When a firm is a natural monopoly, it is less concerned about new entrants eroding its monopoly power. Normally, a firm has trouble maintaining a monopoly position without ownership of a key resource or protection from the government. The monopolist's profit attracts entrants into the market, and these entrants make the market more competitive. By contrast, entering a market in which another firm has a natural monopoly is unattractive. Would-be entrants know that they cannot achieve the same low costs that the monopolist enjoys because, after entry, each firm would have a smaller piece of the market.

In some cases, the size of the market is one determinant of whether an industry is a natural monopoly. Again, consider a bridge across a river. When the population is small, the bridge may be a natural monopoly. A single bridge can satisfy the entire demand for trips across the river at lowest cost. Yet as the population grows and the bridge becomes congested, satisfying the entire demand may require two or more bridges across the same river. Thus, as a market expands, a natural monopoly can evolve into a more competitive market.

QUICK QUIZ   What are the three reasons that a market might have a monopoly? • Give two examples of monopolies and explain the reason for each.

HOW MONOPOLIES MAKE PRODUCTION AND PRICING DECISIONS

Now that we know how monopolies arise, we can consider how a monopoly firm decides how much of its product to make and what price to charge for it. The analysis of monopoly behavior in this section is the starting point for evaluating whether monopolies are desirable and what policies the government might pursue in monopoly markets.

MONOPOLY VERSUS COMPETITION

The key difference between a competitive firm and a monopoly is the monopoly's ability to influence the price of its output. A competitive firm is small relative to the market in which it operates and, therefore, has no power to influence the price of its output. It takes the price as given by market conditions. By contrast, because a monopoly is the sole producer in its market, it can alter the price of its good by adjusting the quantity it supplies to the market.

One way to view this difference between a competitive firm and a monopoly is to consider the demand curve that each firm faces. When we analyzed profit maximization by competitive firms in the preceding chapter, we drew the market price as a horizontal line. Because a competitive firm can sell as much or as little as it wants at this price, the competitive firm faces a horizontal demand curve, as in panel (a) of Figure 2. In effect, because the competitive firm sells a product with many perfect substitutes (the products of all the other firms in its market), the demand curve that any one firm faces is perfectly elastic.

By contrast, because a monopoly is the sole producer in its market, its demand curve is the market demand curve. Thus, the monopolist's demand curve slopes downward for all the usual reasons, as in panel (b) of Figure 2. If the monopolist raises the price of its good, consumers buy less of it. Looked at another way, if the

2 FIGURE

Demand Curves for Competitive and Monopoly Firms

Because competitive firms are price takers, they in effect face horizontal demand curves, as in panel (a). Because a monopoly firm is the sole producer in its market, it faces the downward-sloping market demand curve, as in panel (b). As a result, the monopoly has to accept a lower price if it wants to sell more output.

(a) A Competitive Firm's Demand Curve

(b) A Monopolist's Demand Curve
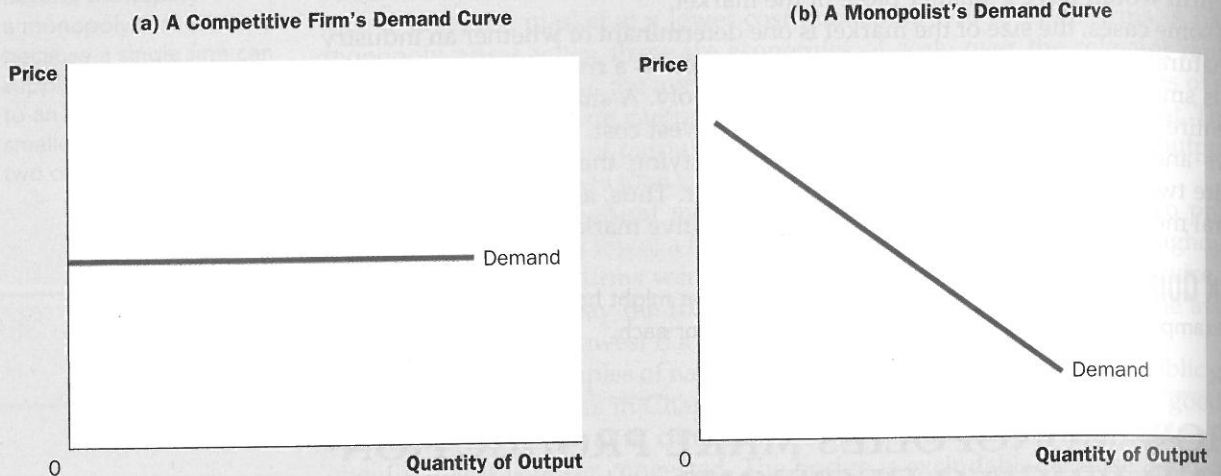
monopolist reduces the quantity of output it produces and sells, the price of its output increases.

