# Chapter 6
# Monopoly and Monopolistic Competition*

In this chapter, we discuss the market structures of monopoly and monopolistic competition. Unlike perfect competition which has many sellers, a monopoly market has just one seller. In this sense, it is the polar opposite of perfect competition.

Monopolistic competition has qualities of both perfect competition and monopoly. Like perfect competition, it has many sellers. Unlike perfect competition, products are not perfect substitutes. Instead, each firm sells a substitute product that has its own unique set of characteristics, which might differ slightly in quality, style, and color. In this case, products are said to be differentiated. In monopolistic competition, firms sell brands that are unique, giving each firm a monopoly over the sale of its particular brand.

Although monopolistic competition is a form of imperfect competition, each firm is so small that its actions have no affect on rival profits. Thus, it provides the one imperfectly competitive model in which game theory is unimportant. It is not until we discuss oligopoly markets later in the book that game theory becomes invaluable. In each of the models that are discussed in this chapter, firms are assumed to be single product producers. We begin with monopoly.

---

*This is a review chapter. You can learn more detail about the basic models that are found in this chapter from any introductory or intermediate microeconomics textbook, such as Frank and Bernanke (2008), Mankiw (2011), Bernheim and Whinston (2008), Pindyck and Rubenfield (2009), and Varian (2010). For more advanced treatments, see Nicholson and Snyder (2012) and Mas-Colell et al. (1995).

V.J. Tremblay and C.H. Tremblay, *New Perspectives on Industrial Organization*, Springer Texts in Business and Economics, DOI 10.1007/978-1-4614-3241-8_6,

## 6.1 Monopoly

As just mentioned, an industry with only one seller is called a monopoly. The term "monopoly" can refer to a firm, an industry, or a market structure. The monopoly model is built on the following assumptions:

1. There is only *one firm* in the market. No other firms produce substitute products.
2. *Barriers to entry* are sufficiently high to allow only one firm in the market.
3. The monopolist's demand is the market demand, which has a negative slope. The firm is a "*price maker*," which means that the firm can raise its price without losing all of its customers.
4. Like perfect competition, the firm is a profit maximizer and there are *no frictions or other forms of market imperfections*.

Monopolies are relatively uncommon in the real world. One reason is that they are illegal under the Sherman Act (1890). Another reason is that when a monopolist earns an economic profit, there is a strong motive for entry. Thus, barriers to entry must be extremely high for there to be just one firm in the industry.

There are three types of barriers to entry. The first is a **natural barrier**, which is due to basic demand and cost conditions. One example occurs when the long-run average cost falls throughout the relevant range of demand. In this case, cost minimization requires a single producer. A market such as this is called a **natural monopoly**. Examples include public utility companies (but they are regulated by government to prohibit excess profits). Another example arises when a firm has sole control of an essential input, such as Alcoa's ownership of all bauxite (aluminum) mines before World War II.

A second type is a **legal barrier to entry**, which is a barrier caused by a government restriction. A patent is one example, which gives an inventor exclusive (monopoly) rights to the production and sale of an invention for 20 years. The purpose of patents is to create property rights for ideas and to stimulate innovation. Without a patent, others can copy an invention and earn profits without incurring the costs of conducting the research necessary to create the invention. A government franchise, which awards selling rights in an area, can also create monopolies. Franchises to taxicab companies in New York City and casino gaming licenses are examples.

Finally, there are **strategic barriers to entry**. Unlike the other types of entry barriers, strategic barriers are endogenous. That is, they involve firm actions that are designed to deter entry. Examples include predatory pricing, where price is set below unit cost, and investments that raise the costs of potential entrants so high that it is too costly to enter the market. We begin with a discussion of an unregulated monopolist and postpone formal discussion of entry barriers until Chap. 8.

### 6.1.1 Firm Behavior and Market Equilibrium

An unregulated monopolist will want to choose the price and output levels that maximize profit, which are determined simultaneously. For a demand function, a given quantity demanded corresponds to a demand price, and a given price corresponds to a level of quantity demanded. Thus, by determining the optimal output (price), the optimal price (output) can be obtained from the demand function. We will see that maximizing profit with respect to output or price produces the same outcome.

The traditional and simplest approach is to let the firm maximize profit with respect to output. This produces the same optimization principle as with perfect competition—the profit-maximizing level of output occurs where marginal revenue (MR) equals marginal cost (MC). In this chapter and later chapters, we focus on the long run, so MC always refers to the firm's long-run marginal cost. The difference in the profit maximizing level of output between perfect competition and monopoly lies in the nature of the firm demand and MR functions. Since firm demand is industry demand in monopoly, the monopolist's demand is negatively sloped. In contrast, the perfectly competitive firm's demand is a horizontal line, where price ($p$) is identical to MR and average revenue (AR). That is, $p \equiv \text{AR} = \text{MR}$ in perfect competition. In monopoly, $p \equiv \text{AR} > \text{MR}$.[1]

To solve the monopoly problem, we consider a simple example where both demand (D) and total cost functions are linear. The inverse demand takes the following form: $p = a - bq$, where $a$ is the price intercept and $b$ is the slope ($a$, $b > 0$). This implies that total revenue is $\text{TR} \equiv pq = aq - bq^2$. The total cost function is $\text{TC} = cq$, where $c$ is long-run marginal and average cost. To assure nonnegative profits, $a > c > 0$. The corresponding marginal functions are $\text{MR} = \partial \text{TR}/\partial q = a - 2bq$ and $\text{MC} = \partial \text{TC}/\partial q = c$. Profits ($\pi$) are given by

$$\begin{aligned} \pi &= \text{TR} - \text{TC}, \\ &= \left(aq - bq^2\right) - (cq). \end{aligned} \tag{6.1}$$

To find the output level that maximizes profit, we take the derivative of $\pi$ with respect to $q$ and set it equal to zero[2]:

$$\begin{aligned} \frac{\partial \pi}{\partial q} &= \frac{\partial \text{TR}}{\partial q} - \frac{\partial \text{TC}}{\partial q}, \\ &= \text{MR} - \text{MC}, \\ &= (a - 2bq) - c = 0. \end{aligned} \tag{6.2}$$

[1] In Chap. 2, we saw that $p > \text{MR}$ for a downward-sloping demand curve. For a linear demand function, we found that MR and demand have the same $y$-intercept but that MR is twice as steep as the demand function.

