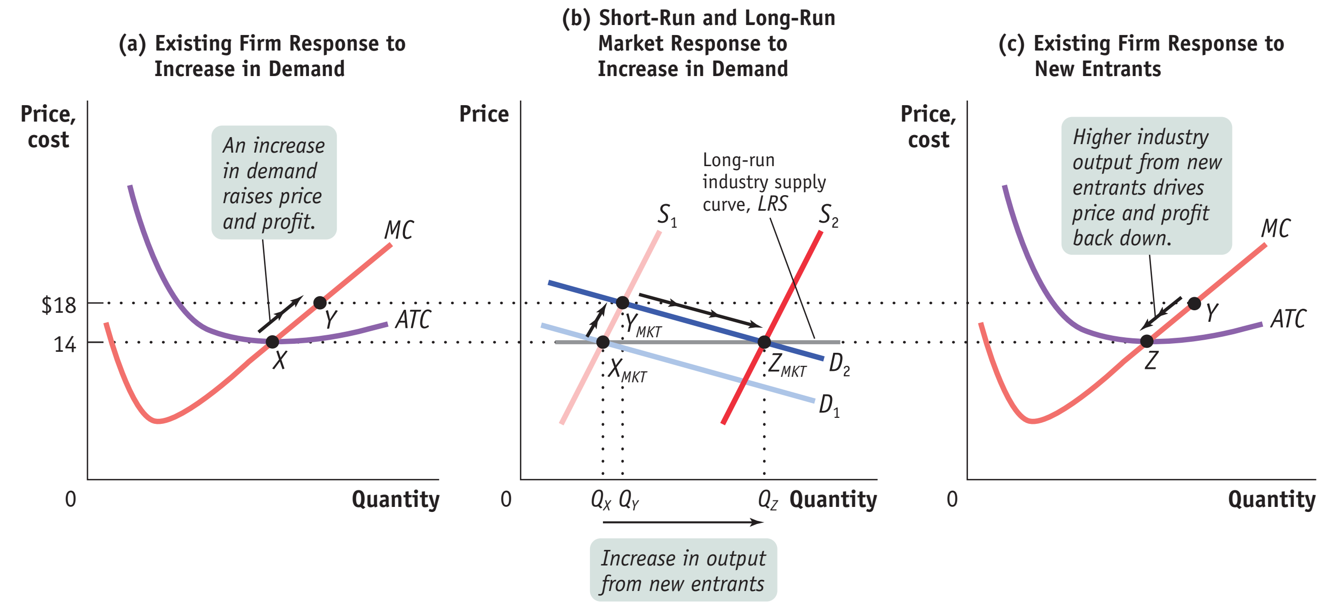
To explore further the difference between short-run and long-run equilibrium, consider the effect of an increase in demand on an industry with free entry that is initially in long-run equilibrium. Panel (b) in Figure 60.3 shows the market adjustment; panels (a) and (c) show how an existing individual firm behaves during the process.

In panel (b) of Figure 60.3, D_1 is the initial demand curve and S_1 is the initial short-run industry supply curve. Their intersection at point X_{MKT} is both a short-run and a long-run market equilibrium because the equilibrium price of \$14 leads to zero economic profit—and therefore neither entry nor exit. It corresponds to point *X* in panel (a), where an individual existing firm is operating at the minimum of its average total cost curve.

Now suppose that the demand curve shifts out for some reason to D_2. As shown in panel (b), in the short run, industry output moves along the short-run industry supply curve, S_1, to the new short-run market equilibrium at Y_{MKT}, the intersection of S_1 and D_2. The market price rises to \$18 per bushel, and industry output increases from Q_X to Q_Y. This corresponds to an existing firm's movement from *X* to *Y* in panel (a) as the firm increases its output in response to the rise in the market price.

But we know that Y_{MKT} is not a long-run equilibrium because \$18 is higher than minimum average total cost, so existing firms are making economic profits. This will lead additional firms to enter the industry. Over time entry will cause the short-run industry supply curve to shift to the right. In the long run, the short-run industry supply curve will have shifted out to S_2, and the equilibrium will be at Z_{MKT}—with the price

figure 60.3 The Effect of an Increase in Demand in the Short Run and the Long Run



Panel (b) shows how an industry adjusts in the short and long run to an increase in demand; panels (a) and (c) show the corresponding adjustments by an existing firm. Initially the market is at point X_{MKT} in panel (b), a short-run and long-run equilibrium at a price of $14 and industry output of Q_X. An existing firm makes zero economic profit, operating at point *X* in panel (a) at minimum average total cost. Demand increases as D_1 shifts rightward to D_2, in panel (b), raising the market price to $18. Existing firms increase their output, and industry output moves along the short-run industry supply curve S_1 to a short-run equilibrium at Y_{MKT}. Correspondingly, the existing firm in panel (a) moves from point *X* to point *Y*. But at a price of $18 existing firms are profitable. As shown in panel (b), in the long run new entrants arrive and the short-run industry supply curve shifts rightward, from S_1 to S_2. There is a new equilibrium at point Z_{MKT}, at a lower price of $14 and higher industry output of Q_Z. An existing firm responds by moving from *Y* to *Z* in panel (c), returning to its initial output level and zero economic profit. Production by new entrants accounts for the total increase in industry output, $Q_Z - Q_X$. Like X_{MKT}, Z_{MKT} is also a short-run and long-run equilibrium: with existing firms earning zero economic profit, there is no incentive for any firms to enter or exit the industry. The horizontal line passing through X_{MKT} and Z_{MKT}, *LRS*, is the *long-run industry supply curve:* at the break-even price of $14, producers will produce any amount that consumers demand in the long run.

falling back to $14 per bushel and industry output increasing yet again, from Q_Y to Q_Z. Like X_{MKT} before the increase in demand, Z_{MKT} is both a short-run and a long-run market equilibrium.

The effect of entry on an existing firm is illustrated in panel (c), in the movement from *Y* to *Z* along the firm's individual supply curve. The firm reduces its output in response to the fall in the market price, ultimately arriving back at its original output quantity, corresponding to the minimum of its average total cost curve. In fact, every firm that is now in the industry—the initial set of firms and the new entrants—will operate at the minimum of its average total cost curve, at point *Z*. This means that the entire increase in industry output, from Q_X to Q_Z, comes from production by new entrants.

The line *LRS* that passes through X_{MKT} and Z_{MKT} in panel (b) is the **long-run industry supply curve.** It shows how the quantity supplied by an industry responds to the price, given that firms have had time to enter or exit the industry.

The **long-run industry supply curve** shows how the quantity supplied responds to the price once producers have had time to enter or exit the industry.

In this particular case, the long-run industry supply curve is horizontal at $14. In other words, in this industry supply is *perfectly elastic* in the long run: given time to enter or exit, firms will supply any quantity that consumers demand at a price of $14.

Perfectly elastic long-run supply is actually a good assumption for many industries. In this case we speak of there being *constant costs across the industry*: each firm, regardless of whether it is an incumbent or a new entrant, faces the same cost structure (that is, they each have the same cost curve). Industries that satisfy this condition are industries in which there is a perfectly elastic supply of inputs—industries like agriculture or bakeries. In other industries, however, even the long-run industry supply curve slopes upward. The usual reason for this is that producers must use some input that is in limited supply (that is, their supply is at least somewhat inelastic). As the industry expands, the price of that input is driven up. Consequently, the cost structure for firms becomes higher than it was when the industry was smaller. An example is beach front resort hotels, which must compete for a limited quantity of prime beachfront property. Industries that behave like this are said to have *increasing costs across the industry*. Finally, it is possible for the long-run industry supply curve to slope downward, a condition that occurs when the cost structure for firms becomes lower as the industry expands. This is the case in industries such as the electric car industry, in which increased output allows for economies of scale in the production of lithium batteries and other specialized inputs, and thus lower input prices. A downward-sloping industry supply curve indicates *decreasing costs across the industry*.

iStockphoto

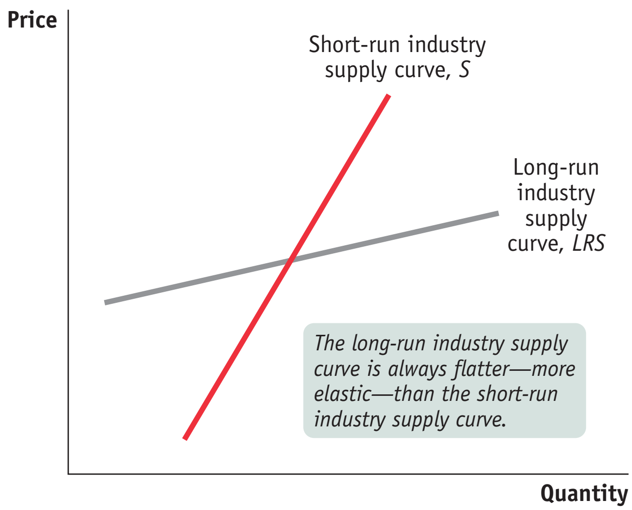
Regardless of whether the long-run industry supply curve is horizontal, upward sloping, or downward sloping, the long-run price elasticity of supply is *higher* than the short-run price elasticity whenever there is free entry and exit. As shown in Figure 60.4, the long-run industry supply curve is always flatter than the short-run industry supply curve. The reason is entry and exit: a high price caused by an increase in demand attracts entry by new firms, resulting in a rise in industry output and an eventual fall in price; a low price caused by a decrease in demand induces existing firms to exit, leading to a fall in industry output and an eventual increase in price.

The distinction between the short-run industry supply curve and the long-run industry supply curve is very important in practice. We often see a sequence of events like that shown in Figure 60.3: an increase in demand initially leads to a large price

figure 60.4

Comparing the Short-Run and Long-Run Industry Supply Curves

The long-run industry supply curve may slope upward, but it is always flatter—more elastic—than the short-run industry supply curve. This is because of entry and exit: a higher price attracts new entrants in the long run, resulting in a rise in industry output and a fall in price; a lower price induces existing producers to exit in the long run, generating a fall in industry output and a rise in price.



increase, but prices return to their initial level once new firms have entered the industry. Or we see the sequence in reverse: a fall in demand reduces prices in the short run, but they return to their initial level as producers exit the industry.

The Cost of Production and Efficiency in Long-Run Equilibrium

Our analysis leads us to three conclusions about the cost of production and efficiency in the long-run equilibrium of a perfectly competitive industry. These results will be important in our upcoming discussion of how monopoly gives rise to inefficiency.

First, in a perfectly competitive industry in equilibrium, the value of marginal cost is the same for all firms. That's because all firms produce the quantity of output at which marginal cost equals the market price, and as price-takers they all face the same market price.

Second, in a perfectly competitive industry with free entry and exit, each firm will have zero economic profit in the long-run equilibrium. Each firm produces the quantity of output that minimizes its average total cost—corresponding to point *Z* in panel (c) of Figure 60.3. So the total cost of producing the industry's output is minimized in a perfectly competitive industry.

The third and final conclusion is that the long-run market equilibrium of a perfectly competitive industry is efficient: no mutually beneficial transactions go unexploited. To understand this, recall a fundamental requirement for efficiency: all consumers who are willing to pay an amount greater than or equal to the sellers' cost actually get the good. We also learned that when a market is efficient (except under certain, well-defined conditions), the market price matches all consumers willing to pay at least the market price with all sellers who have a cost of production that is less than or equal to the market price.

So in the long-run equilibrium of a perfectly competitive industry, production is efficient: costs are minimized and no resources are wasted. In addition, the allocation of goods to consumers is efficient: every consumer willing to pay the cost of producing the good gets it. Indeed, no mutually beneficial transaction is left unexploited. Moreover, this condition tends to persist over time as the environment changes: the force of competition makes producers responsive to changes in consumers' desires and to changes in technology.

fyi

A Crushing Reversal

For some reason, starting in the mid-1990s, Americans began drinking a lot more wine. Part of this increase in demand may have reflected a booming economy, but the surge in wine consumption continued even after the economy stumbled in 2001. By 2006, Americans were consuming 59% more wine than they did in 1993—a total of 2.4 gallons of wine per year per U.S. resident.

At first, the increase in wine demand led to sharply higher prices; between 1993 and 2000, the price of red wine grapes rose approximately 50%, and California grape growers earned high profits. Then, as the discussions of long-run supply foretell, there was a rapid expansion of the industry, both because existing grape growers expanded their capacity and because new growers entered the industry. Between 1994 and 2002, production of red wine grapes almost doubled. The result was predictable: the price of grapes fell as the supply curve shifted out. As demand growth slowed in 2002, prices plunged by 17%. The effect was to end the California wine industry's expansion. In fact, some grape producers began to exit the industry. By 2004, U.S. grape production had fallen by 20% compared to 2002.


John Lee/Aurora Photos

Module 60 AP Review

Solutions appear at the back of the book.

Check Your Understanding

1. Which of the following events will induce firms to enter an industry? Which will induce firms to exit? When will entry or exit cease? Explain your answer.
   a. A technological advance lowers the fixed cost of production of every firm in the industry.
   b. The wages paid to workers in the industry go up for an extended period of time.
   c. A permanent change in consumer tastes increases demand for the good.
   d. The price of a key input rises due to a long-term shortage of that input.

