

WHY MONOPOLIES ARISE

- A firm is a **monopoly** if it is the sole seller of its product and if its product does not have close substitutes. **The fundamental cause of monopoly is *barriers to entry*:** A monopoly remains the only seller in its market because other firms cannot enter the market and compete with it.
- Barriers to entry, in turn, have three main sources:
 - *Monopoly resources*: A key resource required for production is owned by a single firm.
 - *Government regulation*: The government gives a single firm the exclusive right to produce some good or service.
 - *The production process*: A single firm can produce output at a lower cost than can a larger number of producers.

MONOPOLY RESOURCES

- The simplest way for a monopoly to arise is for a single firm to own a key resource.
- For example, consider the market for water in a small town. If dozens of town residents have working wells, the competitive model discussed in the preceding chapter describes the behavior of sellers. As a result of the competition among water suppliers, the price of a gallon is driven to equal the marginal cost of pumping an extra gallon. But if there is only one well in town and it is impossible to get water from anywhere else, then the owner of the well has a monopoly on water. Not surprisingly, the monopolist has much greater market power than any single firm in a competitive market. In

the case of a necessity like water, the monopolist could command quite a high price, even if the marginal cost of pumping an extra gallon is low.

- A classic example of market power arising from the ownership of a key resource is DeBeers, the South African diamond company. Founded in 1888 by Cecil Rhodes, an English businessman (and benefactor for the Rhodes scholarship), DeBeers has at times controlled up to 80 percent of the production from the world's diamond mines. Because its market share is less than 100 percent, DeBeers is not exactly a monopoly, but the company has nonetheless exerted substantial influence over the market price of diamonds.

GOVERNMENT-CREATED MONOPOLIES

- In many cases, monopolies arise because the government has given one person or firm the exclusive right to sell some good or service. Sometimes the monopoly arises from the sheer political clout of the would-be monopolist.
- The **patent and copyright** laws are two important examples. When a pharmaceutical company discovers a new drug, it can apply to the government for a patent.
- If the government deems the drug to be truly original, it approves the patent, which gives the company the exclusive right to manufacture and sell the drug for 20 years. Similarly, when a novelist finishes a book, she can copyright it. The copyright is a government guarantee that no one can print and sell the work without the author's permission. The copyright makes the novelist a monopolist in the sale of her novel.

- The effects of patent and copyright laws are easy to see. Because these laws give one producer a monopoly, they lead to higher prices than would occur under competition. But by allowing these monopoly producers to charge higher prices and earn higher profits, the laws also encourage some desirable behavior. Drug companies are allowed to be monopolists in the drugs they discover to encourage research. Authors are allowed to be monopolists in the sale of their books to encourage them to write more and better books.

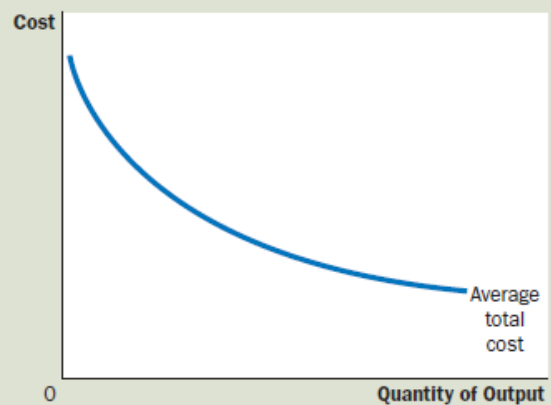
NATURAL MONOPOLIES

- An industry is a natural monopoly when a single firm can supply a good or service to an entire market at a lower cost than could two or more firms.
- A natural monopoly arises when there are economies of scale over the relevant range of output.

1 FIGURE

Economies of Scale as a Cause of Monopoly

When a firm's average-total-cost curve continually declines, the firm has what is called a natural monopoly. In this case, when production is divided among more firms, each firm produces less, and average total cost rises. As a result, a single firm can produce any given amount at the smallest cost.



