CHAPTER 13

Monopolistic Competition and Oligopoly

What are the characteristics of monopolistic competition and oligopoly market structure models?

Chapter Outline

Learning Objectives

- Explain how price and output are determined under monopolistic competition.
- Understand the characteristics of oligopoly.
- Explore several key non-cooperative oligopoly models: Cournot, Stackelberg, and dominant firm.
- Show how price and output are determined under the cooperative oligopoly model of cartels.

Competition and pure monopoly lie at opposite ends of the market spectrum. Competition is characterized by many firms, unrestricted entry, and a homogeneous product, while a pure monopoly is the sole producer of a product. Yet many real-world market structures seem to be incompatible with either the competitive or the pure monopoly model. How do we analyze a market situation, then, where there are a dozen similar but slightly different brands of aspirin or only three companies supplying breakfast cereals?

Falling between competition and pure monopoly are two other types of market structure: monopolistic competition and oligopoly. Monopolistic competition is closer to

competition; it has many firms and unrestricted entry, like the competitive model, but the firms' products are differentiated. Fast-food chains, for example, may be viewed as monopolistic competitors. They supply the same general product, fast food, but one chain's specialty burger, say the Big Mac, is "different" from another's, such as the Whopper. Oligopoly is more like pure monopoly; it is characterized by a small number of large firms producing either a homogeneous product like steel or a differentiated product like cars.

This chapter examines monopolistic competition and oligopoly market structure models, noting the similarities with as well as the differences from perfect competition and pure monopoly. We also explore the case of cartels, whereby firms in an industry attempt to coordinate price and output decisions so as to act, in concert, as a pure monopoly and maximize their joint profit.

13.1 PRICE AND OUTPUT UNDER MONOPOLISTIC COMPETITION

MONOPOLISTIC COMPETITION a market characterized by unrestricted entry and exit and a large number of independent sellers producing differentiated products

DIFFERENTIATED PRODUCT a product that consumers view as different from other similar products

Monopolistic competition resembles both competition and monopoly. As with competition, entry into and exit from the industry are unrestricted, often resulting in a large number of independent sellers. However, in contrast to competition, the firms do not produce a homogeneous product. Instead, their products are heterogeneous, or differentiated. A **differentiated product** is one that consumers view as different from other similar products. For example, Wheaties and Cheerios are differentiated products in the general category of breakfast cereals. Consumers are not indifferent among brands of cereals; they perceive differences in taste, crunchiness, caloric content, and nutritional value. In a competitive market, by contrast, consumers view the product of one firm as identical to (a perfect substitute for) any rival firm's product.

Product differentiation may reflect *real* differences among products (Special K cereal is lower in fiber but higher in protein than Post Raisin Bran), or it may be based only on the *belief* that there are differences (a three-year-old may perceive Fruit Loops to be sweeter than Frosted Flakes but their sugar content is the same). The content of most aspirins is virtually identical, but many consumers believe Bayer to be superior to other brands. In blind taste tests, many consumers claiming to have strong preferences for Coca-Cola over Pepsi cannot select their preferred brand. This outcome doesn't affect the theory, however; the important point is that consumers, or at least a substantial number of them, *believe* the products to be different.

There are many aspects to product differentiation. For example, products may be differentiated by physical features such as function, design, or quality, or by advertising, brand names, logos (such as the rainbow-hued apple on Macintosh computers), or packaging (such as Oscar Mayer Lunchables). They may also be differentiated by conditions related to the sale, such as credit terms, availability, or congeniality of sales help, location (have you ever shopped at a nearby 7-Eleven because of its convenience?), or service. As this list suggests, many of the goods you purchase are differentiated products. Clothing, drugs, cosmetics, restaurant meals, and many types of food products are prominent examples.

Determination of Market Equilibrium

The first step in analyzing monopolistic competition is understanding the demand curve that confronts a single firm. When a firm sells a differentiated product with close substitutes, it has some degree of monopoly power—hence, the "monopolistic" in monopolistic competition. In other words, the demand curve confronting the firm is downward-sloping. However, the degree of monopoly power will typically be slight because of the availability of close substitutes.

For instance, because McDonald's is the only firm selling Big Macs, the quantity of Big Macs sold is unlikely to fall to zero if McDonald's charges a *slightly* higher price than its competitors. But at a higher price for Big Macs, many consumers might switch to a Burger King Whopper or a Wendy's Double Bacon Cheeseburger. Thus, *the demand curve facing each firm in a monopolistically competitive market is downward-sloping but fairly elastic*.

Assume that the market for jeans is monopolistically competitive. In Figure 13.1a, we show the demand curve, *D*, for one firm in this market, Tight Jeans. The demand curve's position depends strongly on the prices of other jeans, as well as the variety available. Thus, in drawing the demand curve for Tight Jeans, we assume that the number of other firms in the industry is fixed. Furthermore, we assume that the prices charged by other firms do not change when Tight Jeans varies its price. (Changes in the prices charged by other firms would cause the demand curve for Tight Jeans to shift.) The basis for assuming other firms' prices as given is that Tight Jeans represents only a small part of the total jeans market. While a lower price for its jeans will cause some customers to shift from other brands, the loss for each brand will be small enough to be unnoticeable, or at least not to provoke a reaction.