2. Assume that the egg industry is perfectly competitive and is in long-run equilibrium with a perfectly elastic long-run industry supply curve. Health concerns about cholesterol then lead to a decrease in demand. Construct a figure similar to Figure 60.3, showing the short-run behavior of the industry and how long-run equilibrium is reestablished.

Tackle the Test: Multiple-Choice Questions

1. In the long run, a perfectly competitive firm will earn
   a. a negative market return.
   b. a positive profit.
   c. a loss.
   d. a normal profit.
   e. excess profit.

2. With perfect competition, efficiency is generally attained in
   a. the short run but not the long run.
   b. the long run but not the short run.
   c. both the short run and the long run.
   d. neither the short run nor the long run.
   e. specific firms only.

3. Compared to the short-run industry supply curve, the long-run industry supply curve will be more
   a. elastic.
   b. inelastic.
   c. steeply sloped.
   d. profitable.
   e. accurate.

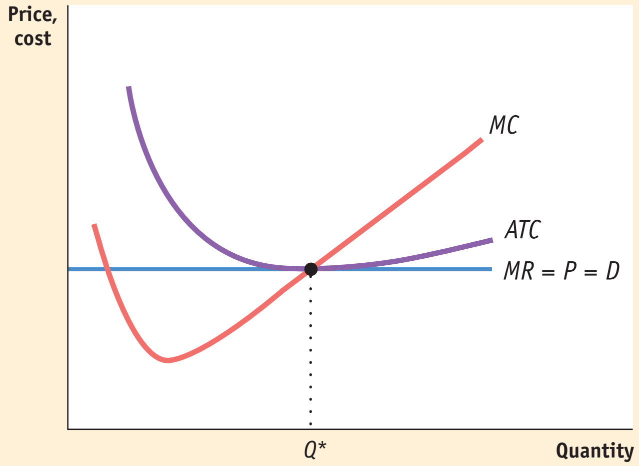
4. Which of the following is generally true for perfect competition?
   I. There is free entry and exit.
   II. Long-run market equilibrium is efficient.
   III. Firms maximize profits at the output level where $P = MC$.
   a. I only
   b. II only
   c. III only
   d. I and II only
   e. I, II, and III

5. Which of the following will happen in response if perfectly competitive firms are earning positive economic profit?
   a. Firms will exit the industry.
   b. The short-run industry supply curve will shift right.
   c. The short-run industry supply curve will shift left.
   d. Firm output will increase.
   e. Market price will increase.

Tackle the Test: Free-Response Questions

1. Draw a correctly labeled graph showing a perfectly competitive firm in long-run equilibrium.

Answer (7 points)



1 point: Axes are correctly labeled.

1 point: Demand curve is horizontal and labeled with some combination of "*P*," "*MR*," or "*D*."

1 point: Marginal cost curve is labeled and slopes upward.

1 point: Profit-maximizing quantity is labeled on horizontal axis where *MC* = *MR*.

1 point: Average total cost curve is labeled and U–shaped.

1 point: Average total cost is equal to price at the profit-maximizing output

1 point: Marginal cost curve crosses the average total cost curve at the lowest point on the average total cost curve

2. Draw correctly labeled side-by-side graphs to show the long-run adjustment that would take place if perfectly competitive firms were earning a profit.

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

- How a monopolist determines the profit-maximizing price and quantity
- How to determine whether a monopoly is earning a profit or a loss

Module 61 Introduction to Monopoly

In this module we turn to monopoly, the market structure at the opposite end of the spectrum from perfect competition. A monopolist's profit-maximizing decision is subtly different from that of a price-taking producer, but it has large implications for the output produced and the welfare created. We will see the crucial role that market demand plays in leading a monopolist to behave differently from a firm in a perfectly competitive industry.

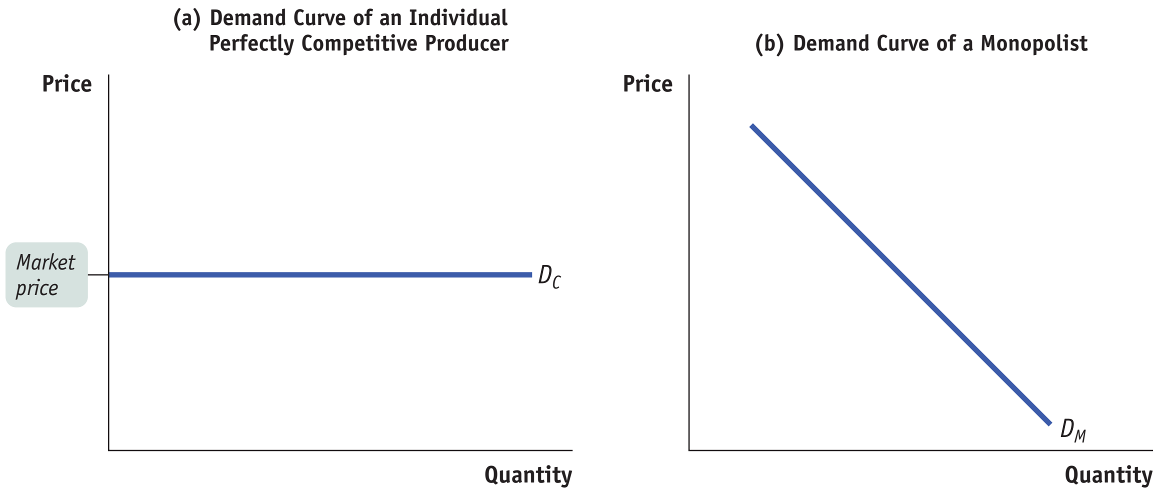
The Monopolist's Demand Curve and Marginal Revenue

Recall the firm's optimal output rule: a profit-maximizing firm produces the quantity of output at which the marginal cost of producing the last unit of output equals marginal revenue—the change in total revenue generated by the last unit of output. That is, $MR = MC$ at the profit-maximizing quantity of output. Although the optimal output rule holds for *all* firms, decisions about price and the quantity of output differ between monopolies and perfectly competitive industries due to differences in the demand curves faced by monopolists and perfectly competitive firms.

We have learned that even though the *market* demand curve always slopes downward, each of the firms that make up a perfectly competitive industry faces a horizontal, *perfectly elastic* demand curve, like D_C in panel (a) of Figure 61.1. Any attempt by an individual firm in a perfectly competitive industry to charge more than the going market price will cause the firm to lose all its sales. It can, however, sell as much as it likes at the market price. We saw that the marginal revenue of a perfectly competitive firm is simply the market price. As a result, the price-taking firm's optimal output rule is to produce the output level at which the marginal cost of the last unit produced is equal to the market price.

A monopolist, in contrast, is the sole supplier of its good. So its demand curve is simply the market demand curve, which slopes downward, like D_M in panel (b) of

figure 61.1 Comparing the Demand Curves of a Perfectly Competitive Producer and a Monopolist



Because an individual perfectly competitive producer cannot affect the market price of the good, it faces a horizontal demand curve D_C, as shown in panel (a). A monopolist, on the other hand, can affect the price. Because it is the sole supplier in the industry, its demand curve is the market demand curve D_M, as shown in panel (b). To sell more output, it must lower the price; by reducing output, it raises the price.

Figure 61.1. This downward slope creates a "wedge" between the price of the good and the marginal revenue of the good. Table 61.1 on the next page shows how this wedge develops. The first two columns of Table 61.1 show a hypothetical demand schedule for De Beers diamonds. For the sake of simplicity, we assume that all diamonds are exactly alike. And to make the arithmetic easy, we suppose that the number of diamonds sold is far smaller than is actually the case. For instance, at a price of $500 per diamond, we assume that only 10 diamonds are sold. The demand curve implied by this schedule is shown in panel (a) of Figure 61.2 on page 611.

The third column of Table 61.1 shows De Beers's total revenue from selling each quantity of diamonds—the price per diamond multiplied by the number of diamonds sold. The last column shows marginal revenue, the change in total revenue from producing and selling another diamond.

Clearly, after the 1st diamond, the marginal revenue a monopolist receives from selling one more unit is less than the price at which that unit is sold. For example, if De Beers sells 10 diamonds, the price at which the 10th diamond is sold is $500. But the marginal revenue—the change in total revenue in going from 9 to 10 diamonds—is only $50.

Why is the marginal revenue from that 10th diamond less than the price? Because an increase in production by a monopolist has two opposing effects on revenue:

- *A quantity effect.* One more unit is sold, increasing total revenue by the price at which the unit is sold (in this case, +$500).
- *A price effect.* In order to sell that last unit, the monopolist must cut the market price on *all* units sold. This decreases total revenue (in this case, by $9 \times -\$50 = -\$450$).

The quantity effect and the price effect are illustrated by the two shaded areas in panel (a) of Figure 61.2. Increasing diamond sales from 9 to 10 means moving down the demand curve from *A* to *B,* reducing the price per diamond from $550 to $500. The green-shaded area represents the quantity effect: De Beers sells the 10th diamond at a price of $500. This is offset, however, by the price effect, represented by the orange-shaded area. In order to

table 61.1

Demand, Total Revenue, and Marginal Revenue for the De Beers Diamond Monopoly

Price of diamond *P*	Quantity of diamonds demanded *Q*	Total revenue $TR = P \times Q$	Marginal revenue $MR = \Delta TR/\Delta Q$
$1,000	0	$0	
			$950
950	1	950	
			850
900	2	1,800	
			750
850	3	2,550	
			650
800	4	3,200	
			550
750	5	3,750	
			450
700	6	4,200	
			350
650	7	4,550	
			250
600	8	4,800	
			150
550	9	4,950	
			50
500	10	5,000	
			–50
450	11	4,950	
			–150
400	12	4,800	
			–250
350	13	4,550	
			–350
300	14	4,200	
			–450
250	15	3,750	
			–550
200	16	3,200	
			–650
150	17	2,550	
			–750
100	18	1,800	
			–850
50	19	950	
			–950
0	20	0	

sell that 10th diamond, De Beers must reduce the price on all its diamonds from $550 to $500. So it loses 9 × $50 = $450 in revenue, the orange-shaded area. So, as point *C* indicates, the total effect on revenue of selling one more diamond—the marginal revenue—derived from an increase in diamond sales from 9 to 10 is only $50.

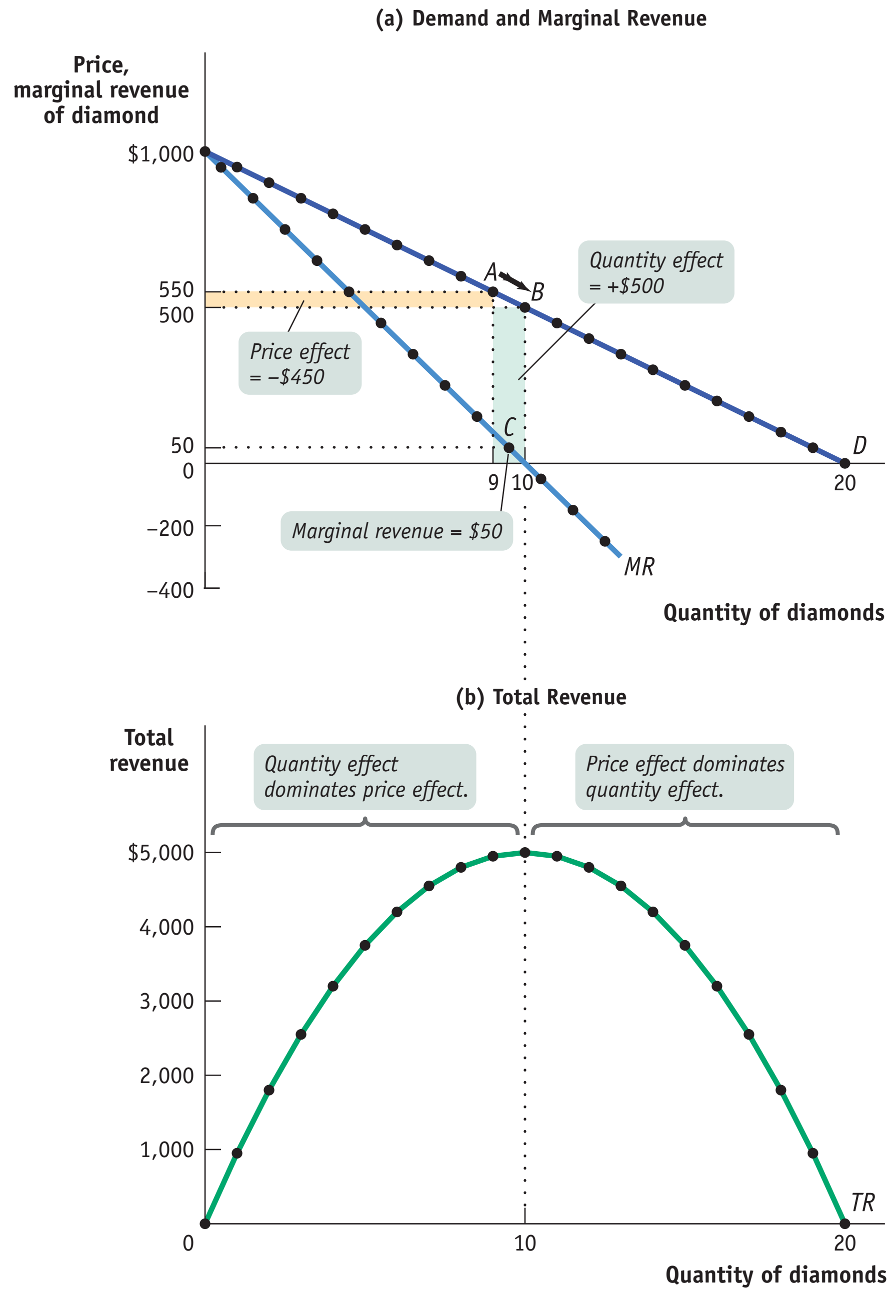
Point *C* lies on the monopolist's marginal revenue curve, labeled *MR* in panel (a) of Figure 61.2 and taken from the last column of Table 61.1. The crucial point about the monopolist's marginal revenue curve is that it is always *below* the demand curve. That's because of the price effect, which means that a monopolist's marginal revenue from selling an additional unit is always less than the price the monopolist receives for that unit. It is the price effect that creates the wedge between the monopolist's marginal revenue curve and the demand curve: in order to sell an additional diamond, De Beers must cut the market price on all units sold.