- Figure 1 shows the average total costs of a firm with economies of scale.

- In this case, a single firm can produce any amount of output at least cost. That is, for any given amount of output, a larger number of firms lead to less output per firm and higher average total cost.
- An example of a natural monopoly is the distribution of water. To provide water to residents of a town, a firm must build a network of pipes throughout the town. If two or more firms were to compete in the provision of this service, each firm would have to pay the fixed cost of building a network. Thus, the average total cost of water is lowest if a single firm serves the entire market.
- 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.

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.

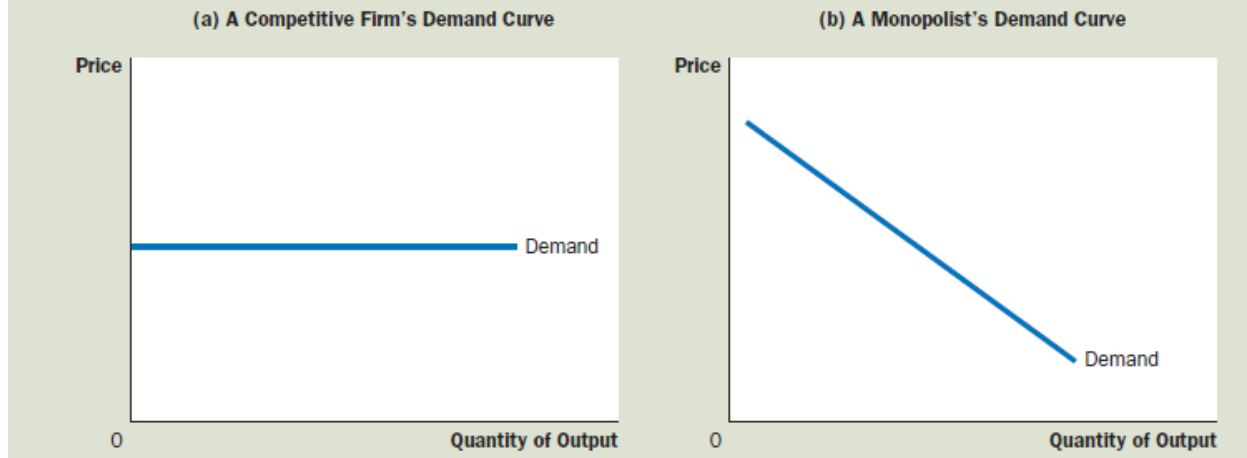
MONOPOLY VERSUS COMPETITION

- 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.

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.



- 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 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.

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.

TABLE 1					A Monopoly's Total, Average, and Marginal Revenue
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	—		
1	10	10	\$10	\$10	
2	9	18	9	8	
3	8	24	8	6	
4	7	28	7	4	
5	6	30	6	2	
6	5	30	5	0	
7	4	28	4	−2	
8	3	24	3	−4	

- 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. 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

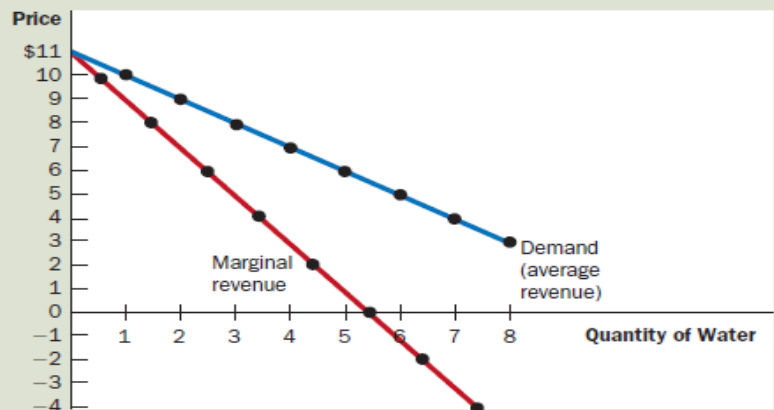
competitive firms. When a monopoly increases the amount it sells, this action has two effects on total revenue ($P \times Q$):

- 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, this 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.)