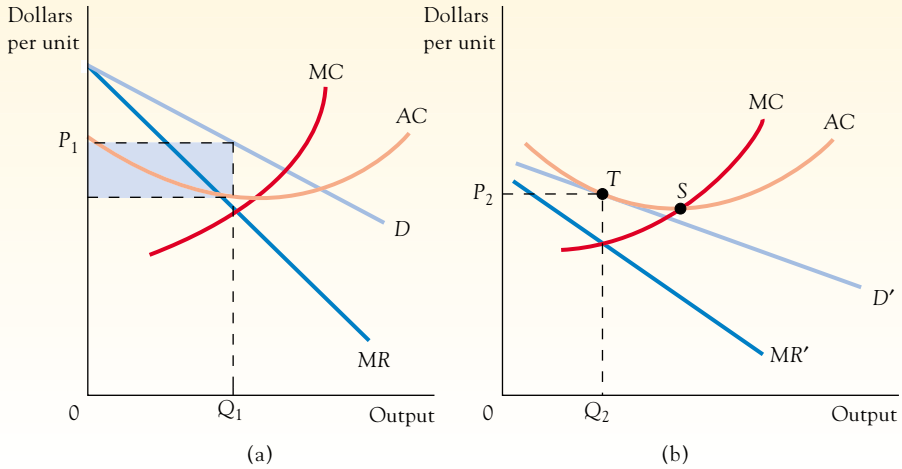
Given the behavior of other firms in the market, let's consider how Tight Jeans determines price and output. Because its demand curve is downward-sloping, marginal revenue is less than price, and profit maximization calls for operating where marginal revenue equals marginal cost. If the firm has the cost curves shown in Figure 13.1a, it produces an output of Q_1 and charges a price of P_1. The resulting economic profit equals the shaded area.

In terms of the diagram, the position of the monopolistically competitive firm resembles that of a monopoly. However, there are two important differences. First, Tight Jeans is only one among a number of firms producing a similar product, and so the demand curve is not

FIGURE 13.1

Monopolistic Competition

(a) In the short run, a firm in a monopolistically competitive market may make a profit. (b) Attracted by the prospect of profits, new firms enter the market. As entry continues, the demand curve for existing firms shifts downward until a zero-profit, long-run equilibrium is attained.



the market demand curve for jeans; it is only the demand curve for jeans produced by one firm. Second, under monopolistic competition, as distinct from pure monopoly, entry into the market is unrestricted. When existing firms are making profits, other firms are attracted to the market. Thus, the equilibrium in Figure 13.1a cannot be a long-run equilibrium because profits are being realized. It could represent a short-run equilibrium, but with the entry of other firms, the demand curves facing each existing firm will shift.

Under monopolistic competition, long-run equilibrium is attained as a result of firms entering (or leaving) the industry in response to profit incentives. In the present example, entry continues to occur until firms in the market are no longer making economic profits. How will the entry of other firms affect existing firms like Tight Jeans in Figure 13.1a? New firms will increase the industry's total output, as well as provide for a wider variety of differentiated products. Both of these effects shift existing firms' demand curves downward, leading to a general reduction in the industry's level of prices and, from that, lower profits. (It is also possible that entry will lead to higher prices for some inputs, causing cost curves to shift upward as in an increasing-cost competitive industry, but we will ignore this possibility.) Entry and output adjustments by existing firms will continue until economic profits are zero; only then will there be no further incentive for other firms to enter the market.

Figure 13.1b shows a position of long-run equilibrium for Tight Jeans. The firm's demand curve has shifted down to D', a position where it is just tangent to the average cost curve at point T. (If the demand curve intersected the average cost curve, then there would be a range of output over which profit would be positive; the final equilibrium must be a tangency.) The profit-maximizing output is now Q_2 with a price of P_2; at this price and output Tight Jeans makes zero economic profit.[1] All rival firms will be in a similar situation, making zero economic profit in long-run equilibrium. Their cost and demand curves, however, need not be identical because they are not producing exactly the same products. For this reason, there may be a range of prices prevailing in equilibrium. Given the similarity among the differentiated products within a monopolistically competitive market, prices are likely to vary over a small range. A Big Mac and a Whopper need not be the exact same price, for example, but it would be surprising if the prices differed substantially.

Firms in a monopolistically competitive industry compete not only on price, but also by variations in their products intended to attract customers. The range of differentiated products in a market is not fixed, and firms often introduce new variations they believe will be profitable. For instance, when Coca-Cola introduced its caffeine-free Coke, it was betting that enough consumers wanted to limit caffeine intake for the line to be profitable. The company was right, and for a time it found itself in the position shown in Figure 13.1a, making a profit. But once it was recognized that this was a profitable way to differentiate cola drinks, other firms followed suit. Coca-Cola's profit eroded as the market moved toward a long-run equilibrium.[2]

Note that the long-run equilibrium position is similar to both the competitive and the monopolistic equilibria. As with perfect competition, each firm's demand curve is tangent to its long-run average cost curve, so economic profit is zero. As with monopoly, the demand curve is downward-sloping, so price exceeds marginal cost at the equilibrium. However, be-

[1]It is not a coincidence that marginal revenue and marginal cost are equal at the output where the demand curve is tangent to the average cost curve; it is a geometric necessity. Try depicting the equilibrium with total revenue and total cost curves to see why.

[2]Not all new product variations, of course, are successful. For example, McDonald's introduced the McLean burger during the 1980s, hoping that it would satisfy the tastes of health-conscious fast-food consumers. The McLean burger never proved profitable and came to be known as the "McFlopper" by industry analysts.

cause the firm's demand curve is relatively elastic, price will normally not exceed marginal cost by very much. For instance, demand elasticities for monopolistically competitive firms can easily exceed 10. If the firm's demand elasticity is 15, we can use the markup formula explained in Chapter 11 [$(P - MC)/P = 1/\eta$] to see that when profit is being maximized, the markup would be only about 7 percent of price.