In fact, this wedge exists for any firm that possesses market power, such as an oligopolist, except in the case of price discrimination as explained in Module 63. Having market

figure 61.2

A Monopolist's Demand, Total Revenue, and Marginal Revenue Curves

Panel (a) shows the monopolist's demand and marginal revenue curves for diamonds from Table 61.1. The marginal revenue curve lies below the demand curve. To see why, consider point *A* on the demand curve, where 9 diamonds are sold at $550 each, generating total revenue of $4,950. To sell a 10th diamond, the price on all 10 diamonds must be cut to $500, as shown by point *B*. As a result, total revenue increases by the green area (the quantity effect: +$500) but decreases by the orange area (the price effect: –$450). So the marginal revenue from the 10th diamond is $50 (the difference between the green and orange areas), which is much lower than its price, $500. Panel (b) shows the monopolist's total revenue curve for diamonds. As output goes from 0 to 10 diamonds, total revenue increases. It reaches its maximum at 10 diamonds—the level at which marginal revenue is equal to 0—and declines thereafter. The quantity effect dominates the price effect when total revenue is rising; the price effect dominates the quantity effect when total revenue is falling.



power means that the firm faces a downward-sloping demand curve. As a result, there will always be a price effect from an increase in output for a firm with market power that charges every customer the same price. So for such a firm, the marginal revenue curve always lies below the demand curve.

Take a moment to compare the monopolist's marginal revenue curve with the marginal revenue curve for a perfectly competitive firm, which has no market power. For such a firm there is no price effect from an increase in output: its marginal revenue curve is simply its horizontal demand curve. So for a perfectly competitive firm, market price and marginal revenue are always equal.

Corbis

To emphasize how the quantity and price effects offset each other for a firm with market power, De Beers's total revenue curve is shown in panel (b) of Figure 61.2. Notice that it is hill-shaped: as output rises from 0 to 10 diamonds, total revenue increases. This reflects the fact that *at low levels of output, the quantity effect is stronger than the price effect:* as the monopolist sells more, it has to lower the price on only very few units, so the price effect is small. As output rises beyond 10 diamonds, total revenue actually falls. This reflects the fact that *at high levels of output, the price effect is stronger than the quantity effect:* as the monopolist sells more, it now has to lower the price on many units of output, making the price effect very large. Correspondingly, the marginal revenue curve lies below zero at output levels above 10 diamonds. For example, an increase in diamond production from 11 to 12 yields only $400 for the 12th diamond, simultaneously reducing the revenue from diamonds 1 through 11 by $550. As a result, the marginal revenue of the 12th diamond is –$150.

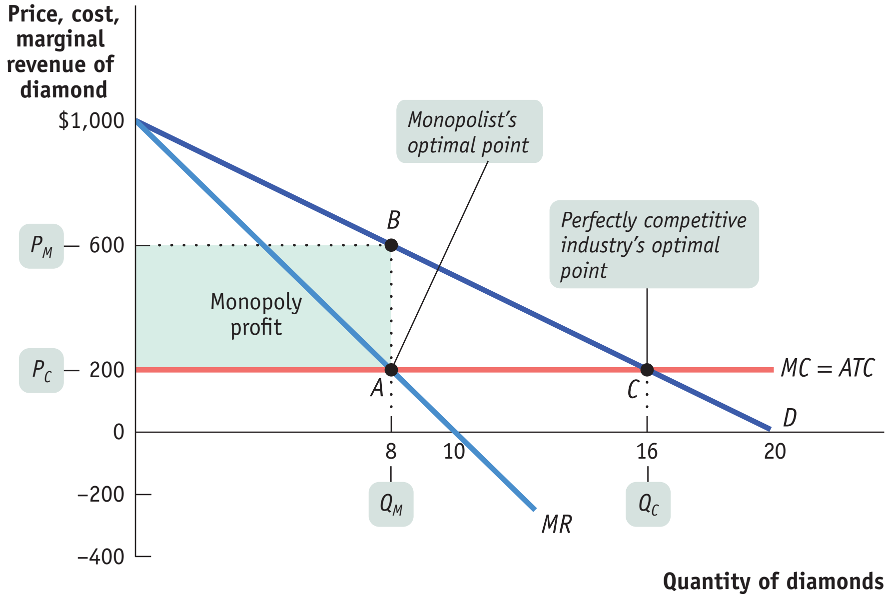
The Monopolist's Profit-Maximizing Output and Price

To complete the story of how a monopolist maximizes profit, we now bring in the monopolist's marginal cost. Let's assume that there is no fixed cost of production; we'll also assume that the marginal cost of producing an additional diamond is constant at $200, no matter how diamonds De Beers produces. Then marginal cost will always equal average total cost, and the marginal cost curve (and the average total cost curve) is a horizontal line at $200, as shown in Figure 61.3.

figure 61.3

The Monopolist's Profit-Maximizing Output and Price

This figure shows the demand, marginal revenue, and marginal cost curves. Marginal cost per diamond is constant at $200, so the marginal cost curve is horizontal at $200. According to the optimal output rule, the profit-maximizing quantity of output for the monopolist is at $MR = MC$, shown by point *A*, where the marginal cost and marginal revenue curves cross at an output of 8 diamonds. The price De Beers can charge per diamond is found by going to the point on the demand curve directly above point *A*, which is point *B* here—a price of $600 per diamond. It makes a profit of $\$400 \times 8 = \$3,200$. A perfectly competitive industry produces the output level at which $P = MC$, given by point *C*, where the demand curve and marginal cost curves cross. So a competitive industry produces 16 diamonds, sells at a price of $200, and makes zero profit.



To maximize profit, the monopolist compares marginal cost with marginal revenue. If marginal revenue exceeds marginal cost, De Beers increases profit by producing more; if marginal revenue is less than marginal cost, De Beers increases profit by producing less. So the monopolist maximizes its profit by using the optimal output rule:

(61-1) $MR = MC$ at the monopolist's profit-maximizing quantity of output

The monopolist's optimal point is shown in Figure 61.3. At *A*, the marginal cost curve, *MC*, crosses the marginal revenue curve, *MR*. The corresponding output level, 8 diamonds, is the monopolist's profit-maximizing quantity of output, Q_M. The price at which consumers demand 8 diamonds is \$600, so the monopolist's price, P_M, is \$600—corresponding to point *B*. The average total cost of producing each diamond is \$200, so the monopolist earns a profit of \$600 – \$200 = \$400 per diamond, and total profit is 8 × \$400 = \$3,200, as indicated by the shaded area.

Monopoly versus Perfect Competition

When Cecil Rhodes consolidated many independent diamond producers into De Beers, he converted a perfectly competitive industry into a monopoly. We can now use our analysis to see the effects of such a consolidation.

Let's look again at Figure 61.3 and ask how this same market would work if, instead of being a monopoly, the industry were perfectly competitive. We will continue to assume that there is no fixed cost and that marginal cost is constant, so average total cost and marginal cost are equal.

If the diamond industry consists of many perfectly competitive firms, each of those producers takes the market price as given. That is, each producer acts as if its marginal revenue is equal to the market price. So each firm within the industry uses the price-taking firm's optimal output rule:

(61-2) $P = MC$ at the perfectly competitive firm's profit-maximizing quantity of output.

In Figure 61.3, this would correspond to producing at *C*, where the price per diamond, P_C, is \$200, equal to the marginal cost of production. So the profit-maximizing output of an industry under perfect competition, Q_C, is 16 diamonds.

But does the perfectly competitive industry earn any profit at *C*? No: the price of \$200 is equal to the average total cost per diamond. So there is no economic profit for this industry when it produces at the perfectly competitive output level.

We've already seen that once the industry is consolidated into a monopoly, the result is very different. The monopolist's marginal revenue is influenced by the price effect, so that marginal revenue is less than the price. That is,

(61-3) $P > MR = MC$ at the monopolist's profit-maximizing quantity of output

As we've already seen, the monopolist produces less than the competitive industry—8 diamonds rather than 16. The price under monopoly is \$600, compared with only \$200 under perfect competition. The monopolist earns a positive profit, but the competitive industry does not.

So, we can see that compared with a competitive industry, a monopolist does the following:

- produces a smaller quantity: $Q_M < Q_C$
- charges a higher price: $P_M > P_C$
- earns a profit

fyi

Monopoly Behavior and the Price Elasticity of Demand

A monopolist faces marginal revenue that is lower than the market price. But how much lower? The answer depends on the *price elasticity of demand.*

Remember that the price elasticity of demand determines how total revenue from sales changes when the price changes. If the price elasticity is greater than 1 (demand is elastic), a fall in the price increases total revenue because the rise in the quantity demanded outweighs the lower price of each unit sold. If the price elasticity is less than 1 (demand is inelastic), a lower price reduces total revenue.

When a monopolist increases output by one unit, it must reduce the market price in order to sell that unit. If the price elasticity of demand is less than 1, this will actually reduce revenue—that is, marginal revenue will be negative. The monopolist can increase revenue by producing more only if the price elasticity of demand is greater than 1; the higher the elasticity, the closer the additional revenue is to the initial market price.

What this tells us is that the difference between monopoly behavior and perfectly competitive behavior depends on the price elasticity of demand. A monopolist that faces highly elastic demand will behave almost like a firm in a perfectly competitive industry.

For example, Amtrak has a monopoly on intercity passenger service in the Northeast Corridor, but it has very little ability to raise prices: potential train travelers will switch to cars and planes. In contrast, a monopolist that faces less elastic demand—like most cable TV companies—will behave very differently from a perfect competitor: it will charge much higher prices and restrict output more.

KAREN BLEIER/AFP/Getty Images

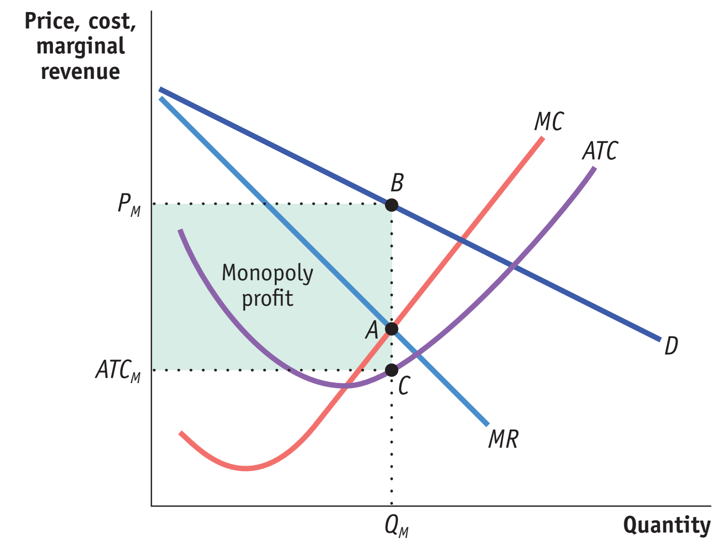
Monopoly: The General Picture

Figure 61.3 involved specific numbers and assumed that marginal cost was constant, there was no fixed cost, and therefore, that the average total cost curve was a horizontal line. Figure 61.4 shows a more general picture of monopoly in action: *D* is the market demand curve; *MR,* the marginal revenue curve; *MC,* the marginal cost curve; and *ATC,* the average total cost curve. Here we return to the usual assumption that the marginal cost curve has a "swoosh" shape and the average total cost curve is U-shaped.

figure 61.4

The Monopolist's Profit

In this case, the marginal cost curve has a "swoosh" shape and the average total cost curve is U-shaped. The monopolist maximizes profit by producing the level of output at which $MR = MC$, given by point *A*, generating quantity Q_M. It finds its monopoly price, P_M, from the point on the demand curve directly above point *A*, point *B* here. The average total cost of Q_M is shown by point *C*. Profit is given by the area of the shaded rectangle.