3 FIGURE

Demand and Marginal-Revenue Curves for a Monopoly

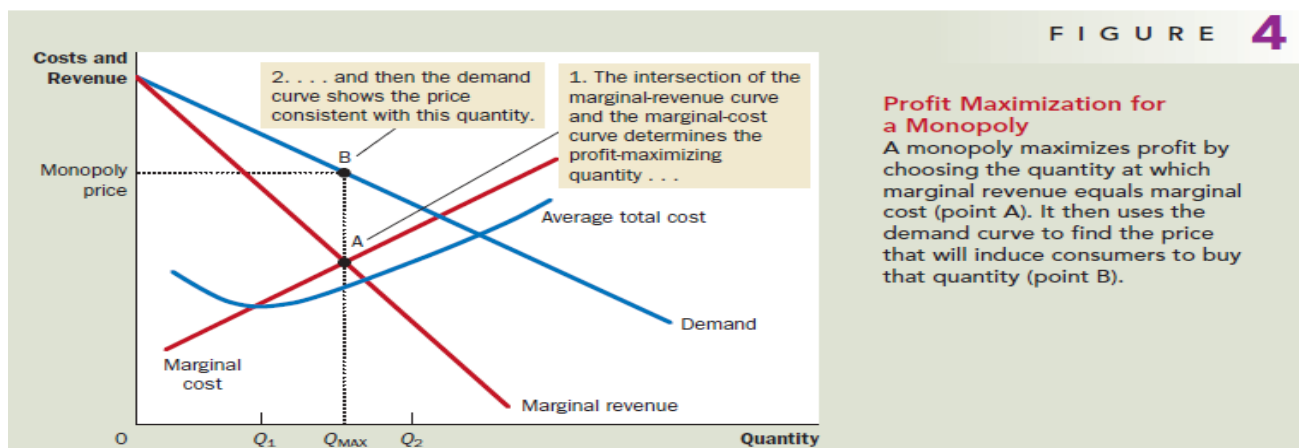
The demand curve shows how the quantity affects the price of the good. The marginal-revenue curve shows how the firm's revenue changes when the quantity increases by 1 unit. Because the price on *all* units sold must fall if the monopoly increases production, marginal revenue is always less than the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



- 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. Quantity

PROFIT MAXIMIZATION

- 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.



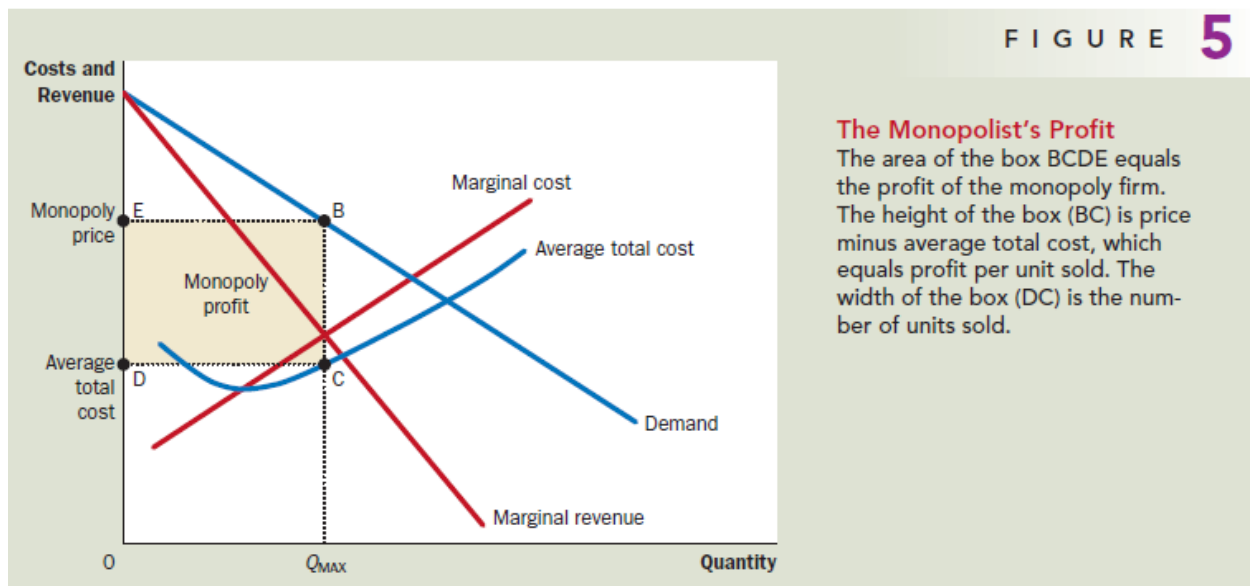
- Suppose, first, that the firm is producing at a low level of output, such as Q_1 . In this case, marginal cost is less than marginal revenue. If the firm increased production by 1 unit, the additional revenue would exceed the additional costs, and profit would rise. Thus, when marginal cost is less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