[2] As in the previous chapter, although $q$ is the only variable in this example, we use $\partial$ instead of $d$ to remind us that there are many other variables that are implicitly assumed to be held fixed.

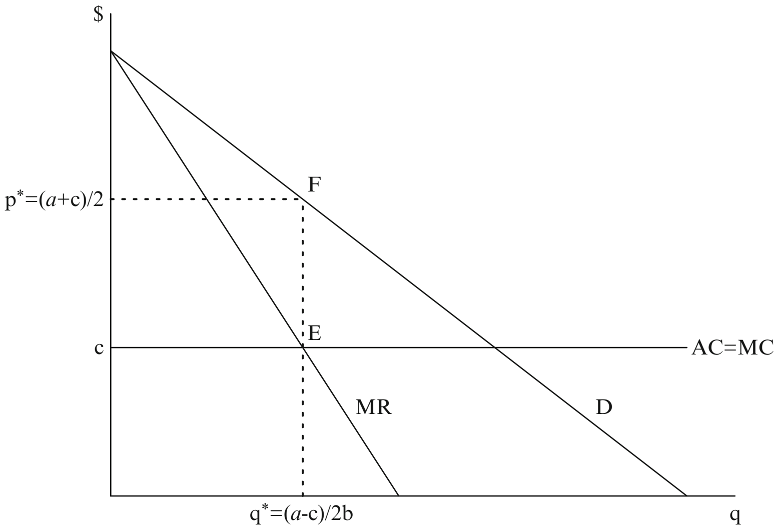


**Fig. 6.1** An example of a monopolist's optimal price and quantity

This is the first-order condition of profit maximization, which tells us that the profit maximizing level of output occurs where MR = MC.[3] Solving (6.2) for $q$ yields the profit maximizing level of output, $q^* = (a - c)/2b$. The profit maximizing price ($p^*$) is determined by substituting $q^*$ into the inverse demand function. Profits are found by substituting $p^*$ and $q^*$ into the profit equation. The optimal values are

$$q^* = \frac{a-c}{2b},$$
$$p^* = \frac{a+c}{2},$$
$$\pi^* = \frac{(a-c)^2}{4b}. \tag{6.3}$$

Figure 6.1 shows demand, marginal revenue, marginal cost, and the profit maximizing price–output pair for the example above. Note that MR = MC at $q^*$, $p^*$ is the maximum price that will just sell $q^*$ (i.e., the demand price at $q^*$), and $\pi^* = \text{TR} - \text{TC}$.

In this example, the monopolist is earning a positive long-run profit; $p^*$ exceeds average cost at $q^*$. Firm profit is area $p^*$cEF in Fig. 6.1. Unlike perfect competition,

[3] This produces a maximum because the profit function for each firm is concave. That is, the second-order condition of profit maximization holds, because the second derivative of the profit function is $-2b < 0$.

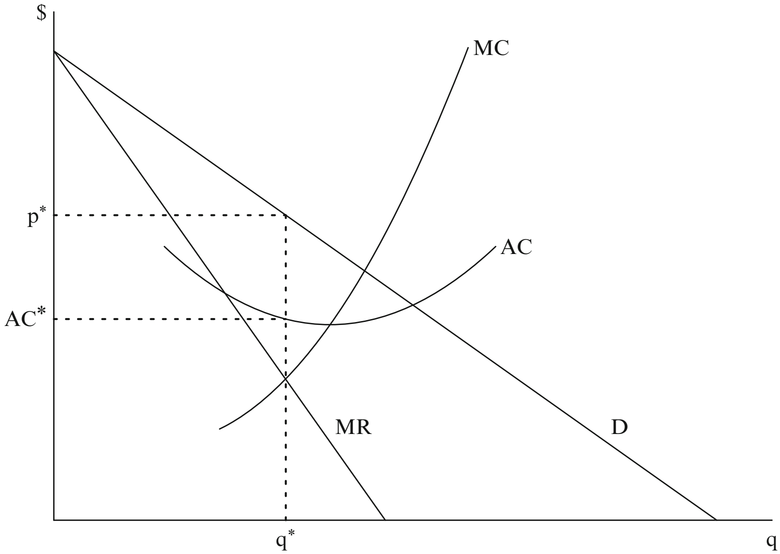


**Fig. 6.2** Equilibrium for a monopolist

long-run profit can be positive because entry barriers prevent the entry that would drive profits to zero.[4]

Figure 6.2 provides another example, this time with a U-shaped average cost function. If we optimize with respect to $q$, the profit maximizing level of output occurs at $q^*$ where MR = MC. The profit maximizing price is the demand price at that level of output, which equals $p^*$. The firm earns a positive profit, because $p^*$ exceeds the average cost at the profit maximizing level of production (AC$^*$).