Monopolistic Competition and Efficiency

In Chapter 10, we saw that a competitive industry tends to be efficient, while in Chapter 11 we saw that a monopoly is inefficient (produces a deadweight loss). Because monopolistic competition combines elements of both monopoly and competition, it is natural to consider whether it is an efficient market structure, like competition, or inefficient, like monopoly.

Monopolistic competition has been charged with inefficiency in two respects. We can examine both with the aid of Figure 13.2, which shows a monopolistically competitive firm in long-run equilibrium (ignore the D^* demand curve for now). The first aspect of the alleged inefficiency involves the fact that the firm does not operate at the minimum point on its long-run average cost curve. In the diagram, the firm operates at point *A*, where average cost per unit is greater than at point *S*. Every firm in the monopolistically competitive industry is in a similar position. By contrast, firms in a competitive industry operate at the minimum points on their long-run *AC* curves. When firms fail to produce at lowest possible average cost, they are sometimes said to have **excess capacity**.

EXCESS CAPACITY the result of firms failing to produce at lowest possible average cost

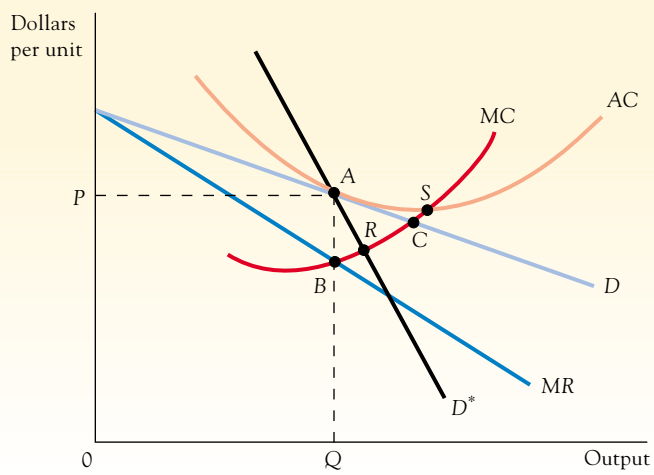
A failure to operate at minimum average cost is potentially inefficient because it is possible to produce the same *industry* output at a lower cost. To verify this notion, suppose that there are currently 10 firms like the one in Figure 13.2, each producing 100 units of output at an average cost of \$15. The total cost of producing the 1,000 units would therefore be \$15,000 ($10 \times 100 \times \$15$). If the average total cost is at a minimum of, say, \$11 per unit at an output of 125 units per firm, then 8 firms could produce the *same* 1,000 total output for less total cost. The total cost would now be \$11,000 ($8 \times 125 \times \$11$).

A monopolistically competitive market has also been alleged to be inefficient because it produces the wrong total output from a social perspective. (Note that in discussing excess capacity we were concerned with an unchanged total output.) Each firm is producing an output where price is greater than marginal cost. This condition suggests, by analogy to the

FIGURE 13.2

Alleged Deadweight Loss of Monopolistic Competition

The monopolistic competitor's demand curve is *D* when it alone varies price; the demand curve D^* is relevant when all firms simultaneously change output. The deadweight loss is shown by area *ARB*; similar areas for the other monopolistically competitive firms can be added to this area to obtain the total deadweight loss due to restricted output in this market.



case of pure monopoly, that additional output is worth more to consumers than the cost of producing it. There is a deadweight loss from producing too little.

Figure 13.2 shows a monopolistically competitive firm producing an output of Q where price, AQ, is greater than marginal cost, BQ. It is tempting to apply the same reasoning we did in the case of pure monopoly and argue that the magnitude of the deadweight loss equals triangular area ACB. By performing the same calculation for each firm in the industry and adding up the results, we could arrive at the deadweight loss for the entire monopolistically competitive industry. Tempting as it is, this procedure is incorrect and overstates the industry's total potential deadweight loss.

To understand this, consider the firm in Figure 13.2 expanding output to the apparently efficient point C on its demand curve. Recall that the firm's demand curve is drawn on the assumption that rival firms keep their prices unchanged. This is the appropriate assumption if we are examining a price change by one firm alone. However, the prospective inefficiency here is that *all* firms in this industry are producing too little. If all expand their output, the demand curve confronting each firm must shift downward. Consequently, it is not desirable for every firm to expand output to the point where marginal cost intersects its *initial* demand curve, since that curve shifts in reaction to output and price changes by the other firms.

There is a complicated interdependence between individual firms' demand curves in an industry composed of several firms, and that interdependence must be accounted for in evaluating efficiency. (Note that this problem did not arise with pure monopoly because there was only one firm in the industry, or with a competitive market where we worked with industry, and not firm, demand curves.) One way to account for the interdependence is to draw the demand curve confronting the firm when it and all other firms in the industry simultaneously expand output. This demand curve, shown as D^* in Figure 13.2, captures the interdependence among the firms and shows that the marginal value, or price, of the firm's product falls more rapidly when rival firms are also producing more units. Point R now represents the efficient output of this firm, and this is consistent with every other firm also having expanded output to the point where their price is equal to marginal cost. We can thus sum the areas like ARB to arrive at the total deadweight loss resulting from each firm producing at a point where price exceeds marginal cost. Of course, the important point is that the industry's total deadweight loss is smaller than if we erroneously sum up the areas like ACB.

While monopolistic competition has been charged with being inefficient, there are three reasons why government intervention probably is not warranted. First, any deadweight loss associated with monopolistic competition is likely to be small, due to the presence of competing firms and free entry. Put differently, each firm's demand curve is relatively elastic, and so the excess of price over marginal cost is typically small. In the case of pure monopoly, this is not necessarily true. (Note that this excess, $P - MC$, is the height of the deadweight loss triangle, AB, in Figure 13.2.) For the same reason, the cost associated with excess capacity will also tend to be small.