Applying the optimal output rule, we see that the profit-maximizing level of output, identified as the quantity at which marginal revenue and marginal cost intersect (see point *A*), is Q_M. The monopolist charges the highest price possible for this quantity, P_M, found at the height of the demand curve at Q_M (see point B). At the profit-maximizing level of output, the monopolist's average total cost is ATC_M (see point *C*).

Recalling how we calculated profit in Equation 59-1, profit is equal to the difference between total revenue and total cost. So we have

$$\textbf{(61-4)}\quad \text{Profit} = TR - TC = (P_M \times Q_M) - (ATC_M \times Q_M) = (P_M - ATC_M) \times Q_M.$$

Profit is equal to the area of the shaded rectangle in Figure 61.4, with a height of $P_M - ATC_M$ and a width of Q_M.

We learned that a perfectly competitive industry can have profits *in the short run but not in the long run.* In the short run, price can exceed average total cost, allowing a perfectly competitive firm to make a profit. But we also know that this cannot persist. In the long run, any profit in a perfectly competitive industry will be competed away as new firms enter the market. In contrast, while a monopoly can earn a profit or a loss in the short run, barriers to entry make it possible for a monopolist to make positive profits in the long run.

Module 61 AP Review

Solutions appear at the back of the book.

Check Your Understanding

1. Use the accompanying total revenue schedule of Emerald, Inc., a monopoly producer of 10-carat emeralds, to calculate the items listed in parts a–d. Then answer part e.

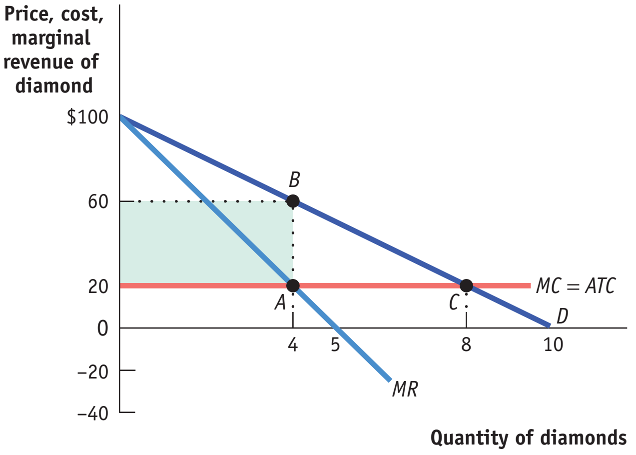
Quantity of emeralds demanded	*Total revenue*
1	$100
2	186
3	252
4	280
5	250

   a. the demand schedule (Hint: the average revenue at each quantity indicates the price at which that quantity would be demanded.)
   b. the marginal revenue schedule
   c. the quantity effect component of marginal revenue at each output level
   d. the price effect component of marginal revenue at each output level
   e. What additional information is needed to determine Emerald, Inc.'s profit-maximizing output?

2. Replicate Figure 61.3 and use your graph to show what happens to the following when the marginal cost of diamond production rises from $200 to $400. Use the information in Table 61.1 to identify specific numbers for prices and quantities on your graph.
   a. the marginal cost curve
   b. the profit-maximizing price and quantity
   c. the profit of the monopolist
   d. the quantity that would be produced if the diamond industry were perfectly competitive, and the associated profit

Tackle the Test: Multiple-Choice Questions

Refer to the graph provided for questions 1–4.

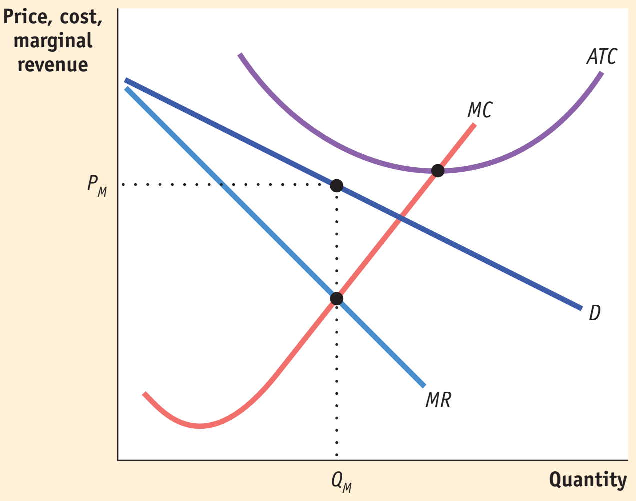


1. The monopolist's profit-maximizing output is
   a. 0.
   b. 4.
   c. 5.
   d. 8.
   e. 10.
2. The monopolist's total revenue equals
   a. $80.
   b. $160.
   c. $240.
   d. $300.
   e. $480.
3. The monopolist's total cost equals
   a. $20.
   b. $80.
   c. $160.
   d. $240.
   e. $480.
4. The monopolist is earning a profit equal to
   a. $0.
   b. $40.
   c. $80.
   d. $160.
   e. $240.
5. How does a monopoly differ from a perfectly competitive industry with the same costs?
   I. It produces a smaller quantity.
   II. It charges a higher price.
   III. It earns normal profits in the long run.
   a. I only
   b. II only
   c. III only
   d. I and II only
   e. I, II, and III

Tackle the Test: Free-Response Questions

1. a. Draw a correctly labeled graph showing a monopoly incurring a loss in the short run.
   b. How can the monopolist determine whether to shut down or produce at a loss in the short run?

Answer (10 points)



1 point: Axes are correctly labeled.

1 point: The demand curve is labeled and negatively sloped.

1 point: The marginal revenue curve is labeled, negatively sloped, and below the demand curve.

1 point: The marginal cost curve is labeled and slopes upward in the shape of a swoosh.

1 point: The profit-maximizing quantity is labeled on the horizontal axis where $MC = MR$.

1 point: Price is determined on the demand curve above the point where $MC = MR$.

1 point: The average total cost curve is labeled and U–shaped.

1 point: Average total cost is above price at the profit-maximizing output.

1 point: The marginal cost curve crosses the average total cost curve at the lowest point on the average total cost curve.

1 point: The firm will produce despite a loss in the short run if $P \geq AVC$.

2. a. Draw a graph showing a monopoly earning a normal profit in the short run.
   b. Can a monopoly earn a normal profit in the long run? Explain.

Module 62 Monopoly and Public Policy

What you will learn in this Module:

- The effects of the difference between perfect competition and monopoly on society's welfare
- How policy-makers address the problems posed by monopoly

It's good to be a monopolist, but it's not so good to be a monopolist's customer. A monopolist, by reducing output and raising prices, benefits at the expense of consumers. But buyers and sellers always have conflicting interests. Is the conflict of interest under monopoly any different from what it is under perfect competition?

The answer is yes, because monopoly is a source of inefficiency: the losses to consumers from monopoly behavior are larger than the gains to the monopolist. Because monopoly leads to net losses for the economy, governments often try either to prevent the emergence of monopolies or to limit their effects. In this module, we will see why monopoly leads to inefficiency and examine the policies governments adopt in an attempt to prevent this inefficiency.

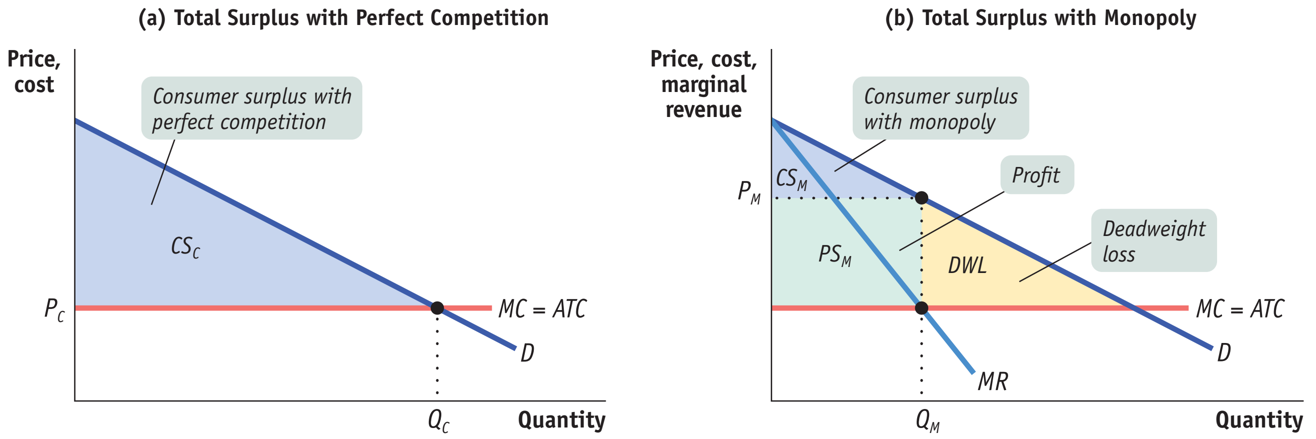
Welfare Effects of Monopoly

By holding output below the level at which marginal cost is equal to the market price, a monopolist increases its profit but hurts consumers. To assess whether this is a net benefit or loss to society, we must compare the monopolist's gain in profit to the consumers' loss. And what we learn is that the consumers' loss is larger than the monopolist's gain. Monopoly causes a net loss for society.

To see why, let's return to the case in which the marginal cost curve is horizontal, as shown in the two panels of Figure 62.1 on the next page. Here the marginal cost curve is *MC,* the demand curve is *D,* and, in panel (b), the marginal revenue curve is *MR.*

Panel (a) shows what happens if this industry is perfectly competitive. Equilibrium output is Q_C; the price of the good, P_C, is equal to marginal cost, and marginal cost is also equal to average total cost because there is no fixed cost and marginal cost is constant. Each firm is earning exactly its average total cost per unit of output, so there is no producer surplus in this equilibrium. The consumer surplus generated by the market is equal to the area of the blue-shaded triangle CS_C shown in panel (a). Since there is no producer surplus when the industry is perfectly competitive, CS_C also represents the total surplus.

figure 62.1 Monopoly Causes Inefficiency



Panel (a) depicts a perfectly competitive industry: output is Q_C, and market price, P_C, is equal to *MC*. Since price is exactly equal to each producer's average total cost of production per unit, there is no producer surplus. So total surplus is equal to consumer surplus, the entire shaded area. Panel (b) depicts the industry under monopoly: the monopolist decreases output to Q_M and charges P_M. Consumer surplus (blue area) has shrunk: a portion of it has been captured as profit (green area), and a portion of it has been lost to deadweight loss (yellow area), the value of mutually beneficial transactions that do not occur because of monopoly behavior. As a result, total surplus falls.

Panel (b) shows the results for the same market, but this time assuming that the industry is a monopoly. The monopolist produces the level of output, Q_M, at which marginal cost is equal to marginal revenue, and it charges the price, P_M. The industry now earns profit—which is also the producer surplus in this case—equal to the area of the green rectangle, PS_M. Note that this profit is part of what was consumer surplus in the perfectly competitive market, and consumer surplus with the monopoly shrinks to the area of the blue triangle, CS_M.

By comparing panels (a) and (b), we see that in addition to the redistribution of surplus from consumers to the monopolist, another important change has occurred: the sum of profit and consumer surplus—total surplus—is *smaller* under monopoly than under perfect competition. That is, the sum of CS_M and PS_M in panel (b) is less than the area CS_C in panel (a). Previously, we analyzed how taxes could cause *deadweight loss* for society. Here we show that a monopoly creates deadweight loss equal to the area of the yellow triangle, *DWL*. So monopoly produces a net loss for society.

This net loss arises because some mutually beneficial transactions do not occur. There are people for whom an additional unit of the good is worth more than the marginal cost of producing it but who don't consume it because they are not willing to pay the monopoly price, P_M. Indeed, by driving a wedge between price and marginal cost, a monopoly acts much like a tax on consumers and produces the same kind of inefficiency.

So monopoly power detracts from the welfare of society as a whole and is a source of market failure. Is there anything government policy can do about it?

Preventing Monopoly Power

Policy toward monopolies depends crucially on whether or not the industry in question is a natural monopoly, one in which increasing returns to scale ensure that a bigger producer has lower average total cost. If the industry is *not* a natural monopoly, the best policy is to prevent a monopoly from arising or break it up if it already exists.

Government policy used to prevent or eliminate monopolies is known as *antitrust policy,* which we will discuss in Module 77. For now, let's focus on the more difficult problem of dealing with a natural monopoly.

Dealing with a Natural Monopoly

Breaking up a monopoly that isn't natural is clearly a good idea: the gains to consumers outweigh the loss to the producer. But it's not so clear whether a natural monopoly, one in which large producers have lower average total costs than small producers, should be broken up, because this would raise average total cost. For example, a town government that tried to prevent a single company from dominating local gas supply—which, as we've discussed, is almost surely a natural monopoly—would raise the cost of providing gas to its residents.