To show that the results are the same whether the firm maximizes profit with respect to $q$ or $p$, we now consider the price problem. In this case, we need to define demand as a function of the choice variable $p$. To convert the inverse demand to a demand function, we solve demand for $q$: $q = a/b - (1/b)p$. Thus, the firm's profit equation becomes

$$\begin{aligned} \pi &= \text{TR} - \text{TC}, \\ &= pq - cq, \\ &= p\left(\frac{a}{b} - \frac{1}{b}p\right) - c\left(\frac{a}{b} - \frac{1}{b}p\right). \end{aligned} \tag{6.4}$$

[4] Of course, demand and cost conditions could be such that long-run profits are zero. This is precisely the long-run equilibrium in the monopolistically competitive model that we discuss subsequently. In the short run, a monopolist can lose money and stay in business as long as its optimal price is above its short-run variable cost, just as with a competitive firm.

To find the price that maximizes profits, we take the derivative of $\pi$ with respect to $p$ and set it equal to zero. This is the first-order condition with respect to price:

$$\begin{aligned}\frac{\partial\pi}{\partial p} &= \frac{\partial \mathrm{TR}}{\partial p} - \frac{\partial \mathrm{TC}}{\partial p}, \\ &= \mathrm{MR}_p - \mathrm{MC}_p, \\ &= \left(\frac{a}{b} - \frac{2}{b}p\right) + \frac{c}{b} = 0. \end{aligned} \tag{6.5}$$

where $\mathrm{MR}_p$ is a different type of marginal revenue, equaling the firm's marginal revenue with respect to a change in $p$. Similarly, $\mathrm{MC}_p$ is the firm's marginal cost with respect to a change in $p$. Solving (6.5) for $p$ yields the profit maximizing price:

$$p^* = \frac{a + c}{2}. \tag{6.6}$$

The profit maximizing output level is determined by substituting $p^*$ into the demand function above, $q^* = a/b - (1/b)p^* = (a - c)/2b$. These are the same optimal values as before, which demonstrates that the optimal price and output are the same whether the firm optimizes over $q$ or $p$. Although this is true for monopoly, it is not true in oligopoly markets, as we will see in Chap. 10.

One way to judge the validity of the model is to check to see if the comparative static results are reasonable. In this model, the equilibrium price rises with an increase in demand (parameter $a$) and an increase in marginal cost ($c$), and the equilibrium output level rises with demand (an increase in $a$ and a decrease in $b$) and decreases with marginal cost. These results are consistent with what we would expect to see in reality.

Finally, we can analyze the monopoly problem with general demand and cost functions. Let $p(q)$ represent the firm's inverse demand function. In this case, $\mathrm{TR} = p(q) \cdot q$ and the firm's profit equation is $\pi = \mathrm{TR} - \mathrm{TC} = p(q) \cdot q - \mathrm{TC}$. The first-order condition of profit maximization is[5]

$$\begin{aligned}\frac{\partial\pi}{\partial q} &= \frac{\partial \mathrm{TR}}{\partial q} - \frac{\partial \mathrm{TC}}{\partial q}, \\ &= \mathrm{MR} - \mathrm{MC}, \\ &= \left(p + \frac{\partial p}{\partial q}q\right) - \mathrm{MC} = 0. \end{aligned} \tag{6.7}$$

This and previous examples illustrates the standard marginal principle or rule of marginalism: to maximize profit with respect to activity $x$ (output in this case), the firm must equate the marginal benefit (MR) with the marginal cost (MC) of activity $x$.

[5] This is an application of the product rule, as discussed in the Mathematics and Econometrics Appendix at the end of the book. That is, if $y = wz$ where $w = f(x)$ and $z = g(x)$, then $\mathrm{d}y/\mathrm{d}x = w(\mathrm{d}z/\mathrm{d}x) + z(\mathrm{d}w/\mathrm{d}x)$. The derivative of the product of two functions equals the first function times the derivative of the second function plus the second function times the derivative of the first function. Because $\mathrm{TR} = p(q) \cdot q$, $\mathrm{MR} = p + (\partial p/\partial q)q$.

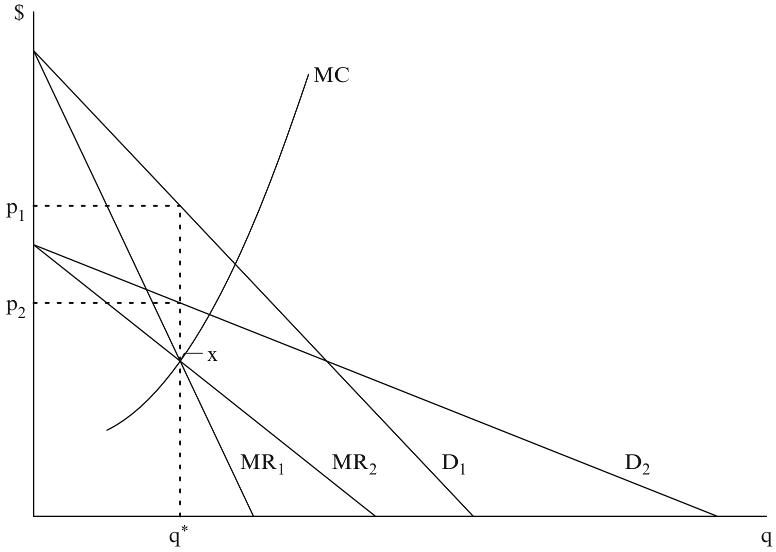


**Fig. 6.3** Different prices for the same level of output

### *6.1.2 No Supply Curve in Monopoly*

You might think that a monopolist's marginal cost is its supply function, as is true in perfect competition. We know that in perfect competition there is a single quantity supplied for a given price, but this is not the case with monopoly. There is not a one-to-one correspondence between price and quantity supplied in a monopoly market, implying that a supply function does not exist for a monopoly firm.