Second, and, perhaps, most important, any possible inefficiency cost must be weighed against the product variety produced by monopolistic competition and the benefits of such variety to consumers. Similarly, it is probably desirable for firms to continue to have a dynamic incentive to introduce new differentiated products that better satisfy consumer tastes, and that incentive could be undermined by regulation.

Third, any sort of intervention has its own costs, which must also be balanced against the potential gain from expanding output. The costs of operating a regulatory agency may exceed the noted deadweight loss associated with monopolistic competition. Moreover, regulators can find it difficult to obtain the information necessary to achieve a more efficient output and mistakes may be made.

Application 13.1 Ready-to-Regulate Ready-to-Eat

In the early 1970s, the Federal Trade Commission (FTC) initiated antitrust proceedings against the three leading ready-to-eat (RTE) cereal manufacturers. The three manufacturers accounted for more than 80 percent of RTE cereal sales and had been instrumental in increasing the number of nationally marketed brands from 27 in 1950 to 74 in 1971. The FTC argued that the product proliferation effectively precluded new entry and relied upon extensive advertising aimed at overly impressionable customers (kids) to differentiate cereals. While the typical industry devoted less than 1 percent of its sales revenue to advertising, the RTE cereal industry allocated more than 11 percent and ranked among the top five industries in terms of advertising intensiveness as of 1977. This is what one would expect from a monopolistically competitive market in which firms compete by varying (and advertising) the nature of their products in an effort to attract customers. The FTC proposed breaking up the three top RTE manufacturers into eight more evenly matched firms, regulating industry expenditures on advertising, and requiring licensing of significant cereal formulas and trademarks.

Although the FTC alleged that there was little fundamental difference between RTE cereal brands, manufacturers argued that the grain bases, shapes, flavors, nutritional values, and so on of the various brands reflected vigorous competition and the desire to better satisfy consumers' preferences for diversified breakfast fare. Furthermore, manufacturers argued that the pricing discretion afforded to any individual brand by product proliferation was minor—that is, the elasticity of demand for any particular brand was fairly high.

After spending several million dollars to prosecute the case, the FTC dropped its proceedings in 1982. The reversal partly resulted from the election of Ronald Reagan as president in 1980 and the appointment of more pro-business commissioners to the FTC. Moreover, the rapid growth in the late 1970s of health-oriented cereals that featured ingredients such as oat bran and were marketed by smaller firms, as well as growth in the number of house brand cereals sold by supermarkets, openly contradicted the FTC claims that product proliferation by the major cereal makers deterred new entrants to the industry.

Application 13.2 Monopolistic Competition Is in the Eye of the Beholder

Refractive eye surgery has become very popular in recent years.[3] Nearly 1 million Americans undergo the procedure each year in order to correct their vision. As the refractive eye surgery industry has grown, it has evidenced all the characteristics of monopolistic competition. Entry and exit into the industry is relatively unrestricted. There are a large number of independent sellers who do not produce a homogeneous product. For example, under the Lasik procedure, a surgeon creates a flap in the eye, then uses a laser on the area underneath to correct the vision. PRK, another form of laser eye surgery, consists of a surgeon using a laser on the eye's surface to correct vision. Some patients opt for corneal rings, prescription inserts that are intended to correct mild nearsightedness.

The various sellers of refractive eye surgery services tout the advantages of their differentiated product over what rivals are offering. It has been estimated that surgery centers are spending $200 per each eye corrected on advertising. For example, TLC Laser Eye Centers rely on Tiger Woods to advertise their service. However, Dr.

[3]This application is based on Randy Tucker, "Cost Cuts Debated by Doctors: Surgery Often 'On Sale'," *Cincinnati Enquirer*, November 15, 2000, p. B10; "Turning Surgery Into a Commodity," *New York Times*, December 9, 2000, pp. B1 and B4; and "Imperfect Vision," *ABCNEWS.com*, July 31, 2001.

Penny Asbell, Director of the Cornea Service and Refractive Surgery Center in New York, urges prospective customers to be cautious when evaluating such promotions. Asbell notes that, "Just because someone is advertising doesn't necessarily mean that they're more qualified." She recommends relying on a surgeon associated with an academic medical center, such as a teaching hospital or one that is well known for advanced technology.

Dr. Steve Updegraff, director of Updegraff Lasik Vision, recommends choosing a doctor belonging to the American College of Surgeons. "The credentialling process there is pretty steep; also that group is diligent about advancing the field of surgery." Dr. Updegraff says that when something goes wrong during the flap cutting stage of Lasik, some less experienced surgeons may go ahead and perform the tissue removal anyway, instead of stopping surgery and trying again at a later date. He says that this is one reason for poor results.

Finally, sellers of refractive eye surgery services appear to also compete on price and respond to profit-based incentives for entry and exit. Whereas, the cost of laser eye surgery was as high as $6,000 in the late 1990s, it had fallen to less than $1,000 for surgery on both eyes as of 2002. Lasik Vision has opened centers across the United States in recent years. When Lasik Vision entered the market in Cincinnati, Ohio, with an introductory offer of $1,000 for both eyes, its chief existing rival in town, LCA-Vision, immediately matched its price.