Yet even in the case of a natural monopoly, a profit-maximizing monopolist acts in a way that causes inefficiency—it charges consumers a price that is higher than marginal cost and, by doing so, prevents some potentially beneficial transactions. Also, it can seem unfair that a firm that has managed to establish a monopoly position earns a large profit at the expense of consumers.

What can public policy do about this? There are two common answers.

Public Ownership

In many countries, the preferred answer to the problem of natural monopoly has been **public ownership.** Instead of allowing a private monopolist to control an industry, the government establishes a public agency to provide the good and protect consumers' interests.

The advantage of public ownership, in principle, is that a publicly owned natural monopoly can set prices based on the criterion of efficiency rather than profit maximization. In a perfectly competitive industry, profit-maximizing behavior *is* efficient because producers set price equal to marginal cost; that is why there is no economic argument for public ownership of, say, wheat farms.

Experience suggests, however, that public ownership as a solution to the problem of natural monopoly often works badly in practice. One reason is that publicly owned firms are often less eager than private companies to keep costs down or offer high-quality products. Another is that publicly owned companies all too often end up serving political interests—providing contracts or jobs to people with the right connections.

David Livingston/Getty Images

Regulation

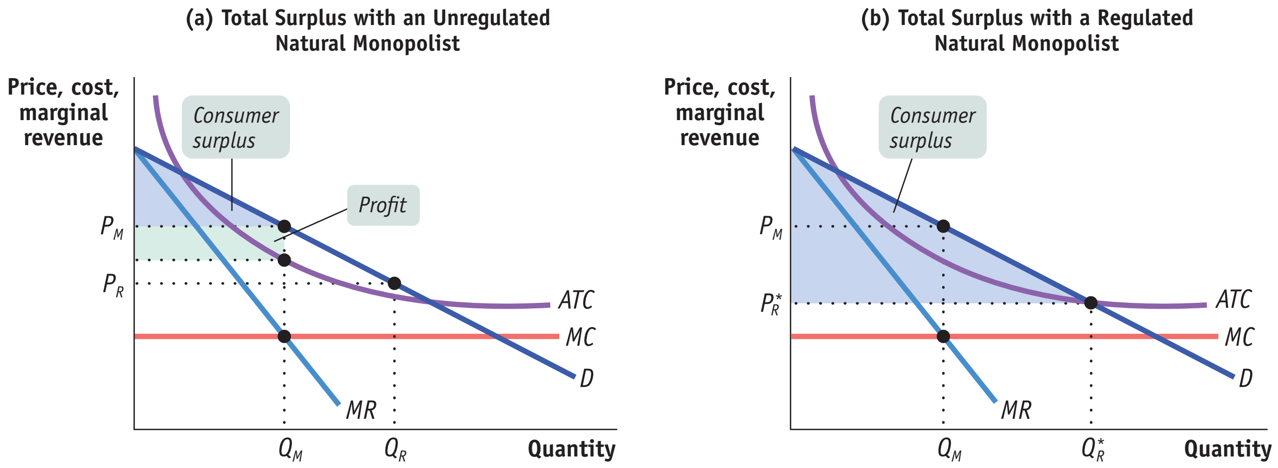
In the United States, the more common answer has been to leave the industry in private hands but subject it to regulation. In particular, most local utilities, like electricity, telephone service, natural gas, and so on, are covered by **price regulation** that limits the prices they can charge.

Figure 62.2 on the next page shows an example of price regulation of a natural monopoly—a highly simplified version of a local gas company. The company faces a demand curve, *D,* with an associated marginal revenue curve, *MR.* For simplicity, we assume that the firm's total cost consists of two parts: a fixed cost and a variable cost that is the same for every unit. So marginal cost is constant in this case, and the marginal cost curve (which here is also the average variable cost curve) is the horizontal line *MC.* The average total cost curve is the downward-sloping curve *ATC;* it slopes downward because the higher the output, the lower the average fixed cost (the fixed cost per unit of output). Because average total cost slopes downward over the range of output relevant for market demand, this is a natural monopoly.

In **public ownership** of a monopoly, the good is supplied by the government or by a firm owned by the government.

Price regulation limits the price that a monopolist is allowed to charge.

figure 62.2 Unregulated and Regulated Natural Monopoly



This figure shows the case of a natural monopolist. In panel (a), if the monopolist is allowed to charge P_M, it makes a profit, shown by the green area; consumer surplus is shown by the blue area. If it is regulated and must charge the lower price, P_R, output increases from Q_M to Q_R and consumer surplus increases. Panel (b) shows what happens when the monopolist must charge a price equal to average total cost, the price P_R^*. Output expands to Q_R^*, and consumer surplus is now the entire blue area. The monopolist makes zero profit. This is the greatest total surplus possible without the monopoly incurring losses.

Panel (a) illustrates a case of natural monopoly without regulation. The unregulated natural monopolist chooses the monopoly output Q_M and charges the price P_M. Since the monopolist receives a price greater than average total cost, she or he earns a profit, represented by the green-shaded rectangle in panel (a). Consumer surplus is given by the blue-shaded triangle.

Now suppose that regulators impose a price ceiling on local gas deliveries—one that falls below the monopoly price P_M but above average total cost, say, at P_R in panel (a). At that price the quantity demanded is Q_R.

Does the company have an incentive to produce that quantity? Yes. If the price the monopolist can charge is fixed at P_R by regulators, the firm can sell any quantity between zero and Q_R for the same price, P_R. Because it doesn't have to lower its price to sell more (up to Q_R), there is no price effect to bring marginal revenue below price, so the regulated price becomes the marginal revenue for the monopoly just like the market price is the marginal revenue for a perfectly competitive firm. With marginal revenue being above marginal cost and price exceeding average cost, the firm expands output to meet the quantity demanded, Q_R. This policy has appeal because at the regulated price, the monopolist produces more at a lower price.

Of course, the monopolist will not be willing to produce at all in the long run if the regulated price means producing at a loss. That is, the price ceiling has to be set high enough to allow the firm to cover its average total cost. Panel (b) shows a situation in which regulators have pushed the price down as far as possible, at the level where the average total cost curve crosses the demand curve. At any lower price the firm loses money. The price here, P_R^*, is the best regulated price: the monopolist is just willing to operate and produces Q_R^*, the quantity demanded at that price. Consumers and society gain as a result.

The welfare effects of this regulation can be seen by comparing the shaded areas in the two panels of Figure 62.2. Consumer surplus is increased by the regulation, with the gains coming from two sources. First, profits are eliminated and added instead to consumer surplus. Second, the larger output and lower price leads to an overall welfare gain—an increase in total surplus. In fact, panel (b) illustrates the largest total surplus possible.

Must Monopoly Be Controlled?

Sometimes the cure is worse than the disease. Some economists have argued that the best solution, even in the case of a natural monopoly, may be to live with it. The case for doing nothing is that attempts to control monopoly will, one way or another, do more harm than good.

The following FYI describes the case of cable television, a natural monopoly that has been alternately regulated and deregulated as politicians change their minds about the appropriate policy.

fyi

Cable Dilemmas

Most price regulation in the United States goes back a long way: electricity, local phone service, water, and gas have been regulated in most places for generations. But cable television is a relatively new industry. Until the late 1970s, only rural areas too remote to support local broadcast stations were served by cable. After 1972, new technology and looser rules made it profitable to offer cable service to major metropolitan areas; new networks like HBO and CNN emerged to take advantage of the possibilities.

Until recently, local cable TV was a natural monopoly: running cable through a town entails large fixed costs that don't depend on how many people actually subscribe. Having more than one cable company would involve a lot of wasteful duplication. But if the local cable company is a monopoly, should its prices be regulated?

At first, most local governments thought so, and cable TV was subject to price regulation. In 1984, however, Congress passed a law prohibiting most local governments from regulating cable prices. (The law was the result both of widespread skepticism about whether price regulation was actually a good idea and of intensive lobbying by the cable companies.)

After the law went into effect, however, cable television rates increased sharply. The resulting consumer backlash led to a new law, in 1992, which once again allowed local governments to set limits on cable prices.

Was the second round of regulation a success? As measured by the prices of "basic" cable service, it was: after rising rapidly during the period of deregulation, the cost of basic service leveled off.

However, price regulation in cable applies only to "basic" service. Cable operators can try to evade the restrictions by charging more for premium channels like HBO or by offering fewer channels in the "basic" package. So some skeptics have questioned whether current regulation has actually been effective.

Yet technological change has begun providing relief to consumers in some areas. Although cable TV is a natural monopoly, there is now another means of delivering video programs to homes: over a high-speed fiber-optic Internet connection. In some locations, fiber-optic Internet providers have begun competing aggressively with traditional cable TV companies. Studies have shown that when a second provider enters a market, prices can drop significantly, as much as 30%. In fact, the United States is currently behind on this front: today 60% of households in Hong Kong watch TV programs delivered over the Internet. What will these changes mean for the cable TV monopolies? Stay tuned.

Stringer/Getty Images

Module 62 AP Review

Solutions appear at the back of the book.

Check Your Understanding

1. What policy should the government adopt in the following cases? Explain.
   a. Internet service in Anytown, OH, is provided by cable. Customers feel they are being overcharged, but the cable company claims it must charge prices that let it recover the costs of laying cable.
   b. The only two airlines that currently fly to Alaska need government approval to merge. Other airlines wish to fly to Alaska but need government-allocated landing slots to do so.

2. True or false? Explain your answer.
   a. Society's welfare is lower under monopoly because some consumer surplus is transformed into profit for the monopolist.
   b. A monopolist causes inefficiency because there are consumers who are willing to pay a price greater than or equal to marginal cost but less than the monopoly price.

3. Suppose a monopolist mistakenly believes that her or his marginal revenue is always equal to the market price. Assuming constant marginal cost and no fixed cost, draw a diagram comparing the level of profit, consumer surplus, total surplus, and deadweight loss for this misguided monopolist compared to a smart monopolist.

Tackle the Test: Multiple-Choice Questions

1. Which of the following statements is true of a monopoly as compared to a perfectly competitive market with the same costs?
   I. Consumer surplus is smaller.
   II. Profit is smaller.
   III. Deadweight loss is smaller.
   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 of a natural monopoly?
   a. It experiences diseconomies of scale.
   b. *ATC* is lower if there is a single firm in the market.
   c. It occurs in a market that relies on natural resources for its production.
   d. There are decreasing returns to scale in the industry.
   e. The government must provide the good or service to achieve efficiency.

3. Which of the following government actions is the most common for a natural monopoly in the United States?
   a. prevent its formation
   b. break it up using antitrust laws
   c. use price regulation
   d. public ownership
   e. elimination of the market

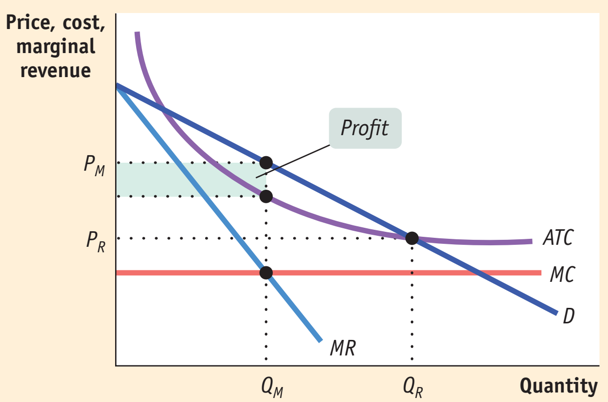
4. Which of the following markets is an example of a regulated natural monopoly?
   a. local cable TV
   b. gasoline
   c. cell phone service
   d. organic tomatoes
   e. diamonds

5. Which of the following is most likely to be higher for a regulated natural monopoly than for an unregulated natural monopoly?
   a. product variety
   b. quantity
   c. price
   d. profit
   e. deadweight loss

Tackle the Test: Free-Response Questions

1. Draw a correctly labeled graph showing a profit-making natural monopoly. On your graph, indicate each of the following:
   a. the monopoly's profit-maximizing output (Q_M)
   b. the monopoly's price (P_M)
   c. the monopoly's profit
   d. the regulated price that would maximize consumer surplus without creating losses for the firm (P_R)

Answer (9 points)



1 point: The axes are correctly labeled.

1 point: The demand curve is labeled and sloped downward.

1 point: The marginal revenue curve is labeled, sloped downward, and below the demand curve.

1 point: The average total cost curve is labeled and downward sloping (not U-shaped!).

1 point: The marginal cost curve is labeled and below the average total cost curve.

1 point: The profit-maximizing output, Q_M, is shown on the horizontal axis where $MC = MR$.

1 point: The profit-maximizing price is found on the demand curve above the point where $MC = MR$.

1 point: The monopoly profit area is correctly shaded and identified.

1 point: The regulated price is labeled on the vertical axis to the left of the point where the demand curve crosses the average total cost curve.