A MONOPOLY'S PROFIT



- To see a monopoly firm's profit through the above graph, recall that profit equals total revenue (TR) minus total costs (TC):

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

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

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

- Consider the shaded box in Figure 5. The height of the box (the segment BC) is price minus average total cost, $(P - \text{ATC})$, 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.

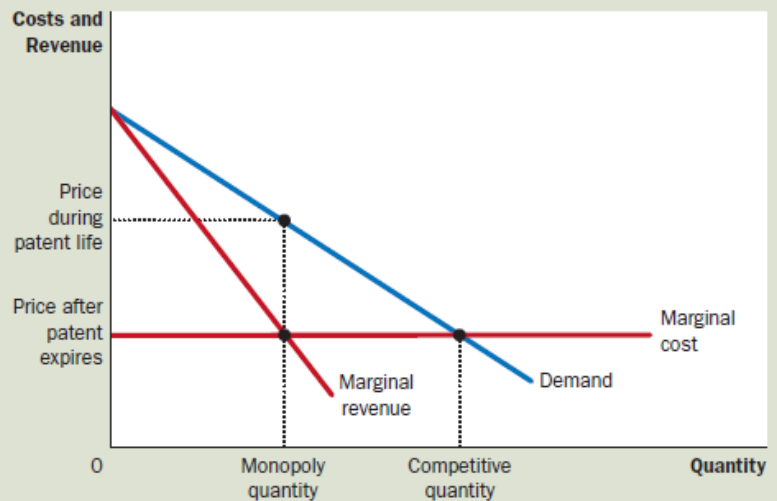
MONOPOLY DRUGS VERSUS GENERIC DRUGS

- 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.

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.



- In this figure, the marginal cost of producing 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.
- 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.

THE WELFARE COST OF MONOPOLIES

- Is monopoly a good way to organize a market?
- We can answer this question using the tools of welfare economics. The total surplus measures the economic well-being of buyers and sellers in a market.
- Total surplus is the sum of consumer surplus and producer surplus. Consumer surplus is consumers' willingness to pay for a good minus the amount they actually pay for it. Producer surplus is the amount producers receive for a good minus their costs of producing it. In this case, there is a single producer—the monopolist.
- The invisible hand of the market leads to an allocation of resources that makes total surplus as large as it can be. Because a monopoly leads to an allocation of resources different from that in a competitive market, the outcome must, in some way, fail to maximize total economic well-being.