13.2 OLIGOPOLY AND THE COURNOT MODEL

OLIGOPOLY
an industry structure characterized by a few firms producing all or most of the output of some good that may or may not be differentiated

Oligopoly is an industry structure characterized by a few firms producing all, or most, of the output of some good that may or may not be differentiated. The number of competitors is the distinguishing feature of this market structure. With competition (and monopolistic competition), there are "many" sellers, whereas with pure monopoly there is only one seller. Oligopoly falls between these extremes. In the United States, there are a number of examples of oligopolistic industries, including aluminum, cellular phone service, network television, and military aircraft.

When there are a small number of competitors, their market decisions will exhibit *strong mutual interdependence*, and this characteristic of oligopoly is what makes analysis of it difficult. By mutual interdependence, we mean that a firm's actions (setting price, for example) have a noticeable effect on its rivals, and so they are likely to react in some way. In this way, the firms are interdependent. As an example, suppose that General Motors (GM) is considering a 10 percent cut in the price of its Buick line. This action will have a significant effect on Ford and DaimlerChrysler. If they maintain their prices, they will lose sales to GM. If they cut prices, they can avoid losing sales, but they will make a smaller profit per car. To complicate matters further, Ford and DaimlerChrysler have the options of cutting prices by more or less than the 10 percent cut by GM.

Now consider what this situation means for GM: the results of its own 10 percent price cut cannot be predicted without knowing how its rivals will respond. For instance, GM's sales will rise more if Ford and DaimlerChrysler maintain their prices than they would if those companies also reduce their cars prices. GM must base its decisions on some guess, or conjecture, about its competitors' responses. What guess should it make? The problem for GM, and also for us as we try to understand the market, is that it is far from clear which prediction is appropriate. The market functions differently depending on which predictions about responses each firm makes and acts on. Furthermore, over time the firm may learn that some of its predictions were wrong and alter its behavior accordingly. But its competitors will also be learning and trying to outguess it. Complex questions of strategy arise in this setting.

In view of this complicated interdependence, it is perhaps not surprising that we do not have one agreed-upon theory of oligopoly. In fact, dozens of models have been suggested. Some of them appear to successfully explain the behavior in *some* industries over *some* periods of time, but none appears to explain all oligopolistic behavior. We will discuss a few of the more important models that have been developed, but be forewarned that determining when each model applies (if at all) is often difficult.

In addition to smallness in the numbers of competitors, there are two other features of oligopoly that are likely to have a bearing on how the market performs. First is whether the product is homogeneous or differentiated. Some oligopolies produce a homogeneous product, like aluminum or steel, while others produce differentiated products, like diapers or airline service in smaller city-pair markets. When the product is differentiated, advertising (which we will discuss in more detail in the following chapter) is likely to become a more important influence in the market.

A second important oligopoly feature is the nature of barriers to entry, if any. Oligopolistic firms are often thought to realize economic profits, and whenever there are profits there is incentive for entry. Something must impede entry for profits to persist. Moreover, just as in the case of monopoly (see Chapter 11), potential entry can influence oligopolists' pricing behavior.

The Cournot Model

We begin our discussion of oligopoly by considering one of the earliest models, introduced by French economist Augustin Cournot in 1838.[4] Cournot considered a **duopoly**, an industry with just two firms. To illustrate his analysis, Cournot assumed that the two firms sold water from the only two mineral springs in the area. To follow tradition, we will consider two firms that sell bottled water, Artesia and Utopia. No entry of new firms is possible. The bottled water is a homogeneous product, so that only one price can prevail in the market; that price is determined by the *combined* output of the two firms in conjunction with the market (industry) demand curve for bottled water. To further simplify the analysis, we assume that both firms have constant and equal long-run marginal cost curves, and that the market demand curve is linear.

DUOPOLY
an industry with just two firms

The key element in the **Cournot model** is that *each firm determines its output based on the assumption that any other firms will not change their outputs*. This assumption (and it may be an unreasonable one, as we will see) allows us to determine the market price and output. To see how we can do this, consider Figure 13.3, where the *market* demand and marginal revenue curves are shown as *D* and *MR*, and each firm's marginal cost (*MC*) and average total cost (*ATC*) curves are assumed to be constant. Now let's examine Artesia's output decision. Artesia's most profitable output will depend on how much Utopia is producing, so first we consider how much Artesia will produce for each possible output of Utopia.

COURNOT MODEL
a model of oligopoly that assumes each firm determines its output based on the assumption that any other firms will not change their outputs

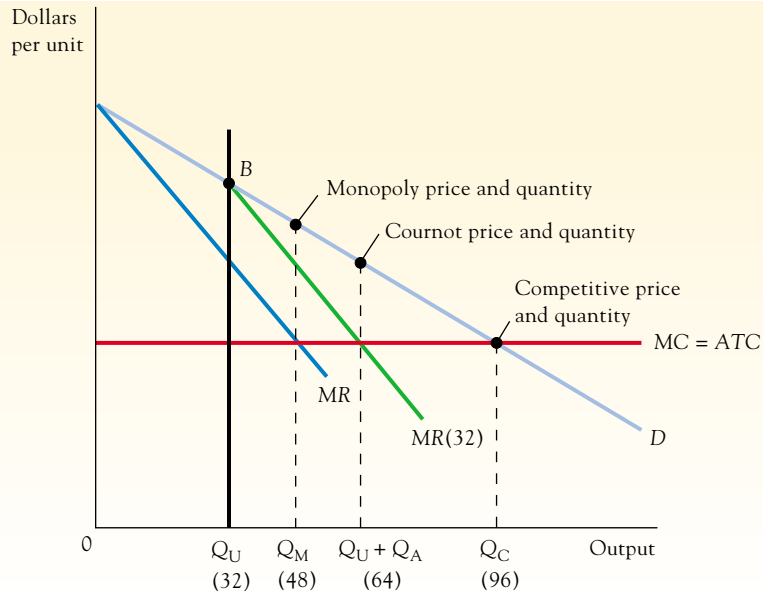
Suppose that Utopia produces nothing. In the Cournot model, Artesia assumes Utopia will continue to produce nothing whatever output Artesia chooses. In this situation, Artesia confronts the entire market demand curve and behaves as a monopolist, producing Q_M (48), where Artesia's marginal revenue curve (the same here as the market marginal revenue curve) intersects marginal cost. In the analysis that follows, it will be helpful to remember that with linear demand and constant marginal cost, the marginal revenue curve intersects marginal cost at an output half as large as that at which marginal cost intersects the demand curve. In this case, Artesia's output of 48 is half as large as the output that