Notice that different optimal prices can correspond to the same output level, as in Fig. 6.3. For the demand curve, $D_1$, and corresponding marginal revenue curve, $MR_1$, we know that the monopolist will equate MC and $MR_1$ and produce $q^*$ at $p_1$. One way to map out a supply function in a perfectly competitive market is to allow demand to shift and identify all optimal points, which maps out the supply function. In the case of monopoly, however, this process does not map out a function. To illustrate, suppose that demand and marginal revenue change so that marginal revenue rotates around a fixed point, $x$, to produce $D_2$ and $MR_2$ in Fig. 6.3. Notice that $q^*$ remains the same but the optimal price decreases to $p_2$. If we hold costs fixed but change demand functions once again so that $x$ moves up the firm's marginal cost function, we will identify a greater $q^*$ and two new optimal prices. This demonstrates that there is not a unique optimal price for a given optimal level of output. Thus, a supply function does not exist in monopoly. In fact, a supply function only exists in perfect competition.

### 6.1.3 Allocative Inefficiency

Recall from the previous chapter that for a market to be allocatively efficient, price must equal marginal cost. This is clearly not true in monopoly. As Figs. 6.1 and 6.2 show, $p^* > \text{MC}$. Unlike a competitive firm, a monopolist is a price maker and has the power to raise price above marginal cost, leading to a price that is too high and a production level that is too low from society's perspective. Thus, monopoly is one type of market failure.

The ability of a monopolist to profitably maintain price above marginal cost is called **monopoly power**. An index of exerted monopoly power was developed by Lerner (1934), which is defined as

$$\mathcal{L} \equiv \frac{p - \text{MC}}{p}. \tag{6.8}$$

The Lerner index ranges from 0 to 1. When price equals marginal cost, there is no monopoly power and $\mathcal{L} = 0$. A higher value of $\mathcal{L}$ implies greater monopoly power.[6]

Lerner also showed that this index is related to the price elasticity of demand ($\eta$). Recall from Chap. 2 the following relationship between marginal revenue and $\eta$:

$$\text{MR} = p\left(1 - \frac{1}{\eta}\right). \tag{6.9}$$

For the profit maximizing monopolist, $\text{MR} = \text{MC}$, implying that

$$\text{MC} = p\left(1 - \frac{1}{\eta}\right). \tag{6.10}$$

Rearranging terms gives

$$\mathcal{L} \equiv \frac{p - \text{MC}}{p} = \frac{1}{\eta}. \tag{6.11}$$

Thus, there is an inverse relationship between $\eta$ and $\mathcal{L}$. Monopoly power increases as demand becomes more inelastic (i.e., as $\eta$ falls). This is clear from Fig. 6.3, where $q^*$ is the same for $D_1$ and $D_2$. At $q^*$, $D_1$ is relatively more inelastic than $D_2$, and the markup of price over marginal cost is greater for $D_1$ than $D_2$.

The connection between the Lerner index and the price elasticity of demand tells us something further about monopoly power. Note that even a monopolist faces products that are imperfect substitutes. There may be only one ice cream parlor in a

[6] It is also possible for a monopolist to have unexerted monopoly power, where the firm has the ability to raise price but chooses not to for public relations reasons or to avoid an antitrust challenge, for example.

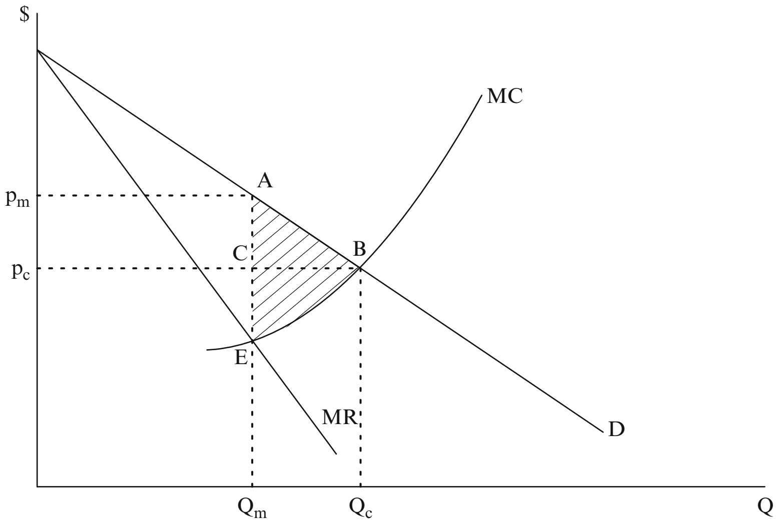


**Fig. 6.4** Equilibrium in monopoly and perfect competition and deadweight loss

small town, giving it monopoly status, but a neighboring bakery provides consumers with alternatives to ice cream. With more and more products that become closer and closer substitutes, the firm's demand function becomes more elastic, diminishing monopoly power. Market power falls as the firm faces increased competition from a greater number of substitute goods. As a numerical example, in the highly elastic case, if $\eta = 20$, $\mathcal{L} = 0.05$, whereas if demand is less elastic, $\eta = 2$, $\mathcal{L} = 0.50$, a considerably higher markup of price over marginal cost.

The Lerner index also tells us about the pass-through rate, the increase in price due to a small increase in MC. To demonstrate, we solve (6.11) for price: $p = \text{MC}\,[\eta/(\eta - 1)]$. For a monopolist, $1 < \eta < \infty$, so that $[\eta/(\eta - 1)] > 1$ and $p > \text{MC}$. If $\eta$ is constant for a small change in price, a \$1 increase in MC will cause $p$ to rise by more than \$1, indicating that the pass-through rate for a monopolist is greater than 1. Notice that in perfect competition, $\eta \to \infty$, $p = \text{MC}$, and the pass-through rate is 1.