The market demand curve provides a constraint on a monopoly's ability to profit from its market power. A monopolist would prefer, if it were possible, to charge a high price and sell a large quantity at that high price. The market demand curve makes that outcome impossible. In particular, the market demand curve describes the combinations of price and quantity that are available to a monopoly firm. By adjusting the quantity produced (or equivalently, the price charged), the monopolist can choose any point on the demand curve, but it cannot choose a point off the demand curve.

What price and quantity of output will the monopolist choose? As with competitive firms, we assume that the monopolist's goal is to maximize profit. Because the firm's profit is total revenue minus total costs, our next task in explaining monopoly behavior is to examine a monopolist's revenue.

A MONOPOLY'S REVENUE

Consider a town with a single producer of water. Table 1 shows how the monopoly's revenue might depend on the amount of water produced.

The first two columns show the monopolist's demand schedule. If the monopolist produces 1 gallon of water, it can sell that gallon for $10. If it produces 2 gallons, it must lower the price to $9 to sell both gallons. If it produces 3 gallons, it must lower the price to $8. And so on. If you graphed these two columns of numbers, you would get a typical downward-sloping demand curve.

Quantity of Water (Q)	Price (P)	Total Revenue (TR = P × Q)	Average Revenue (AR = TR / Q)	Marginal Revenue (MR = ΔTR / ΔQ)
0 gallons	$11	$ 0	—	
				$10
1	10	10	$10	
				8
2	9	18	9	
				6
3	8	24	8	
				4
4	7	28	7	
				2
5	6	30	6	
				0
6	5	30	5	
				−2
7	4	28	4	
				−4
8	3	24	3	

The third column of the table presents the monopolist's *total revenue*. It equals the quantity sold (from the first column) times the price (from the second column). The fourth column computes the firm's *average revenue*, the amount of revenue the firm receives per unit sold. We compute average revenue by taking the number for total revenue in the third column and dividing it by the quantity of output in the first column. As we discussed in the previous chapter, average revenue always equals the price of the good. This is true for monopolists as well as for competitive firms.

The last column of Table 1 computes the firm's *marginal revenue*, the amount of revenue that the firm receives for each additional unit of output. We compute marginal revenue by taking the change in total revenue when output increases by 1 unit. For example, when the firm is producing 3 gallons of water, it receives total revenue of $24. Raising production to 4 gallons increases total revenue to $28. Thus, marginal revenue from the sale of the fourth gallon is $28 minus $24, or $4.

Table 1 shows a result that is important for understanding monopoly behavior: *A monopolist's marginal revenue is always less than the price of its good.* For example, if the firm raises production of water from 3 to 4 gallons, it will increase total revenue by only $4, even though it will be able to sell each gallon for $7. For a monopoly, marginal revenue is lower than price because a monopoly faces a downward-sloping demand curve. To increase the amount sold, a monopoly firm must lower the price it charges to all customers. Hence, to sell the fourth gallon of water, the monopolist will get $1 less revenue for each of the first three gallons. This $3 loss accounts for the difference between the price of the fourth gallon ($7) and the marginal revenue of that fourth gallon ($4).

Marginal revenue for monopolies is very different from marginal revenue for competitive firms. When a monopoly increases the amount it sells, this action has two effects on total revenue ($P \times Q$):

- *The output effect:* More output is sold, so Q is higher, which tends to increase total revenue.
- *The price effect:* The price falls, so P is lower, which tends to decrease total revenue.

Because a competitive firm can sell all it wants at the market price, there is no price effect. When it increases production by 1 unit, it receives the market price for that unit, and it does not receive any less for the units it was already selling. That is, because the competitive firm is a price taker, its marginal revenue equals the price of its good. By contrast, when a monopoly increases production by 1 unit, it must reduce the price it charges for every unit it sells, and this cut in price reduces revenue on the units it was already selling. As a result, a monopoly's marginal revenue is less than its price.

Figure 3 graphs the demand curve and the marginal-revenue curve for a monopoly firm. (Because the firm's price equals its average revenue, the demand curve is also the average-revenue curve.) These two curves always start at the same point on the vertical axis because the marginal revenue of the first unit sold equals the price of the good. But for the reason we just discussed, the monopolist's marginal revenue on all units after the first is less than the price of the good. Thus, a monopoly's marginal-revenue curve lies below its demand curve.