[4]Augustin Cournot, *Réchêrches sur les Principes Mathématiques de la Théorie des Richesses* (Paris, 1838), trans. Nathaniel Bacon (New York: Macmillan, 1897).

FIGURE 13.3

The Cournot Model

When Utopia's output is 32, the vertical axis relevant for Artesia's output decision is BQ_U, and Artesia's demand curve is the *BD* portion of the market demand curve. Artesia's marginal revenue curve is then *MR*(32), and its most profitable output is 32, so combined output is 64.



would be produced under competition, 96, as shown by the intersection of demand and marginal cost.

Artesia's output depends on how much Utopia produces. We have just seen that Artesia will produce 48 units when Utopia produces nothing. Alternatively, suppose that Utopia produces 32 units. Then how much will Artesia produce? Artesia believes Utopia will continue to produce 32 units regardless of how much Artesia produces (and thus regardless of what happens to the market price, which will be determined by the two firms' combined output). At each price, Artesia can sell 32 fewer units than total quantity demanded as shown by the market demand curve. So Artesia's demand curve is the market demand curve shifted leftward by 32 units. This idea can be shown in a simpler, yet equivalent fashion by moving Artesia's vertical axis rightward by 32 units without repositioning the demand curve. Taking the origin for Artesia now to be Q_U, the demand curve confronting Artesia is the *BD* portion of the original demand curve. This makes sense. If Artesia produces nothing, total output will be 32 (Utopia's output), price will be BQ_U, and as Artesia produces and adds to Utopia's output, price will fall along the *BD* portion of the demand curve.

Confronted with the demand curve *BD*, Artesia's marginal revenue curve is *MR*(32), the marginal revenue curve when Utopia's output is fixed at 32. In this situation, Artesia produces where its marginal revenue curve, *MR*(32), intersects marginal cost; thus Artesia's output is 32 units, while the total output of the two firms is 64 units. Note that Artesia's output is half the difference between the competitive output (96) and Utopia's output (32); this is because *MR*(32) intersects MC halfway between the new vertical axis for Artesia and the output at which MC intersects the demand curve.

We can now see how Artesia's output depends on how much Utopia produces. For each possible output by Utopia, Artesia will produce half the difference between Utopia's output and the output at which MC intersects *D* (96 units). If Utopia produces nothing, Artesia will produce 48; if Utopia produces 10, Artesia will produce 43; if Utopia produces 20, Artesia will produce 38; and so on. Now that we know what Artesia will do, what about Utopia? Because the firms have the same costs and because we also make the Cournot assumption for Utopia (that is, it will take Artesia's output as a given in determining its output), the same

relationship holds for Utopia. In other words, Utopia will produce 48 units if Artesia produces nothing, 43 if Artesia produces 10, and so on.

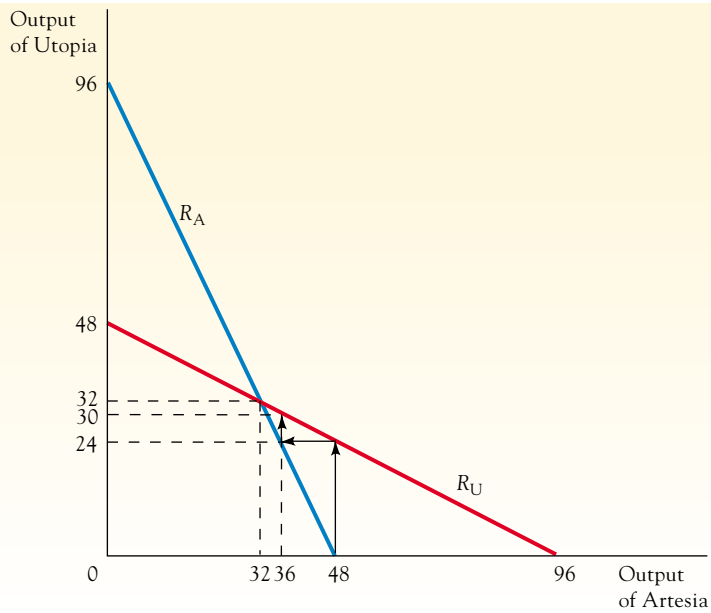
So where will the market equilibrium be? Equilibrium is reached when neither firm has any incentive to change its output. This occurs when each firm is producing the output it prefers given the other firm's output. In this example, that occurs only when both firms produce 32 units. To check this, we note in Figure 13.3 that Artesia's most profitable output when Utopia produces 32 units is also 32 units. Because Utopia has the same marginal cost curve, it will also maximize profit by producing 32 units when Artesia produces 32. Neither firm has any incentive to change its output of 32 when the other firm is producing 32. (The implication of equal output here arises because the firms have the same costs; if costs differ, outputs will differ, but the reasoning remains the same.)