2. Draw a correctly labeled graph of a natural monopoly. Use your graph to identify each of the following:
   a. consumer surplus if the market were somehow able to operate as a perfectly competitive market
   b. consumer surplus with the monopoly
   c. monopoly profit
   d. deadweight loss with the monopoly

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

- The meaning of price discrimination
- Why price discrimination is so prevalent when producers have market power

Module 63 Price Discrimination

Up to this point, we have considered only the case of a monopolist who charges all consumers the same price. However, monopolists want to maximize their profits and often they do so by charging different prices for the same product. In this module we look at how monopolists increase their profits by engaging in *price discrimination*.

Price Discrimination Defined

A monopolist who charges everyone the same price is known as a **single-price monopolist.** As the term suggests, not all monopolists do this. In fact, many monopolists find that they can increase their profits by selling the same good to different customers for different prices: they practice **price discrimination.**

An example of price discrimination that travelers encounter regularly involves airline tickets. Although there are a number of airlines, most routes in the United States are serviced by only one or two carriers, which, as a result, have market power and can influence prices. So any regular airline passenger quickly becomes aware that the simple question "How much will it cost me to fly there?" rarely has a simple answer. If you are willing to buy a nonrefundable ticket a month in advance and stay over a Saturday night, the round trip may cost only $150—or less if you are a senior citizen or a student. But if you have to go on a business trip tomorrow, which happens to be Tuesday, and want to come back on Wednesday, the same round trip might cost $550. Yet the business traveler and the visiting grandparent receive the same product.

You might object that airlines are not usually monopolies—that in most flight markets the airline industry is an oligopoly. In fact, price discrimination takes place under oligopoly and monopolistic competition as well as monopoly. But it doesn't happen under perfect competition. And once we've seen why monopolists sometimes price-discriminate, we'll be in a good position to understand why it happens in other cases, too.

A **single-price monopolist** charges all consumers the same price.

Sellers engage in **price discrimination** when they charge different prices to different consumers for the same good.

The Logic of Price Discrimination

To get a preliminary view of why price discrimination might be more profitable than charging all consumers the same price, imagine that Air Sunshine offers the only nonstop flights between Bismarck, North Dakota, and Ft. Lauderdale, Florida. Assume that there are no capacity problems—the airline can fly as many planes as the number of passengers warrants. Also assume that there is no fixed cost. The marginal cost to the airline of providing a seat is $125 however many passengers it carries.

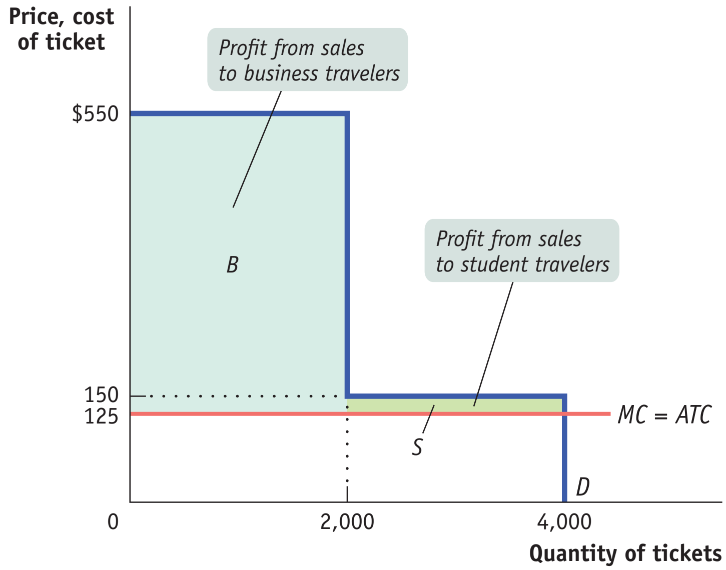
Further assume that the airline knows there are two kinds of potential passengers. First, there are business travelers, 2,000 of whom want to travel between the destinations each week. Second, there are high school students, 2,000 of whom also want to travel each week.

Will potential passengers take the flight? It depends on the price. The business travelers, it turns out, really need to fly; they will take the plane as long as the price is no more than $550. Since they are flying purely for business, we assume that cutting the price below $550 will not lead to any increase in business travel. The students, however, have less money and more time; if the price goes above $150, they will take the bus. The implied demand curve is shown in Figure 63.1.

figure 63.1

Two Types of Airline Customers

Air Sunshine has two types of customers, business travelers willing to pay at most $550 per ticket and students willing to pay at most $150 per ticket. There are 2,000 of each kind of customer. Air Sunshine has constant marginal cost of $125 per seat. If Air Sunshine could charge these two types of customers different prices, it would maximize its profit by charging business travelers $550 and students $150 per ticket. It would capture all of the consumer surplus as profit.



So what should the airline do? If it has to charge everyone the same price, its options are limited. It could charge $550; that way it would get as much as possible out of the business travelers but lose the student market. Or it could charge only $150; that way it would get both types of travelers but would make significantly less money from sales to business travelers.

We can quickly calculate the profits from each of these alternatives. If the airline charged $550, it would sell 2,000 tickets to the business travelers, earning a total revenue of $2{,}000 \times \$550 = \$1.1$ million and incurring costs of $2{,}000 \times \$125 = \$250{,}000$; so its profit would be $850,000, illustrated by the shaded area B in Figure 63.1. If the airline charged only $150, it would sell 4,000 tickets, receiving revenue of $4{,}000 \times \$150 = \$600{,}000$ and incurring costs of $4{,}000 \times \$125 = \$500{,}000$; so its profit would be $100,000. If the airline must charge everyone the same price, charging the higher price and forgoing sales to students is clearly more profitable.

What the airline would really like to do, however, is charge the business travelers the full $550 but offer $150 tickets to the students. That's a lot less than the price paid by business travelers, but it's still above marginal cost; so if the airline could sell those extra 2,000 tickets to students, it would make an additional $50,000 in profit. That is, it would make a profit equal to the areas *B* plus *S* in Figure 63.1.

It would be more realistic to suppose that there is some "give" in the demand of each group: at a price below $550, there would be some increase in business travel; and at a price above $150, some students would still purchase tickets. But this, it turns out, does not do away with the argument for price discrimination. The important point is that the two groups of consumers differ in their *sensitivity to price*—that a high price has a larger effect in discouraging purchases by students than by business travelers. As long as different groups of customers respond differently to the price, a monopolist will find that it can capture more consumer surplus and increase its profit by charging them different prices.

Price Discrimination and Elasticity

A more realistic description of the demand that airlines face would not specify particular prices at which different types of travelers would choose to fly. Instead, it would distinguish between the groups on the basis of their sensitivity to the price—their price elasticity of demand.

Suppose that a company sells its product to two easily identifiable groups of people—business travelers and students. It just so happens that business travelers are very insensitive to the price: there is a certain amount of the product they just have to have whatever the price, but they cannot be persuaded to buy much more than that no matter how cheap it is. Students, though, are more flexible: offer a good enough price and they will buy quite a lot, but raise the price too high and they will switch to something else. Which approach is best for the company in this case?

The answer is the one already suggested by our simplified example: the company should charge business travelers, with their low price elasticity of demand, a higher price than it charges students, with their high price elasticity of demand.

The actual situation of the airlines is very much like this hypothetical example. Business travelers typically place a high priority on being in the right place at the right time and are not very sensitive to the price. But leisure travelers are fairly sensitive to the price: faced with a high price, they might take the bus, drive to another airport to get a lower fare, or skip the trip altogether.

William Thomas Cain/Getty

On many airline routes, the fare you pay depends on the type of traveler you are.

So why doesn't an airline simply announce different prices for business and leisure customers? First, this would probably be illegal. (U.S. law places some limits on the ability of companies to practice blatant price discrimination.) Second, even if it were legal, it would be a hard policy to enforce: business travelers might be willing to wear casual clothing and claim they were visiting family in Ft. Lauderdale in order to save $400.

So what the airlines do—quite successfully—is impose rules that indirectly have the effect of charging business and leisure travelers different fares. Business travelers usually travel during the week and want to be home on the weekend, so the round-trip fare is much higher if you don't stay over a Saturday night. The requirement of a weekend stay for a cheap ticket effectively separates business travelers from leisure travelers. Similarly, business travelers often visit several cities in succession rather than make a simple round trip; so round-trip fares are much lower than twice the one-way fare. Many business trips are scheduled on short notice, so fares are much lower if you book far in advance. Fares are also lower if you travel standby, taking your chances on whether you actually get a seat—business travelers have to make it to that meeting; people visiting their relatives don't. And because customers must show their ID at check-in, airlines make sure there are no resales of tickets between the two groups that would undermine their ability to price-discriminate—students can't buy cheap tickets and resell them to business travelers. Look at the rules that govern ticket pricing, and you will see an ingenious implementation of profit-maximizing price discrimination.

Perfect Price Discrimination

Perfect price discrimination takes place when a monopolist charges each consumer his or her willingness to pay—the maximum that the consumer is willing to pay.

Let's return to the example of business travelers and students traveling between Bismarck and Ft. Lauderdale, illustrated in Figure 63.1, and ask what would happen if the airline could distinguish between the two groups of customers in order to charge each a different price.

Clearly, the airline would charge each group its *willingness to pay*—that is, the maximum that each group is willing to pay. For business travelers, the willingness to pay is $550; for students, it is $150. As we have assumed, the marginal cost is $125 and does not depend on output, making the marginal cost curve a horizontal line. And as we noted earlier, we can easily determine the airline's profit: it is the sum of the areas of rectangle *B* and rectangle *S*.

In this case, the consumers do not get any consumer surplus! The entire surplus is captured by the monopolist in the form of profit. When a monopolist is able to capture the entire surplus in this way, we say that the monopolist achieves **perfect price discrimination.**

In general, the greater the number of different prices charged, the closer the monopolist is to perfect price discrimination. Figure 63.2 on the next page shows a monopolist facing a downward-sloping demand curve, a monopolist who we assume is able to charge different prices to different groups of consumers, with the consumers who are willing to pay the most being charged the most. In panel (a) the monopolist charges two different prices; in panel (b) the monopolist charges three different prices. Two things are apparent:

- The greater the number of prices the monopolist charges, the lower the lowest price—that is, some consumers will pay prices that approach marginal cost.
- The greater the number of prices the monopolist charges, the more money extracted from consumers.

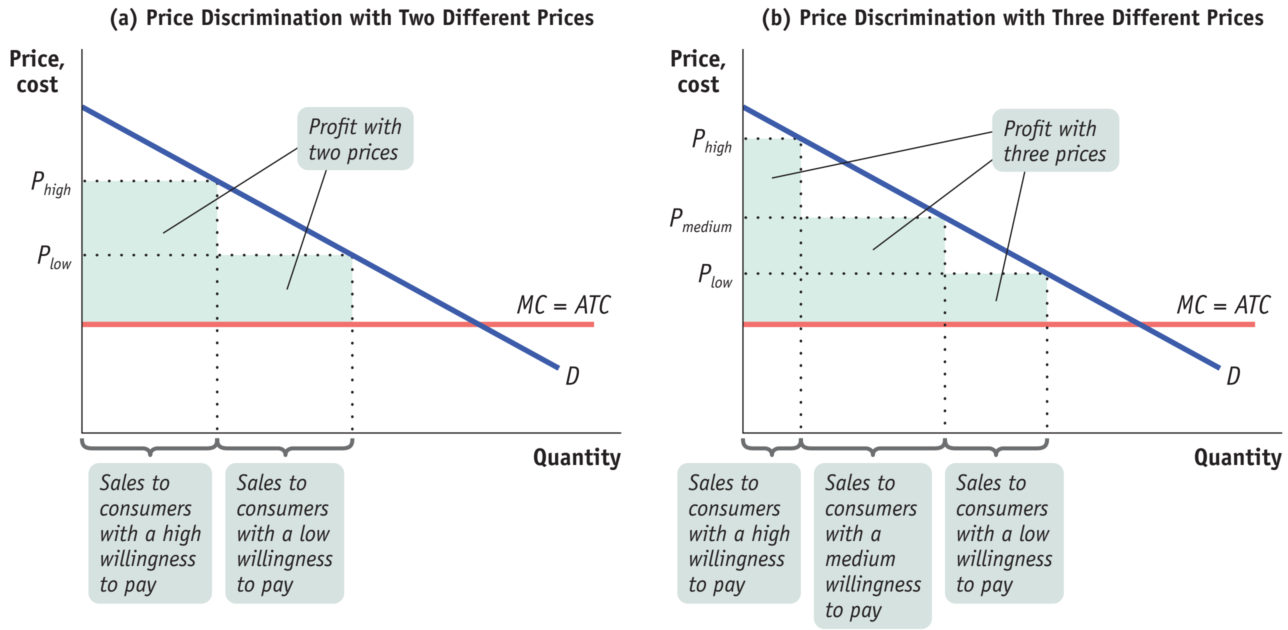
With a very large number of different prices, the picture would look like panel (c), a case of perfect price discrimination. Here, every consumer pays the most he or she is willing to pay, and the entire consumer surplus is extracted as profit.