You can see in the figure (as well as in Table 1) that marginal revenue can even become negative. Marginal revenue is negative when the price effect on revenue is greater than the output effect. In this case, when the firm produces an extra unit of output, the price falls by enough to cause the firm's total revenue to decline, even though the firm is selling more units.

3 **FIGURE**

Demand and Marginal-Revenue Curves for a Monopoly
The demand curve shows how the quantity affects the price of the good. The marginal-revenue curve shows how the firm's revenue changes when the quantity increases by 1 unit. Because the price on *all* units sold must fall if the monopoly increases production, marginal revenue is always less than the price.
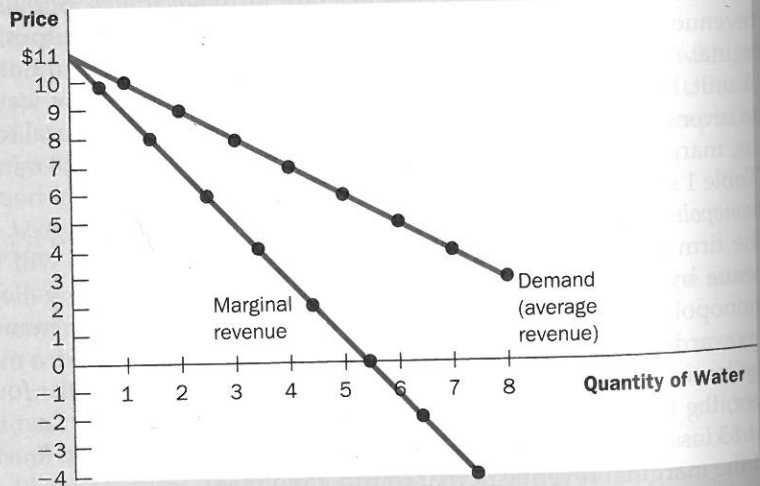
PROFIT MAXIMIZATION

Now that we have considered the revenue of a monopoly firm, we are ready to examine how such a firm maximizes profit. Recall from Chapter 1 that one of the *Ten Principles of Economics* is that rational people think at the margin. This lesson is as true for monopolists as it is for competitive firms. Here we apply the logic of marginal analysis to the monopolist's decision about how much to produce.

Figure 4 graphs the demand curve, the marginal-revenue curve, and the cost curves for a monopoly firm. All these curves should seem familiar: The demand and marginal-revenue curves are like those in Figure 3, and the cost curves are like those we encountered in the last two chapters. These curves contain all the information we need to determine the level of output that a profit-maximizing monopolist will choose.

Suppose, first, that the firm is producing at a low level of output, such as Q_1. In this case, marginal cost is less than marginal revenue. If the firm increased production by 1 unit, the additional revenue would exceed the additional costs, and profit would rise. Thus, when marginal cost is less than marginal revenue, the firm can increase profit by producing more units.

A similar argument applies at high levels of output, such as Q_2. In this case, marginal cost is greater than marginal revenue. If the firm reduced production by 1 unit, the costs saved would exceed the revenue lost. Thus, if marginal cost is greater than marginal revenue, the firm can raise profit by reducing production.

In the end, the firm adjusts its level of production until the quantity reaches Q_{MAX}, at which marginal revenue equals marginal cost. *Thus, the monopolist's profit-maximizing quantity of output is determined by the intersection of the marginal-revenue curve and the marginal-cost curve.* In Figure 4, this intersection occurs at point A.

You might recall from the previous chapter that competitive firms also choose the quantity of output at which marginal revenue equals marginal cost. In
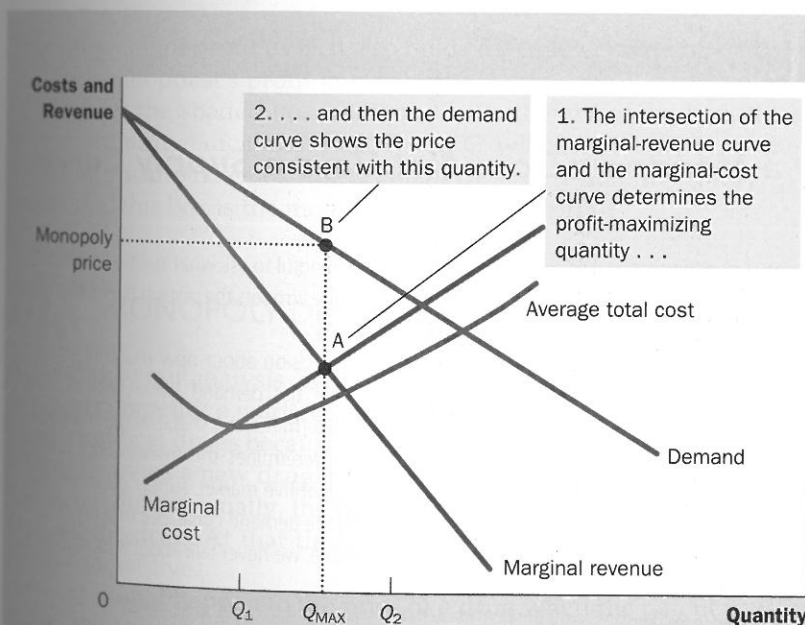