The monopoly solution is inefficient because the optimal price exceeds marginal cost. Consider the monopoly problem in Fig. 6.4. The allocative efficient solution is where price equals marginal cost at the point where MC crosses demand (point B at $p_c$ and $Q_c$, the perfectly competitive outcome). The monopoly solution is at point A ($p_m$ and $q_m = Q_m$). This shows that the monopolist produces too little output and charges too high a price from society's perspective.

We can get a sense of the magnitude of the efficiency loss by investigating the loss in total (consumer plus producer) surplus due to monopoly. Recall from Chap. 5 that a market is allocatively efficient when total surplus is maximized, which occurs at point B in Fig. 6.4 (i.e., where demand equals supply, which is identical to marginal cost).

By moving from the allocatively efficient point to the monopoly optimum point (A), consumers lose and producers gain. Overall, the decrease in output from $Q_c$ to $Q_m$ causes total surplus to fall by the shaded area AEB, which is called the deadweight loss or efficiency loss due to monopoly. This area measures the dollar value of the welfare loss that is caused by the monopolist producing $Q_m$ instead of $Q_c$.

### *6.1.4 X-inefficiency and Rent-Seeking Behavior*

To this point, we have assumed that market structure has no effect on firm costs. Profit maximization assures cost minimization and, therefore, economic efficiency. All profit maximizing firms, regardless of market structure, will operate on (not above) their cost function. In reality, managers and workers may not always work as hard as they could and may make mistakes. Regarding behavioral factors, overconfidence can lead to risky investment decisions by managers and risky behavior of workers regarding workplace safety. Other contributing cognitive issues, discussed in Chap. 4, include confirmation bias and cognitive dissonance. Problems such as these cause firm costs to be higher than they would be if the firm were fully efficient.

Cognitive errors and insufficient effort that push up firm costs are less likely to be an issue in competitive markets, because firms with higher costs than their competitors will go out of business in the long run. This is a natural selection argument, where only the fittest (most efficient) institutions survive in the long run (Alchian 1950). Nevertheless, inefficient firms may survive if there is insufficient competition. Thus, inefficiency due to lax work effort and cognitive errors is more likely to occur in a monopoly market. As Hicks (1935, 8) pointed out, "the best of all monopoly profits is the quiet life." Leibenstein (1966) calls this **X-inefficiency**.

X-inefficiency can be viewed as a deadweight loss in the sense that less is produced with no offsetting gain in consumer or producer surplus. Protection from competitive pressure can facilitate X-inefficiency in the public sector and in other industries besides monopolies.

Another type of inefficiency due to monopoly is **rent seeking**, the act of investing resources into nonproductive activities to obtain and maintain monopoly power.[7] This normally takes the form of lobbying efforts and campaign contributions to government officials in exchange for creating and maintaining legal barriers to entry. Tax breaks, subsidies, tariff protection, and licensing laws are common ways of shielding firms from competition. Rent-seeking behavior is socially wasteful, because it is costly and does not lead to an increase in output.

[7] These are normally associated with legal activities and exclude rents deriving from corruption and illegal bribes. For further discussion, see Tullock (1967) and Posner (1975).

In fact, when it increases monopoly power, it leads to less production and greater deadweight loss. Thus, the full social cost of monopoly power must include rent seeking expenditures as well as the traditional deadweight loss.[8]

### 6.1.5 Dynamic Considerations: Addiction and Product Durability

Thus far, we have analyzed monopolistic markets in a static world where decisions, costs, and benefits occur in a single period. In some circumstances, however, it benefits the monopolist to consider more than one period when making a decision. Demand may be interdependent from one period to the next as in the case of addictive commodities. In addition, greater production today may lower costs tomorrow when workers gain from experience or learning by doing. Investment in research and development today may also bring expected benefits in the future. Finally, an increase in product durability affects a consumer's need to replace a good tomorrow. These are examples of dynamic markets where actions in one period affect profits in another period. In this section we introduce two cases where dynamic considerations are important, addiction and product durability.

For addictive commodities like cigarettes or addictive drugs, consuming the good today increases the probability of consuming the good tomorrow. As is dramatized in Hollywood movies, a monopoly dealer will give addictive drugs away for free in the first period, increasing future demand and allowing the dealer to substantially raise price once consumers are addicted. If we were to ignore the dynamic nature of the market, it would look like there is no monopoly power in the first period, because the price is not above MC. Thus, the measurement of monopoly power is a bit more complex in a dynamic market. We discuss this issue in Chap. 12.

In a dynamic market, the firm will want to choose a level of production today that maximizes the sum of profits today and into the future. With two periods, the current period (I) and next period (II), total profit is

$$\begin{aligned} \Pi &= \pi_{\mathrm{I}} + \pi_{\mathrm{II}}, \\ &= \mathrm{TR}_{\mathrm{I}} - \mathrm{TC}_{\mathrm{I}} + \pi_{\mathrm{II}}, \end{aligned} \tag{6.12}$$

where $\pi_i$ is profit in period i, I or II, $\mathrm{TR}_{\mathrm{I}}$ is total revenue in period I, and $\mathrm{TC}_{\mathrm{I}}$ is total cost in period I. The monopolist's first-order condition of profit maximization with respect to production in period I is

[8] Rent-seeking expenditures themselves may be viewed as simple transfers from monopolies to politicians. Nevertheless, not all is transferred and rent seeking that effectively increases market power raises the deadweight loss associated with monopoly.