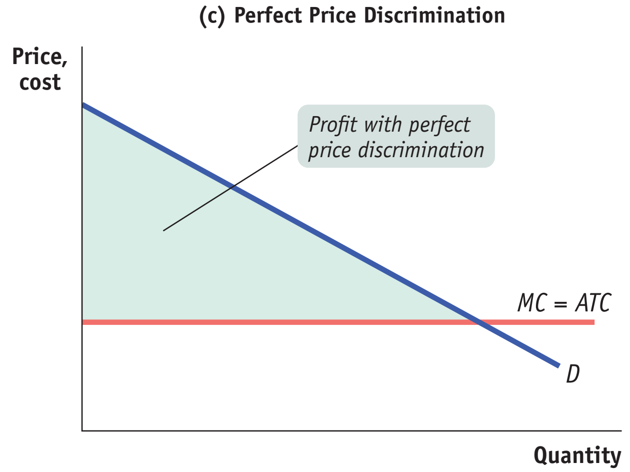
Both our airline example and the example in Figure 63.2 can be used to make another point: a monopolist who can engage in perfect price discrimination doesn't cause any inefficiency! The reason is that the source of inefficiency is eliminated: all potential consumers who are willing to purchase the good at a price equal to or above marginal cost are able to do so. The perfectly price-discriminating monopolist manages to "scoop up" all consumers by offering some of them lower prices than others.

Perfect price discrimination is almost never possible in practice. At a fundamental level, the inability to achieve perfect price discrimination is a problem of prices as economic signals. When prices work as economic signals, they convey the information needed to ensure that all mutually beneficial transactions will indeed occur: the market price signals the seller's cost, and a consumer signals willingness to pay by purchasing the good whenever that willingness to pay is at least as high as the market price. The problem in reality, however, is that prices are often not perfect signals: a consumer's true willingness to pay can be disguised, as by a business traveler who claims to be a student when buying a ticket in order to obtain a lower fare. When such disguises work, a monopolist cannot achieve perfect price discrimination. However, monopolists do try to move in the direction of perfect price discrimination through a variety of pricing strategies. Common techniques for price discrimination include the following:

- *Advance purchase restrictions.* Prices are lower for those who purchase well in advance (or in some cases for those who purchase at the last minute). This separates those who are likely to shop for better prices from those who won't.
- *Volume discounts.* Often the price is lower if you buy a large quantity. For a consumer who plans to consume a lot of a good, the cost of the last unit—the marginal cost to the consumer—is considerably less than the average price. This separates those who plan to buy a lot, and so are likely to be more sensitive to price, from those who don't.

figure 63.2 Price Discrimination





Panel (a) shows a monopolist that charges two different prices; its profit is shown by the shaded area. Panel (b) shows a monopolist that charges three different prices; its profit, too, is shown by the shaded area. It is able to capture more of the consumer surplus and to increase its profit. That is, by increasing the number of different prices charged, the monopolist captures more of the consumer surplus and makes a larger profit. Panel (c) shows the case of perfect price discrimination, where a monopolist charges each consumer his or her willingness to pay; the monopolist's profit is given by the shaded triangle.

Alamy

- *Two-part tariffs.* In a discount club like Costco or Sam's Club (which are not monopolists but monopolistic competitors), you pay an annual fee (the first part of the tariff) in addition to the price of the item(s) you purchase (the second part of the tariff). So the full price of the first item you buy is in effect much higher than that of subsequent items, making the two-part tariff behave like a volume discount.

Our discussion also helps explain why government policies on monopoly typically focus on preventing deadweight loss, not preventing price discrimination—unless it causes serious issues of equity. Compared to a single-price monopolist, price discrimination—even when it is not perfect—can increase the efficiency of the market. When a single, medium-level price is replaced by a high price and a low price, some consumers who were formerly priced out of the market will be able to purchase the good. The price discrimination increases efficiency because more of the units for which the willingness to pay (as indicated by the height of the demand curve) exceeds the marginal cost are produced and sold. Consider a drug that is disproportionately prescribed to senior citizens, who are often on fixed incomes and so are very sensitive to price. A policy that allows a drug company to charge senior citizens a low price and everyone else a high price will serve more consumers and create more total surplus than a situation in which everyone is charged the same price. But price discrimination that creates serious concerns about equity is likely to be prohibited—for example, an ambulance service that charges patients based on the severity of their emergency.

Module 63 AP Review

Solutions appear at the back of the book.

Check Your Understanding

1. True or false? Explain your answer.
   a. A single-price monopolist sells to some customers that would not find the product affordable if purchasing from a price-discriminating monopolist.
   b. A price-discriminating monopolist creates more inefficiency than a single-price monopolist because it captures more of the consumer surplus.
   c. Under price discrimination, a customer with highly elastic demand will pay a lower price than a customer with inelastic demand.
2. Which of the following are cases of price discrimination and which are not? In the cases of price discrimination, identify the consumers with high price elasticity of demand and those with low price elasticity of demand.
   a. Damaged merchandise is marked down.
   b. Restaurants have senior citizen discounts.
   c. Food manufacturers place discount coupons for their merchandise in newspapers.
   d. Airline tickets cost more during the summer peak flying season.

Tackle the Test: Multiple-Choice Questions

1. Which of the following characteristics is necessary in order for a firm to price discriminate?
   a. free entry and exit
   b. differentiated product
   c. many sellers
   d. some control over price
   e. horizontal demand curve
2. Price discrimination
   a. is the opposite of volume discounts.
   b. is a practice limited to movie theaters and the airline industry.
   c. can lead to increased efficiency in the market.
   d. rarely occurs in the real world.
   e. helps to increase the profits of perfect competitors.
3. With perfect price discrimination, consumer surplus
   a. is maximized.
   b. equals zero.
   c. is increased.
   d. cannot be determined.
   e. is the area below the demand curve above MC.
4. Which of the following is a technique used by price discriminating monopolists?
   I. advance purchase restrictions
   II. two-part tariffs
   III. volume discounts
   a. I only
   b. II only
   c. III only
   d. I and II only
   e. I, II, and III
5. A price discriminating monopolist will charge a higher price to consumers with
   a. a more inelastic demand.
   b. a less inelastic demand.
   c. higher income.
   d. lower willingness to pay.
   e. less experience in the market.

Tackle the Test: Free-Response Questions

1. a. Define price discrimination.
   b. Why do firms price discriminate?
   c. In which market structures can firms price discriminate? Explain why.
   d. Give an example of price discrimination.

Answer (5 points)

1 point: Price discrimination is the practice of charging different prices to different customers for the same product.

1 point: Firms price discriminate to increase their profits.

1 point: In order to price discriminate, firms must be in the monopoly, oligopoly, or monopolistic competition market structure.

1 point: Because rather than being price-takers, firms in these market structures have some degree of market power, which gives them the ability to charge more than one price.

1 point: An example is different prices for movie tickets charged for people of different ages.

2. Draw a correctly labeled graph showing a monopoly practicing perfect price discrimination. On your graph, identify the monopoly's profit. What does consumer surplus equal in this case? Explain.

Section 11 Review

Summary

1. A producer chooses output according to the **price-taking firm's optimal output rule:** produce the quantity at which price equals marginal cost. However, a firm that produces the optimal quantity may not be profitable.

2. A firm is profitable if total revenue exceeds total cost or, equivalently, if the market price exceeds its **break-even price**—minimum average total cost. If market price exceeds the break-even price, the firm is profitable. If market price is less than minimum average total cost, the firm is unprofitable. If market price is equal to minimum average total cost, the firm breaks even. When profitable, the firm's per-unit profit is $P - ATC$; when unprofitable, its per-unit loss is $ATC - P$.

3. Fixed cost is irrelevant to the firm's optimal short-run production decision. The short-run production decision depends on the firm's **shut-down price**—its minimum average variable cost—and the market price. When the market price is equal to or exceeds the shut-down price, the firm produces the output quantity at which marginal cost equals the market price. When the market price falls below the shut-down price, the firm ceases production in the short run. This generates the firm's **short-run individual supply curve.**

4. Fixed cost matters over time. If the market price is below minimum average total cost for an extended period of time, firms will exit the industry in the long run. If market price is above minimum average total cost, existing firms are profitable and new firms will enter the industry in the long run.

5. The **industry supply curve** depends on the time period (short run or long run). When the number of firms is fixed, the **short-run industry supply curve** applies. The **short-run market equilibrium** occurs where the short-run industry supply curve and the demand curve intersect.

6. With sufficient time for entry into and exit from an industry, the **long-run industry supply curve** applies. The **long-run market equilibrium** occurs at the intersection of the long-run industry supply curve and the demand curve. At this point, no producer has an incentive to enter or exit. The long-run industry supply curve is often horizontal. It may slope upward if there is limited supply of an input, resulting in increasing costs across the industry. It may even slope downward, as in the case of decreasing costs across the industry. But the long-run industry supply curve is always more elastic than the short-run industry supply curve.
7. In the long-run market equilibrium of a competitive industry, profit maximization leads each firm to produce at the same marginal cost, which is equal to the market price. Free entry and exit means that each firm earns zero economic profit—producing the output corresponding to its minimum average total cost. So the total cost of production of an industry's output is minimized. The outcome is efficient because every consumer with willingness to pay greater than or equal to marginal cost gets the good.
8. The key difference between a monopoly and a perfectly competitive industry is that a single, perfectly competitive firm faces a horizontal demand curve but a monopolist faces a downward-sloping demand curve. This gives the monopolist market power, the ability to raise the market price by reducing output.
9. The marginal revenue of a monopolist is composed of a quantity effect (the price received from the additional unit) and a price effect (the reduction in the price at which all units are sold). Because of the price effect, a monopolist's marginal revenue is always less than the market price, and the marginal revenue curve lies below the demand curve.
10. At the monopolist's profit-maximizing output level, marginal cost equals marginal revenue, which is less than market price. At the perfectly competitive firm's profit-maximizing output level, marginal cost equals the market price. So in comparison to perfectly competitive industries, monopolies produce less, charge higher prices, and can earn profits in both the short run and the long run.
11. A monopoly creates deadweight losses by charging a price above marginal cost: the loss in consumer surplus exceeds the monopolist's profit. This makes monopolies a source of market failure and governments often make policies to prevent or end them.
12. Natural monopolies also cause deadweight losses. To limit these losses, governments sometimes impose **public ownership** and at other times impose **price regulation.** A price ceiling on a monopolist, as opposed to a perfectly competitive industry, need not cause shortages and can increase total surplus.
13. Not all monopolists are **single-price monopolists.** Monopolists, as well as oligopolists and monopolistic competitors, often engage in **price discrimination** to make higher profits, using various techniques to differentiate consumers based on their sensitivity to price and charging those with less elastic demand higher prices. A monopolist that achieves **perfect price discrimination** charges each consumer a price equal to his or her willingness to pay and captures the total surplus in the market. Although perfect price discrimination creates no inefficiency, it is practically impossible to implement.

Key Terms

Price-taking firm's optimal output rule, p. 585
Break-even price, p. 592
Shut-down price, p. 593
Short-run individual supply curve, p. 594
Industry supply curve, p. 599
Short-run industry supply curve, p. 600
Short-run market equilibrium, p. 601
Long-run market equilibrium, p. 602
Long-run industry supply curve, p. 603
Public ownership, p. 619
Price regulation, p. 619
Single-price monopolist, p. 624
Price discrimination, p. 624
Perfect price discrimination, p. 627

Problems

1. For each of the following, is the industry perfectly competitive? Referring to market share, standardization of the product, and/or free entry and exit, explain your answers.

   a. aspirin

   b. Alicia Keys concerts

   c. SUVs

2. Kate's Katering provides catered meals, and the catered meals industry is perfectly competitive. Kate's machinery costs $100 per day and is the only fixed input. Her variable cost consists of the wages paid to the cooks and the food ingredients. The variable cost per day associated with each level of output is given in the accompanying table.

Quantity of meals	*VC*
0	$0
10	200
20	300
30	480
40	700
50	1,000

   a. Calculate the total cost, the average variable cost, the average total cost, and the marginal cost for each quantity of output.

   b. What is the break-even price? What is the shut-down price?

   c. Suppose that the price at which Kate can sell catered meals is $21 per meal. In the short run, will Kate earn a profit? In the short run, should she produce or shut down?

   d. Suppose that the price at which Kate can sell catered meals is $17 per meal. In the short run, will Kate earn a profit? In the short run, should she produce or shut down?

   e. Suppose that the price at which Kate can sell catered meals is $13 per meal. In the short run, will Kate earn a profit? In the short run, should she produce or shut down?

3. Bob produces DVD movies for sale, which requires a building and a machine that copies the original movie onto a DVD. Bob rents a building for $30,000 per month and rents a machine for $20,000 a month. Those are his fixed costs. His variable costs per month are given in the accompanying table.

Quantity of DVDs	*VC*
0	$0
1,000	5,000
2,000	8,000
3,000	9,000
4,000	14,000
5,000	20,000
6,000	33,000
7,000	49,000
8,000	72,000
9,000	99,000
10,000	150,000

   a. Calculate Bob's average variable cost, average total cost, and marginal cost for each quantity of output.

   b. There is free entry into the industry, and anyone who enters will face the same costs as Bob. Suppose that currently the price of a DVD is $25. What will Bob's profit be? Is this a long-run equilibrium? If not, what will the price of DVD movies be in the long run?