FIGURE **4**

Profit Maximization for a Monopoly

A monopoly maximizes profit by choosing the quantity at which marginal revenue equals marginal cost (point A). It then uses the demand curve to find the price that will induce consumers to buy that quantity (point B).

following this rule for profit maximization, competitive firms and monopolies are alike. But there is also an important difference between these types of firms: The marginal revenue of a competitive firm equals its price, whereas the marginal revenue of a monopoly is less than its price. That is,

For a competitive firm: $P = MR = MC$.
For a monopoly firm: $P > MR = MC$.

The equality of marginal revenue and marginal cost at the profit-maximizing quantity is the same for both types of firms. What differs is the relationship of the price to marginal revenue and marginal cost.

How does the monopoly find the profit-maximizing price for its product? The demand curve answers this question because the demand curve relates the amount that customers are willing to pay to the quantity sold. Thus, after the monopoly firm chooses the quantity of output that equates marginal revenue and marginal cost, it uses the demand curve to find the highest price it can charge and sell that quantity. In Figure 4, the profit-maximizing price is found at point B.

We can now see a key difference between markets with competitive firms and markets with a monopoly firm: *In competitive markets, price equals marginal cost. In monopolized markets, price exceeds marginal cost.* As we will see in a moment, this finding is crucial to understanding the social cost of monopoly.

A MONOPOLY'S PROFIT

How much profit does a monopoly make? To see a monopoly firm's profit in a graph, recall that profit equals total revenue (*TR*) minus total costs (*TC*):

$$\text{Profit} = TR - TC.$$

### FYI

Why a Monopoly Does Not Have a Supply Curve

You may have noticed that we have analyzed the price in a monopoly market using the market demand curve and the firm's cost curves. We have not made any mention of the market supply curve. By contrast, when we analyzed prices in competitive markets beginning in Chapter 4, the two most important words were always *supply* and *demand*.

What happened to the supply curve? Although monopoly firms make decisions about what quantity to supply (in the way described in this chapter), a monopoly does not have a supply curve. A supply curve tells us the quantity that firms choose to supply at any given price. This concept makes sense when we are analyzing competitive firms, which are price takers. But a monopoly firm is a price maker,

not a price taker. It is not meaningful to ask what such a firm would produce at any price because the firm sets the price at the same time it chooses the quantity to supply.

Indeed, the monopolist's decision about how much to supply is impossible to separate from the demand curve it faces. The shape of the demand curve determines the shape of the marginal-revenue curve, which in turn determines the monopolist's profit-maximizing quantity. In a competitive market, supply decisions can be analyzed without knowing the demand curve, but that is not true in a monopoly market. Therefore, we never talk about a monopoly's supply curve.

FIGURE 5
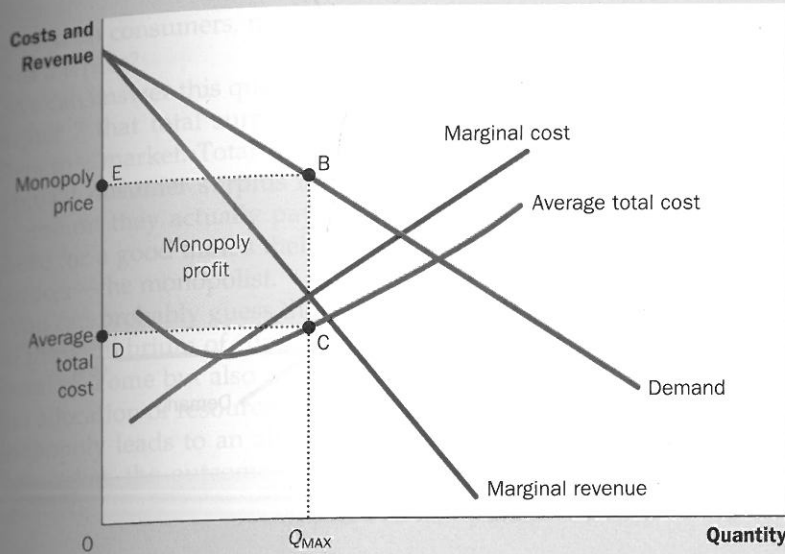
The Monopolist's Profit
The area of the box BCDE equals the profit of the monopoly firm. The height of the box (BC) is price minus average total cost, which equals profit per unit sold. The width of the box (DC) is the number of units sold.