There is another way to arrive at this conclusion—by using reaction curves. Each firm's **reaction curve** shows its profit-maximizing output for each possible output by the other firm. In fact, we have already explained the relationships above. In Figure 13.4, R_A is Artesia's reaction curve. It shows that Artesia will produce 48 units when Utopia's output (measured on the vertical axis) is zero, will produce 36 when Utopia's output is 24, and so on. Utopia's reaction curve is R_U; it is the same relationship as for Artesia but looks different in the graph because the firms' outputs are on different axes. We can see how equilibrium can be attained in a step-by-step process, although this should not be thought of as the actual adjustment process, because if both firms started producing 32, there would be no reason for either to change. To begin, if Utopia produces nothing, then Artesia produces 48. When Artesia produces 48, however, we can see by looking directly above 48 to Utopia's reaction curve that it will produce 24. With Utopia producing 24, Artesia would prefer to change its output to 36. And with Artesia producing 36, Utopia will produce 30. The adjustments follow the arrows, and the firms are *both* not satisfied with their outputs until they reach the point where each is producing 32 units. Put differently, the Cournot equilibrium occurs at the *intersection* of the two reaction curves.

REACTION CURVE
a relationship showing one firm's most profitable output as a function of the output chosen by other firms

FIGURE 13.4

The Cournot Model with Reaction Curves
Each reaction curve shows one firm's most profitable output as a function of the other firm's output. For example, when Artesia's output is 48, R_U shows 24 to be Utopia's most profitable output. The Cournot equilibrium is shown by the intersection of the reaction curves, with each firm producing its most profitable output given the output of the other firm.



As depicted in Figure 13.3, the Cournot equilibrium involves a combined output of 64 units by the two firms in the industry. It is important to note that this amount is greater than the pure monopoly output (48 units) and less than the competitive output (96 units). A total output lying between that of pure monopoly and competition is characteristic of most oligopoly models. Some other things are clear from inspection of Figure 13.3. Price exceeds marginal cost, and, because average cost equals marginal cost, price is above average cost and both firms realize economic profit. However, their combined profit is less than the maximum combined profit possible if the firms together produce the monopoly output. This fact is significant—it means that if the firms colluded instead of behaving independently as Cournot duopolists, they could increase their combined profit.

Evaluation of the Cournot Model

Is it reasonable for a firm to assume, in choosing its output, that the output of a rival remains constant? Not in the duopoly setting we have just studied, if the market is still adjusting toward the Cournot equilibrium. Away from the Cournot equilibrium—that is, when Artesia changes its output on the assumption that Utopia will keep its output fixed—it will observe that the assumption is wrong: Utopia does change its output in response to Artesia's actions. Yet at each step in the adjustment process, the firms continue to behave based on an assumption they can see is wrong. Thus, the key assumption of the Cournot model, that each firm takes the other firms' outputs as given, appears to be suspect if the market is still adjusting toward equilibrium.

While this criticism is significant, there are some things that can be said in defense of the Cournot model. First, note that if the equilibrium is somehow established, firms will not see the assumption invalidated. When Artesia sees Utopia producing 32 units, and decides on 32 for itself, based on the assumption that Utopia will not change its output, it will be right. The assumption becomes implausible only for adjustments to the equilibrium.

Second, the assumption is more plausible the larger the number of firms in the market. (The Cournot model can readily accommodate any number of firms greater than two and, in general, the greater the number of firms, the larger the total industry output as a percentage of the competitive output.) With 10 equal-sized firms, if one changes its output by, say, 10 percent, it will represent only a 1 percent change in industry output, which will have a small effect on price. The other firms may not associate such a small price change with the actions of one firm because other things, like shifting market demand, can also affect price.

APPLICATION 13.3 STRATEGIC INTERACTION ON DUOPOLY AIR ROUTES

A 1993 study examined the interaction between American and United Airlines over 1984–1988 on 16 Chicago-based air routes on which the two carriers could reasonably be characterized as a symmetric duopoly.[5] On these routes, the two carriers held a combined market share of over 90 percent, accounted for more than one-third of the total passenger traffic each, and had very similar costs.

The study found that Cournot behavior most frequently characterized the interaction between the two air carriers on the selected duopoly routes. The Cournot outcome, however, did not always prevail. Changes in underlying costs and market demand influenced the strategic interaction between the two carriers. For example, a strike in 1985 by the pilots of United Airlines triggered an apparent price war and an outcome more

[5]James A. Brander and Anming Zhang, "Dynamic Oligopoly Behavior in the Airline Industry," *International Journal of Industrial Organization*, 11 (1993), pp. 407–435.

consistent with competition than with the Cournot model. The strike appeared to have increased the uncertainty each firm had (and thereby assumptions it made) about its rival's costs and strategic intentions.

The study acknowledged that some other factors may explain why, in terms of total output, the most frequently observed market outcome (consistent with Cournot behavior) lies between the pure monopoly and the perfectly competitive outcome. For example, because American and United Airlines interact in more than one market (that is, on different routes), such multiple points of contact may serve to restrain competition between the two firms on any given route. In other words, American may be wary of aggressively expanding output and lowering price on the Chicago–Indianapolis route for fear that United may retaliate in kind across all the routes on which the two airlines compete.

On the other hand, while only two carriers may operate on any particular route, there is always the possibility of new entry. Just as we saw with monopoly in Chapter 11, the possibility of entry can strongly affect the operation of a market, and the same is certainly true in oligopoly markets. If American and United recognize that entry will occur if the price rises too much above cost, it may influence their output decisions. The threat of entry thus serves to push the observed market output away from the pure monopoly outcome and closer to the competitive outcome.