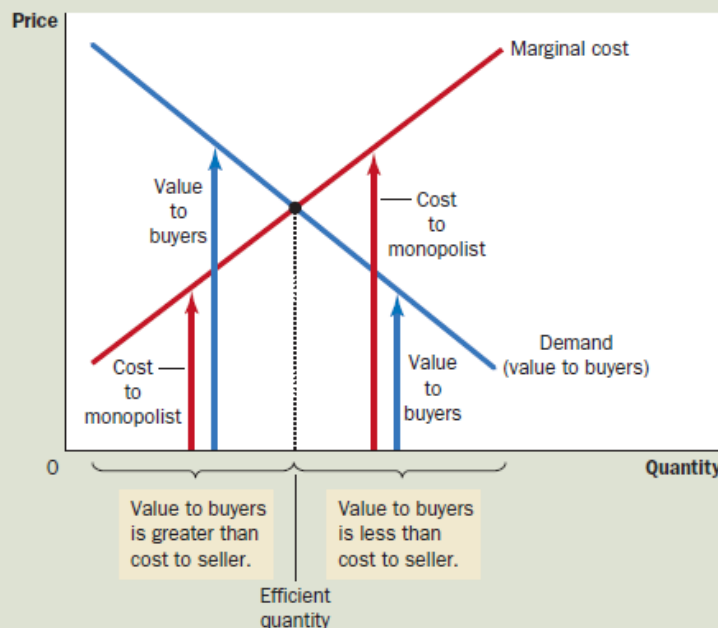
THE DEADWEIGHT LOSS

- We begin by considering what the monopoly firm would do if it were run by a benevolent social planner.
- The social planner cares not only about the profit earned by the firm's owners but also about the benefits received by the firm's consumers. The planner tries to maximize total surplus, which equals producer surplus (profit) plus consumer surplus. Keep in mind that total surplus equals the value of the good to consumers minus the costs of making the good incurred by the monopoly producer.

7 FIGURE

The Efficient Level of Output

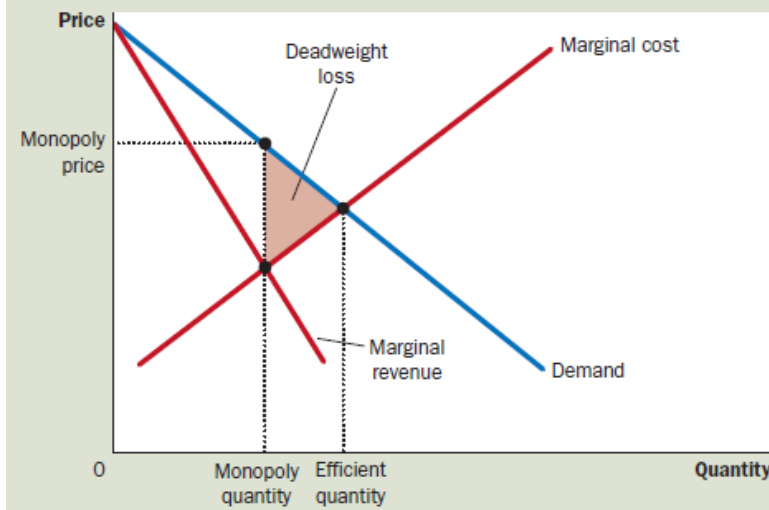
A benevolent social planner who wanted to maximize total surplus in the market would choose the level of output where the demand curve and marginal-cost curve intersect. Below this level, the value of the good to the marginal buyer (as reflected in the demand curve) exceeds the marginal cost of making the good. Above this level, the value to the marginal buyer is less than marginal cost.



- Figure 7 analyzes how a benevolent social planner would choose the monopoly's level of output.
- The demand curve reflects the value of the good to consumers, as measured by their willingness to pay for it. The marginal-cost curve reflects the costs of the monopolist. Thus, the socially efficient quantity is found where the demand curve and the marginal-cost curve intersect.
- Below this quantity, the value of an extra unit to consumers exceeds the cost of providing it, so increasing output would raise total surplus. Above this quantity, the cost of producing an extra unit exceeds the value of that unit to consumers, so decreasing output would raise total surplus. At the optimal quantity, the value of an extra unit to consumers exactly equals the marginal cost of production.