4. Consider Bob's DVD company described in Problem 3. Assume that DVD production is a perfectly competitive industry. For each of the following questions, explain your answers.

   a. What is Bob's break-even price? What is his shut-down price?

   b. Suppose the price of a DVD is $2. What should Bob do in the short run?

   c. Suppose the price of a DVD is $7. What is the profit-maximizing quantity of DVDs that Bob should produce? What will his total profit be? Will he produce or shut down in the short run? Will he stay in the industry or exit in the long run?

   d. Suppose instead that the price of DVDs is $20. Now what is the profit-maximizing quantity of DVDs that Bob should produce? What will his total profit be now? Will he produce or shut down in the short run? Will he stay in the industry or exit in the long run?

5. Consider again Bob's DVD company described in Problem 3.

a. Draw Bob's marginal cost curve.

b. Over what range of prices will Bob produce no DVDs in the short run?

c. Draw Bob's individual supply curve.

6. a. A profit-maximizing business incurs an economic loss of $10,000 per year. Its fixed cost is $15,000 per year. Should it produce or shut down in the short run? Should it stay in the industry or exit in the long run?

b. Suppose instead that this business has a fixed cost of $6,000 per year. Should it produce or shut down in the short run? Should it stay in the industry or exit in the long run?

7. The first sushi restaurant opens in town. Initially, people are very cautious about eating tiny portions of raw fish, as this is a town where large portions of grilled meat have always been popular. Soon, however, an influential health report warns consumers against grilled meat and suggests that they increase their consumption of fish, especially raw fish. The sushi restaurant becomes very popular and its profit increases.

a. What will happen to the short-run profit of the sushi restaurant? What will happen to the number of sushi restaurants in town in the long run? Will the first sushi restaurant be able to sustain its short-run profit over the long run? Explain your answers.

b. Local steakhouses suffer from the popularity of sushi and start incurring losses. What will happen to the number of steakhouses in town in the long run? Explain your answer.

8. A perfectly competitive firm has the following short-run total costs:

Quantity	*TC*
0	$5
1	10
2	13
3	18
4	25
5	34
6	45

Market demand for the firm's product is given by the following market demand schedule:

Price	Quantity demanded
$12	300
10	500
8	800
6	1,200
4	1,800

a. Calculate this firm's marginal cost and, for all output levels except zero, the firm's average variable cost and average total cost.

b. There are 100 firms in this industry that all have costs identical to those of this firm. Draw the short-run industry supply curve. In the same diagram, draw the market demand curve.

c. What is the market price, and how much profit will each firm make?

9. A new vaccine against a deadly disease has just been discovered. Presently, 55 people die from the disease each year. The new vaccine will save lives, but it is not completely safe. Some recipients of the shots will die from adverse reactions. The projected effects of the inoculation are given in the accompanying table:

Percent of population inoculated	Total deaths due to disease	Total deaths due to inoculation	Marginal benefit of inoculation	Marginal cost of inoculation	"Profit" of inoculation
0	55	0			
			—	—	—
10	45	0			
			—	—	—
20	36	1			
			—	—	—
30	28	3			
			—	—	—
40	21	6			
			—	—	—
50	15	10			
			—	—	—
60	10	15			
			—	—	—
70	6	20			
			—	—	—
80	3	25			
			—	—	—
90	1	30			
			—	—	—
100	0	35			

a. What are the interpretations of "marginal benefit" and "marginal cost" here? Calculate marginal benefit and marginal cost per each 10% increase in the rate of inoculation. Write your answers in the table.

b. What proportion of the population should optimally be inoculated?

c. What is the interpretation of "profit" here? Calculate the profit for all levels of inoculation.

10. The production of agricultural products like wheat is one of the few examples of a perfectly competitive industry. In this question, we analyze results from a study released by the U.S. Department of Agriculture about wheat production in the United States in 1998 and make some comparisons to wheat production in 2010.

a. The average variable cost per acre planted with wheat was $107 per acre. Assuming a yield of 50 bushels per acre, calculate the average variable cost per bushel of wheat.

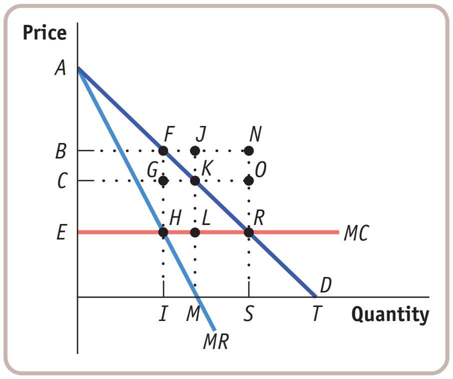
b. The average price of wheat received by a farmer in 1998 was $2.65 per bushel. Do you think the average farm would have exited the industry in the short run? Explain.

c. With a yield of 50 bushels of wheat per acre, the average total cost per farm was $3.80 per bushel. The harvested acreage for rye (a type of wheat) in the United States fell from 418,000 acres in 1998 to 250,000 in 2010. Using the information on prices and costs here and in parts a and b, explain why this might have happened.

d. Using the above information, do you think the prices of wheat were higher or lower prior to 1998? Why?

11. Skyscraper City has a subway system for which a one-way fare is $1.50. There is pressure on the mayor to reduce the fare by one-third, to $1.00. The mayor is dismayed, thinking that this will mean Skyscraper City is losing one-third of its revenue from sales of subway tickets. The mayor's economic adviser reminds her that she is focusing only on the price effect and ignoring the quantity effect. Explain why the mayor's estimate of a one-third loss of revenue is likely to be an overestimate. Illustrate with a diagram.

12. Consider an industry with the demand curve (*D*) and marginal cost curve (*MC*) shown in the accompanying diagram. There is no fixed cost. If the industry is a single-price monopoly, the monopolist's marginal revenue curve would be *MR*. Answer the following questions by naming the appropriate points or areas.



a. If the industry is perfectly competitive, what will be the total quantity produced? At what price?

b. Which area reflects consumer surplus under perfect competition?

c. If the industry is a single-price monopoly, what quantity will the monopolist produce? Which price will it charge?

d. Which area reflects the single-price monopolist's profit?

e. Which area reflects consumer surplus under single-price monopoly?

f. Which area reflects the deadweight loss to society from single-price monopoly?

g. If the monopolist can price-discriminate perfectly, what quantity will the perfectly price-discriminating monopolist produce?

13. Bob, Bill, Ben, and Brad Baxter have just made a documentary movie about their basketball team. They are thinking about making the movie available for download on the Internet, and they can act as a single-price monopolist if they choose to. Each time the movie is downloaded, their Internet service provider charges them a fee of $4. The Baxter brothers are arguing about which price to charge customers per download. The accompanying table shows the demand schedule for their film.

Price of download	Quantity of downloads demanded
$10	0
8	1
6	3
4	6
2	10
0	15

a. Calculate the total revenue and the marginal revenue per download.

b. Bob is proud of the film and wants as many people as possible to download it. Which price would he choose? How many downloads would be sold?

c. Bill wants as much total revenue as possible. Which price would he choose? How many downloads would be sold?

d. Ben wants to maximize profit. Which price would he choose? How many downloads would be sold?

e. Brad wants to charge the efficient price. Which price would he choose? How many downloads would be sold?

14. Suppose that De Beers is a single-price monopolist in the market for diamonds. De Beers has five potential customers: Raquel, Jackie, Joan, Mia, and Sophia. Each of these customers will buy at most one diamond—and only if the price is just equal to, or lower than, her willingness to pay. Raquel's willingness to pay is $400; Jackie's, $300; Joan's, $200; Mia's, $100; and Sophia's, $0. De Beers's marginal cost per diamond is $100. This leads to the demand schedule for diamonds shown in the accompanying table.

Price of diamond	Quantity of diamonds demanded
$500	0
400	1
300	2
200	3
100	4
0	5

a. Calculate De Beers's total revenue and its marginal revenue. From your calculation, draw the demand curve and the marginal revenue curve.

b. Explain why De Beers faces a downward-sloping demand curve.

c. Explain why the marginal revenue from an additional diamond sale is less than the price of the diamond.

d. Suppose De Beers currently charges $200 for its diamonds. If it lowers the price to $100, how large is the price effect? How large is the quantity effect?

e. Add the marginal cost curve to your diagram from part a, and determine which quantity maximizes the company's profit and which price De Beers will charge.

15. Use the demand schedule for diamonds given in Problem 14. The marginal cost of producing diamonds is constant at $100. There is no fixed cost.

a. If De Beers charges the monopoly price, how large is the individual consumer surplus that each buyer experiences? Calculate total consumer surplus by summing the individual consumer surpluses. How large is producer surplus?

Suppose that upstart Russian and Asian producers enter the market and the market becomes perfectly competitive.

b. What is the perfectly competitive price? What quantity will be sold in this perfectly competitive market?

c. At the competitive price and quantity, how large is the consumer surplus that each buyer experiences? How large is total consumer surplus? How large is producer surplus?

d. Compare your answer to part c to your answer to part a. How large is the deadweight loss associated with monopoly in this case?

16. Use the demand schedule for diamonds given in Problem 14. De Beers is a monopolist, but it can now price-discriminate perfectly among all five of its potential customers. De Beers's marginal cost is constant at $100. There is no fixed cost.

a. If De Beers can price-discriminate perfectly, to which customers will it sell diamonds and at what prices?

b. How large is each individual consumer surplus? How large is total consumer surplus? Calculate producer surplus by summing the producer surplus generated by each sale.

17. Download Records decides to release an album by the group Mary and the Little Lamb. It produces the album with no fixed cost, but the total cost of downloading an album to a CD and paying Mary her royalty is $6 per album. Download Records can act as a single-price monopolist. Its marketing division finds that the demand schedule for the album is as shown in the accompanying table.

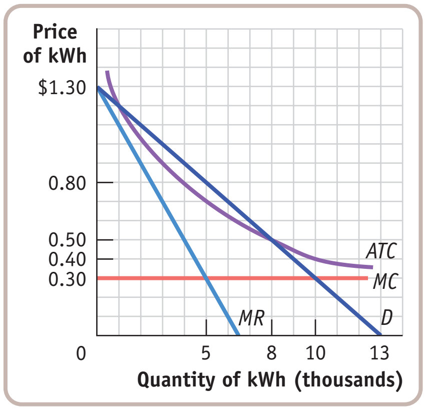
Price of album	Quantity of albums demanded
$22	0
20	1,000
18	2,000
16	3,000
14	4,000
12	5,000
10	6,000
8	7,000

a. Calculate the total revenue and the marginal revenue per album.

b. The marginal cost of producing each album is constant at $6. To maximize profit, what level of output should Download Records choose, and which price should it charge for each album?

c. Mary renegotiates her contract and now needs to be paid a higher royalty per album. So the marginal cost rises to be constant at $14. To maximize profit, what level of output should Download Records now choose, and which price should it charge for each album?

18. The accompanying diagram illustrates your local electricity company's natural monopoly. The diagram shows the demand curve for kilowatt-hours (kWh) of electricity, the company's marginal revenue (*MR*) curve, its marginal cost (*MC*) curve, and its average total cost (*ATC*) curve. The government wants to regulate the monopolist by imposing a price ceiling.



a. If the government does not regulate this monopolist, which price will it charge? Illustrate the inefficiency this creates by shading the deadweight loss from monopoly.

b. If the government imposes a price ceiling equal to the marginal cost, $0.30, will the monopolist make a profit or lose money? Shade the area of profit (or loss) for the monopolist. If the government does impose this price ceiling, do you think the firm will continue to produce in the long run?

c. If the government imposes a price ceiling of $0.50, will the monopolist make a profit, lose money, or break even?

19. The movie theater in Collegetown serves two kinds of customers: students and professors. There are 900 students and 100 professors in Collegetown. Each student's willingness to pay for a movie ticket is $5. Each professor's willingness to pay for a movie ticket is $10. Each will buy at most one ticket. The movie theater's marginal cost per ticket is constant at $3, and there is no fixed cost.

a. Suppose the movie theater cannot price-discriminate and needs to charge both students and professors the same price per ticket. If the movie theater charges $5, who will buy tickets and what will the movie theater's profit be? How large is consumer surplus?

b. If the movie theater charges $10, who will buy movie tickets and what will the movie theater's profit be? How large is consumer surplus?

c. Now suppose that, if it chooses to, the movie theater can price-discriminate between students and professors by requiring students to show their student ID. If the movie theater charges students $5 and professors $10, how much profit will the movie theater make? How large is consumer surplus?

20. A monopolist knows that in order to expand the quantity of output it produces from 8 to 9 units, it must lower the price of its output from $2 to $1. Calculate the quantity effect and the price effect. Use these results to calculate the monopolist's marginal revenue of producing the 9th unit. The marginal cost of producing the 9th unit is positive. Is it a good idea for the monopolist to produce the 9th unit?