$$\begin{aligned}\frac{\partial \Pi}{\partial q_{\mathrm{I}}} &= \frac{\partial \mathrm{TR}_{\mathrm{I}}}{\partial q_{\mathrm{I}}} - \frac{\partial \mathrm{TC}_{\mathrm{I}}}{\partial q_{\mathrm{I}}} + \frac{\partial \pi_{\mathrm{II}}}{\partial q_{\mathrm{I}}},\\ &= \mathrm{MR}_{\mathrm{I}} - \mathrm{MC}_{\mathrm{I}} + \frac{\partial \pi_{\mathrm{II}}}{\partial q_{\mathrm{I}}} = 0,\end{aligned} \tag{6.13}$$

where $\mathrm{MR_I}$ is marginal revenue in period I and $\mathrm{MC_I}$ is marginal cost in period I. The term $\partial \pi_{\mathrm{II}}/\partial q_{\mathrm{I}}$ represents the effect of a marginal increase in output in period I on profit in period II. Notice that if this were a static problem, $\partial \pi_{\mathrm{II}}/\partial q_{\mathrm{I}} = 0$, and we get the usual condition of profit maximization, $\mathrm{MR_I} = \mathrm{MC_I}$. The market is static because an increase in $q_{\mathrm{I}}$ has no effect on future profit.

With an addictive commodity, however, future profits will be positively related to current consumption, i.e., $\partial \pi_{\mathrm{II}}/\partial q_{\mathrm{I}} > 0$. As $\partial \pi_{\mathrm{II}}/\partial q_{\mathrm{I}}$ gets larger, the marginal benefit of increasing production today ($\mathrm{MR_I} + \partial \pi_{\mathrm{II}}/\partial q_{\mathrm{I}}$) goes up, and the firm will produce more output in period I (i.e., charge a lower price in period I), just as in our Hollywood movie example.

The opposite happens in a durable goods market. If a good that is produced in period I is still available to consumers in period II, then an increase in current production and consumption will lower demand tomorrow. In this situation, $\partial \pi_{\mathrm{II}}/\partial q_{\mathrm{I}} < 0$ and the monopolist will produce less of a durable good in period I. Furthermore, the firm will produce less and less in period I as the good becomes more durable. If product durability is under the control of the monopolist, "planned obsolescence"—purposefully designing products that will wear out or become obsolete more quickly—may be a profitable strategy. We take up this topic and formal methods to solve dynamic problems in Chap. 11.

### *6.1.6 Social Benefits of Monopoly*

Although a monopoly is not allocatively efficient, it may be productively efficient. The classic example is the natural monopoly, where industry output is produced at lowest cost by a single firm. Such a market is characterized by substantial scale economies, as depicted in Fig. 6.5. Notice that AC is lowest with one firm producing all industry output. Thus, productive efficiency requires that there be just one firm in a market when there are pronounced scale economies.

Even though an unregulated natural monopoly is productively efficient, it will be allocatively inefficient. Production will take place where MR = MC, at $q^*$ and $p^*$ in Fig. 6.5. In terms of public policy, a monopoly is required to assure productive efficiency, but price is typically regulated to minimize allocative inefficiency.

To completely eliminate allocative inefficiency, the price would need to be regulated so that it equals marginal cost at the point where it crosses demand (point A in Fig. 6.5). At this price, however, the firm is losing money because $p = \mathrm{MC} < \mathrm{AC}$. Thus, a subsidy would be required to keep the firm in business. To avoid an administratively costly subsidy, in practice price is generally capped so that that firm earns zero profit in the long run. That is, price is set equal to average cost where it crosses demand (at point B in Fig. 6.5). These regulatory issues will be discussed in more detail in Chap. 20.

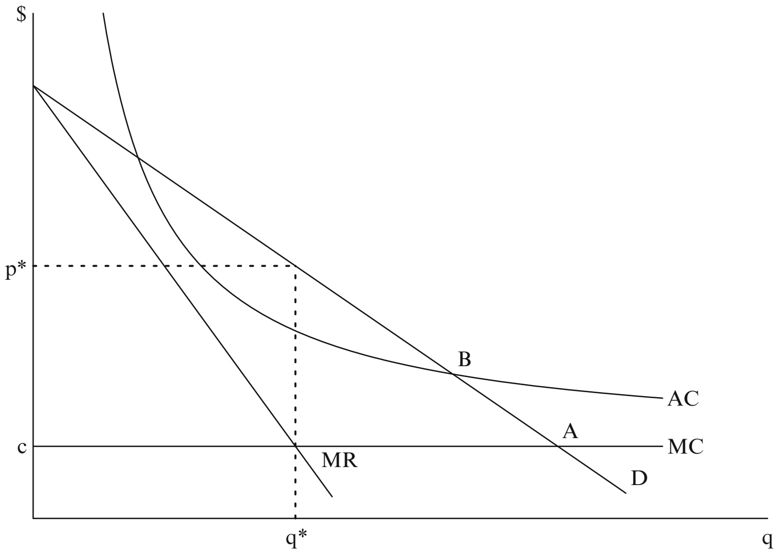


**Fig. 6.5** Scale economies and natural monopoly

Finally, monopoly may be more dynamically efficient than other market structures. Schumpeter (1942) hypothesized that large corporations are necessary for dynamic efficiency because they are more likely to invest in the research and development that drives technological change and economic progress. In addition, as we discussed in Chap. 1, Demsetz (1973) argued that many firms gained their monopoly positions by being superior firms. Thus, monopoly profits can be a reward for success, and such rewards encourage effort and innovative activity. We discuss these dynamic issues more fully in Chaps. 17, 19, and 20.

## 6.2 Monopolistic Competition

Next, we investigate the model of monopolistic competition (Chamberlin 1933). The name derives from the fact that it has features in common with perfect competition and with monopoly. A distinctive aspect of this imperfectly competitive model is that firms are assumed to be so small that strategic interaction is nonexistent. Thus, this model does not require sophisticated game theory tools.