- If the social planner were running the monopoly, the firm could achieve this efficient outcome by charging the price found at the intersection of the demand and marginal-cost curves. Thus, like a competitive firm and unlike a profit maximizing monopoly, a social planner would charge a price equal to marginal cost. Because this price would give consumers an accurate signal about the cost of producing the good, consumers would buy the efficient quantity.
- We can evaluate the welfare effects of monopoly by comparing the level of output that the monopolist chooses to the level of output that a social planner would choose. As we have seen, the monopolist chooses to produce and sell the quantity of output at which the marginal-revenue and marginal-cost curves intersect; the social planner would choose the quantity at which the demand and marginal-cost curves intersect.
- **We can also view the inefficiency of monopoly in terms of the monopolist's price.** Because the market demand curve describes a negative relationship between the price and quantity of the good, a quantity that is inefficiently low is equivalent to a price that is inefficiently high. When a monopolist charges a price above marginal cost, some potential consumers value the good at more than its marginal cost but less than the monopolist's price. These consumers do not buy the good.
- Because the value these consumers place on the good is greater than the cost of providing it to them, this result is inefficient. Thus, monopoly pricing prevents some mutually beneficial trades from taking place.
- The inefficiency of monopoly can be measured with a deadweight loss triangle, as illustrated in Figure 8.

FIGURE 8



The Inefficiency of Monopoly

Because a monopoly charges a price above marginal cost, not all consumers who value the good at more than its cost buy it. Thus, the quantity produced and sold by a monopoly is below the socially efficient level. The deadweight loss is represented by the area of the triangle between the demand curve (which reflects the value of the good to consumers) and the marginal-cost curve (which reflects the costs of the monopoly producer).

- Because the demand curve reflects the value to consumers and the marginal-cost curve reflects the costs to the monopoly producer, the area of the deadweight loss triangle between the demand curve and the marginal-cost curve equals the total surplus lost because of monopoly pricing.
- It is the reduction in economic well-being that results from the monopoly's use of its market power.
- The deadweight loss caused by monopoly is similar to the deadweight loss caused by a tax. Indeed, a monopolist is like a private tax collector. As we know, a tax on a good places a wedge between consumers' willingness to pay (as reflected in the demand curve) and producers' costs (as reflected in the supply curve). Because a monopoly exerts its market power by charging a price above marginal cost, it places a similar wedge.
- In both cases, the wedge causes the quantity sold to fall short of the social optimum. The difference between the two cases is that the government gets the revenue from a tax, whereas a private firm gets the monopoly profit.

THE MONOPOLY'S PROFIT: A SOCIAL COST?

- Welfare in a monopolized market, like all markets, includes the welfare of both consumers and producers. Whenever a consumer pays an extra dollar to a producer because of a monopoly price, the consumer is worse off by a dollar, and the producer is better off by the same amount. This transfer from the consumers of the good to the owners of the monopoly does not affect the market's total surplus—the sum of consumer and producer surplus.
- In other words, the monopoly profit itself represents not a reduction in the size of the economic pie but merely a bigger slice for producers and a smaller slice for consumers. Unless consumers are for some reason more deserving than producers—a normative judgment about equity that goes beyond the realm of economic efficiency—the monopoly profit is not a social problem.
- The problem in a monopolized market arises because the firm produces and sells a quantity of output below the level that maximizes total surplus. The deadweight loss measures how much the economic pie shrinks as a result.
- This inefficiency is connected to the monopoly's high price: Consumers buy fewer units when the firm raises its price above marginal cost. But keep in mind that the profit earned on the units that continue to be sold is not the problem.
- The problem stems from the inefficiently low quantity of output. Put differently, if the high monopoly price did not discourage some consumers from buying the good, it would raise producer surplus by exactly the amount

it reduced consumer surplus, leaving total surplus the same as could be achieved by a benevolent social planner.

- There is, however, a possible exception to this conclusion. Suppose that a monopoly firm has to incur additional costs to maintain its monopoly position.
- For example, a firm with a government-created monopoly might need to hire lobbyists to convince lawmakers to continue its monopoly. In this case, the monopoly may use up some of its monopoly profits paying for these additional costs.
- If so, the social loss from monopoly includes both these costs and the deadweight loss resulting from a price above marginal cost.