We can rewrite this as

$$\text{Profit} = (TR/Q - TC/Q) \times Q.$$

TR/Q is average revenue, which equals the price, P, and TC/Q is average total cost, ATC. Therefore,

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

This equation for profit (which also holds for competitive firms) allows us to measure the monopolist's profit in our graph.

Consider the shaded box in Figure 5. The height of the box (the segment BC) is price minus average total cost, $P - ATC$, which is the profit on the typical unit sold. The width of the box (the segment DC) is the quantity sold, Q_{MAX}. Therefore, the area of this box is the monopoly firm's total profit.

## CASE STUDY MONOPOLY DRUGS VERSUS GENERIC DRUGS

According to our analysis, prices are determined differently in monopolized markets and competitive markets. A natural place to test this theory is the market for pharmaceutical drugs because this market takes on both market structures. When a firm discovers a new drug, patent laws give the firm a monopoly on the sale of that drug. But eventually, the firm's patent runs out, and any company can make and sell the drug. At that time, the market switches from being monopolistic to being competitive.

What should happen to the price of a drug when the patent runs out? Figure 6 shows the market for a typical drug. In this figure, the marginal cost of producing

6 FIGURE

The Market for Drugs
When a patent gives a firm a monopoly over the sale of a drug, the firm charges the monopoly price, which is well above the marginal cost of making the drug. When the patent on a drug runs out, new firms enter the market, making it more competitive. As a result, the price falls from the monopoly price to marginal cost.
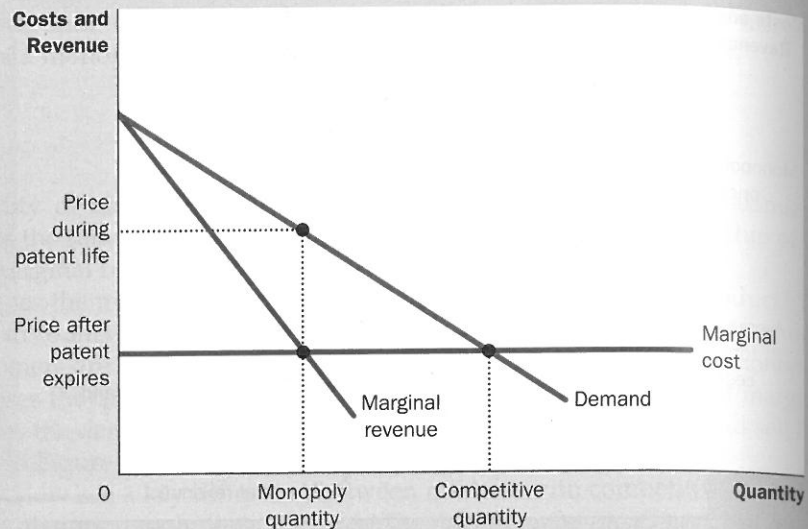
the drug is constant. (This is approximately true for many drugs.) During the life of the patent, the monopoly firm maximizes profit by producing the quantity at which marginal revenue equals marginal cost and charging a price well above marginal cost. But when the patent runs out, the profit from making the drug should encourage new firms to enter the market. As the market becomes more competitive, the price should fall to equal marginal cost.

Experience is, in fact, consistent with our theory. When the patent on a drug expires, other companies quickly enter and begin selling so-called generic products that are chemically identical to the former monopolist's brand-name product. And just as our analysis predicts, the price of the competitively produced generic drug is well below the price that the monopolist was charging.

The expiration of a patent, however, does not cause the monopolist to lose all its market power. Some consumers remain loyal to the brand-name drug, perhaps out of fear that the new generic drugs are not actually the same as the drug they have been using for years. As a result, the former monopolist can continue to charge a price at least somewhat above the price charged by its new competitors. ●

QUICK QUIZ Explain how a monopolist chooses the quantity of output to produce and the price to charge.

THE WELFARE COST OF MONOPOLIES

Is monopoly a good way to organize a market? We have seen that a monopoly, in contrast to a competitive firm, charges a price above marginal cost. From the standpoint of consumers, this high price makes monopoly undesirable. At the same time, however, the monopoly is earning profit from charging this high price.