The model of monopolistic competition derives from the following assumptions. Some are similar to perfect competition and others are similar to monopoly:

1. There are *many identical firms* in the industry, and each firm is so small that strategic interaction is zero.
2. Firms produce *differentiated products*. That is, each firm produces a product that performs the same basic function but has slight differences from rival products.

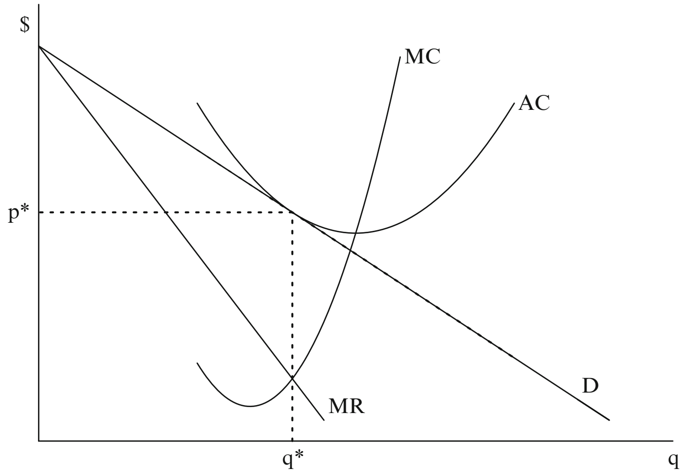


**Fig. 6.6** Equilibrium in monopolistic competition

3. There are *no barriers to entry or exit*.
4. Each firm is a profit maximizer, and there are *no frictions or other forms of market imperfections*.
5. There are *economies of scale* in production.

The assumptions of many sellers, no entry barriers, profit maximization, and no market imperfections are identical to those of perfect competition. The key difference is that there is product differentiation.[9] Levi and Lee jeans perform the same function but differ in styling, for example. A differentiated product gives each monopolistically competitive firm a monopoly over the sale of its own brand. Only the Levi Company can sell Levi brand jeans. The importance of the assumption that there are economies of scale will become apparent subsequently.

The long run equilibrium in monopolistic competition has two key features. First, because each firm has a monopoly over the sale of its own brand, firms face negatively sloped demand functions (i.e., they are price makers). We will see that this gives firms monopoly or market power.[10] At the same time, this power is limited by the presence of many close substitutes. Second, free entry ensures that long-run profits are zero, that is, price ($p^*$) equals long-run average cost (AC). This occurs at the point where demand (D) is tangent to AC, as depicted in Fig. 6.6.

[9] In the next chapter, we discuss the different types of product differentiation. At this point, all that matters is that products are different in the eyes of consumers.

[10] Because we are not talking about a true monopoly firm, it may be better to call this market power than monopoly power. Carlton and Perloff (2005, 93) suggest that we define monopoly power as the case where $p^* > \text{MC}$ and firm long-run profits are positive and define market power as the case when $p^* > \text{MC}$ and long-run profits are zero. However, these terms are generally used interchangeably.

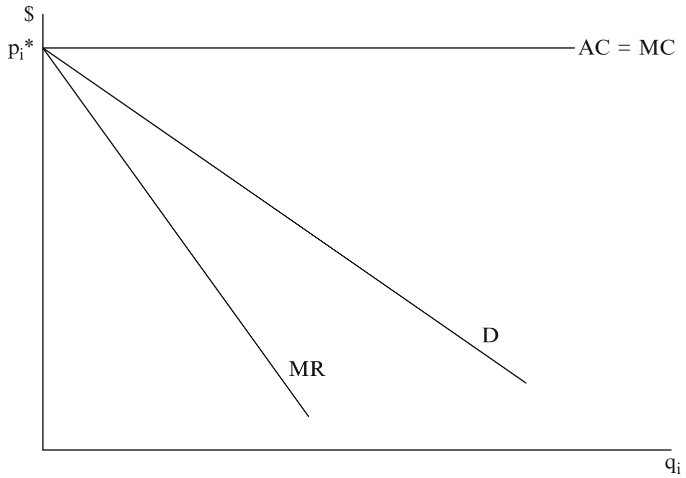


**Fig. 6.7** Constant returns to scale and monopolistic competition

Given the presence of scale economies, long-run marginal cost is less than long-run average cost at the optimum. Thus, $p_i^* = \text{AC} > \text{MC}$; profits are zero even though the firm has monopoly power. For the market to be in equilibrium, these conditions must prevail for every firm.

The welfare implications of this model are complex. On the minus side, the market is allocatively inefficient, because price exceeds marginal cost. Further, production is below efficient scale, as it takes place in the region of economies of scale. Production costs would be lower if there were fewer firms that each produced greater output. With many close substitutes, however, demand is relatively flat. Price will be close to marginal cost, and production will be close to minimum AC.[11] Thus, the social cost of these factors is relatively small. On the plus side, this market structure brings consumers greater product variety than one would find in perfect competition or in monopoly. With fewer firms each producing more output but at lower AC, product variety would diminish. Therefore, it is unlikely that the welfare effect of monopolistic competition is sufficiently negative to warrant a policy response.

The final question we address is: why does the model require economies of scale? It turns out that as scale economies diminish, the market outcome approaches that of perfect competition (Färe et al. 2012). Suppose that there are constant returns to scale. Then AC would be a horizontal line and equal MC. Free entry assures zero profit, $p_i^* = \text{AC}$. Thus, $p_i^* = \text{AC} = \text{MC}$. This produces a corner solution, as depicted in Fig. 6.7, where it is assumed that each firm produces an infinitesimally

[11] We exaggerate the steepness of demand (and the markup of price over marginal cost) to make it easier to see the tangency point in the figure.

small amount of output. This demonstrates how the nature of technology can be an important determinant of market power.

In essence, without any economies of scale, each producer becomes identical. For example, if there were no cost savings from the mass production of clothing, we would all shop at a tailor for custom made clothing. In effect, product differentiation is eliminated, as each firm provides the same service of producing custom made clothing. Thus, some degree of scale economies is required for a market to be monopolistically competitive.

## 6.3 Summary

1. A monopoly exists when there is only one seller in an industry.
2. A monopoly is protected from competition by barriers to entry. There are three types of barriers to entry. A **natural barrier** exists when the presence of scale economies limits the number of competitors that can profitably enter a market. A **legal barrier to entry** is due to a government restriction. A **strategic barrier to entry** is due firm actions that are designed to deter entry.
3. Demand and marginal revenue functions are downward sloping in monopoly because firm demand is industry demand.
4. The monopolist will produce the level of output that maximizes profit where marginal revenue equals marginal cost. Price will be determined along the demand curve at the optimal level of output.
5. In monopoly, price can exceed average cost and economic profit can be sustained in the long run. Unlike perfect competition, new firms cannot enter the market and erode profits because of entry barriers.
6. In equilibrium, $p^* > \text{MC}$, indicating that the monopoly result is **allocative inefficiency**.
7. We cannot specify a supply curve in monopoly as there is no one-to-one correspondence between price and quantity.
8. The **Lerner Index** is the difference between price and marginal cost as a fraction of price. It measures the degree of monopoly power and allocative inefficiency. The Lerner Index is inversely related to the price elasticity of demand. When demand is relatively inelastic, the index is higher indicating that the monopolist has greater power over price.
9. In monopoly, output is lower and price is higher than in perfect competition.
10. Social welfare, as measured by total surplus, is lower under monopoly than under perfect competition. The lost surplus due to monopoly is called the **deadweight loss**.
11. **X-inefficiency** is inefficiency that arises from insulation from competitive pressure on the monopolist and the workers. It leads to higher costs.
12. Resources are also wasted by **rent-seeking**, efforts and monies expended by the firm to protect its monopoly position.

13. Addiction, learning-by-doing, research and development, and product durability lead to **dynamic** problems, where production or investment today can affect profits tomorrow. The firm will maximize profits across periods. When output today positively relates to output tomorrow (e.g., an addictive commodity), the monopolist will produce more today. When output today negatively relates to output tomorrow (e.g., a durable product), the firm will produce less today.
14. A **natural monopoly** arises when industry output is produced at lowest cost by one firm. It is productively efficient but allocatively inefficient.
15. Monopoly may be dynamically efficient if it leads to greater investment in research and development. Monopoly power may serve as an incentive for firms to perform in a superior way.
16. **Monopolistic competition** is characterized by many identical firms that are profit maximizers, product differentiation, free entry and exit, no frictions or other market imperfections, and economies of scale.
17. The assumption of **product differentiation** is that all firms produce goods that serve the same basic function but are slightly different from one another.
18. In long-run equilibrium in monopolistic competition, firm profits are zero but there is allocative inefficiency, i.e., $p^* = \text{AC} > \text{MC}$. Because each firm's demand is relatively elastic in monopolistically competitive markets, the degree of monopoly power is limited.

## 6.4 Review Questions

1. Suppose that you are the owner of a metals-producing firm that is an unregulated monopoly. You find that your marginal cost curve can be approximated by a straight line, $\text{MC} = 60 + 2q$, where MC is marginal cost (in dollars) and $q$ is output. Inverse demand is $p = 100 - q$, where $p$ is the product price. What is the equation of your MR curve? What are your profit maximizing $q$ and $p$?
2. A textbook author sells the rights to a book to a publisher, and copyright laws give the publisher a monopoly over the sale of the book. Authors are typically paid a percent of total revenues. If the publisher is a profit maximizer, show that the author will prefer to sell more books than the publisher.
3. Do unregulated monopolists always make positive economic profits? Use a graph to show that a monopolist could earn zero economic profit in the long run.
4. A monopoly producer charges a price of $1 for its product. Assuming that the monopoly is maximizing profits and the absolute value of the price elasticity of demand $\eta = 2$ at that price, calculate the monopolist's MR and MC.
5. Show graphically the deadweight loss associated with monopoly when costs are constant, i.e., $\text{AC} = \text{MC} = c$. Point out differences in consumer surplus and producer surplus (if any) between the perfectly competitive and monopoly outcomes.
6. Provide an example of rent-seeking behavior and of X-inefficiency.

7. Give an example of how the behavior of managers can lead to X-inefficiency in a firm with monopoly power.
8. Learning-by-doing occurs when workers and management become more productive as they gain experience from producing more output and running the company. When learning-by-doing plays a role in production, what do you expect will be the sign of $\partial \pi_{II}/\partial q_{I}$? ($q_I$ is the level of production in period I, and $\pi_{II}$ is profit in period II.) Do you think that the firm should produce more or less of the good in the current period?
9. Suppose that the total cost function in an industry is given by TC $= c \cdot q$. Do you think that there could be a natural monopoly in this industry? Why or why not?
10. Consider an industry with a linear inverse demand, $p = 100 - Q$, and MC = AC = \$10. Solve for industry output, price, and profits if the industry is:

    A. Perfectly competitive
    B. Monopolistic

11. This question relates to the Lerner Index, $\mathcal{L}$.

    A. Based on the formula for the Lerner Index in (6.8), how would the value of $\mathcal{L}$ compare for perfect competition versus monopolistic competition? How do you suspect that the value of $\mathcal{L}$ would compare for monopolistic competition versus monopoly?
    B. Based on the relationship between $\mathcal{L}$ and $\eta$ in (6.11), do you agree with your responses in part (A)? Explain.

12. Consider an established monopoly firm. Explain how the behavioral concept of the endowment effect relates to barriers to entry in the industry.