13.3 Other Oligopoly Models

The Cournot model serves as a good introduction to oligopoly models by highlighting the importance of how firms handle the mutual interdependence in such markets. In this section we explore two other models of oligopoly. Although they by no means represent all the models that have been suggested, they do indicate some different assumptions a firm in an oligopoly market might make about rival firms' actions.

The Stackelberg Model

Stackelberg model
a model of oligopoly in which a leader firm selects its output first, taking the reactions of follower firms into account

Residual demand curve
a firm's demand curve based on the assumption that the firm knows how much output rivals will produce for each output the firm may choose

Recall that in the Cournot model each firm takes other firms' outputs as constant in determining its own output. We saw, however, that this assumption may not be valid. So now suppose that in the same two-firm example, we have one firm that continues to behave in the naive Cournot fashion, while the other firm wises up and realizes that it should not assume its rival's output doesn't change. In fact, let's assume that Artesia realizes how Utopia chooses its output (from its reaction curve) and see whether Artesia can use that information to realize greater profit. Artesia is the "leader" firm in this case; it chooses its best output taking Utopia's reaction into account. Utopia is the "follower" firm; it selects output in exactly the same way as in the Cournot model, by taking the output of the other firm as given. This is the essence of the **Stackelberg model**: a leader firm selects its output first, taking the reactions of naive Cournot follower firms into account.[6]

Figure 13.5 illustrates the Stackelberg model as it operates for Artesia and Utopia. The marginal cost, average total cost, and market demand curves are shown in Figure 13.5a; they are the same as in Figure 13.3. Figure 13.5a shows how Artesia selects its output. Given Artesia's output, Utopia's output can be read off its reaction curve, R_U, reproduced directly below in Figure 13.5b. Remember that we are assuming that Artesia knows Utopia's reaction curve, so that it knows how much output Utopia will produce for each output Artesia may choose.

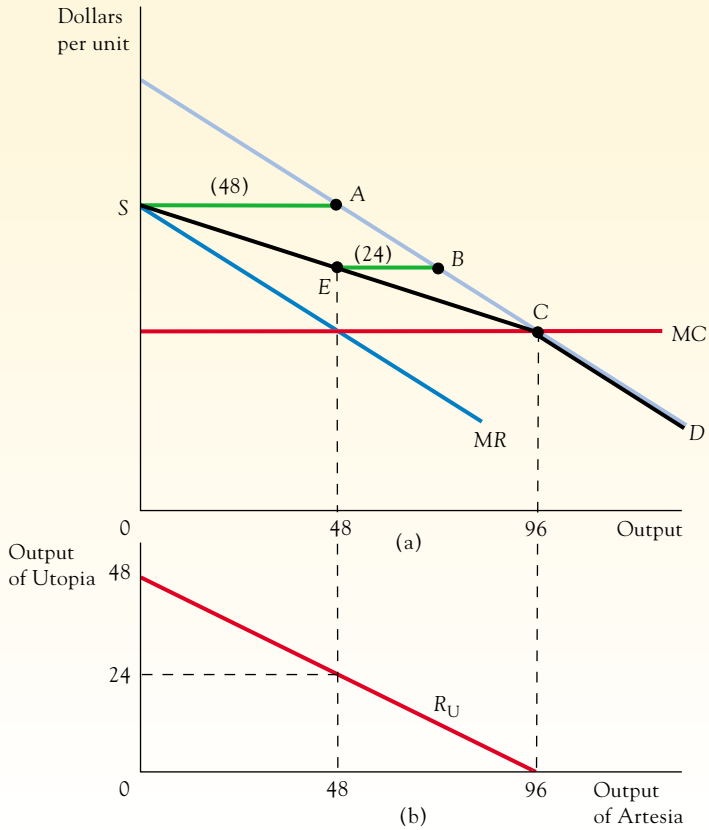
Our first task is to determine Artesia's demand curve under these conditions. This will not be the market demand curve, but what is referred to as a **residual demand curve**, which shows how much Artesia can sell at each price. The amount that Artesia can sell at each

[6]Heinrich von Stackelberg, *Marktform und Gleichewicht* (Vienna: Julius Springer, 1934). As with the Cournot model, the Stackelberg model can readily be adapted to account for a larger number of firms.

FIGURE 13.5

The Stackelberg Model

When Artesia knows that Utopia will choose output as in the Cournot model, Artesia confronts the demand curve *SCD*, and its most profitable output is 48. Utopia will produce 24, so total output is 72, higher than when both firms behave as Cournot duopolists.



price is less than total quantity demanded (as shown by the market demand curve) by the amount that Utopia produces. For example, suppose that Artesia produces zero. From Utopia's reaction curve, Artesia knows that Utopia will then produce 48 units. Thus, total (combined) output is 48 units when Artesia produces nothing, and the market price in Figure 13.5a will be *S*. At the other extreme, if Artesia produces 96 units, Utopia will produce nothing, and Artesia will be at point C in Figure 13.5a. That gives two points on Artesia's residual demand curve, *S* and *C*. Artesia's residual demand curve (with the assumed linear demand and cost conditions) is just the straight line connecting these points between outputs of zero and 96 units. Beyond 96 units of output, Artesia's residual demand curve coincides with the market demand curve (along *CD*), since Utopia will produce zero if Artesia produces in excess of 96 units.

To see that straight-line segment *SC* represents Artesia's residual demand curve for outputs between zero and 96 units, suppose that Artesia produces 48 units. From Utopia's reaction curve, Artesia knows that Utopia will produce 24, so total output will be 72. When total output is 72, the price is given by point *B* on the market demand curve. Thus, Artesia can sell 48 units when price is at the height of point *E* (the same height as point *B*), which gives us a third point on Artesia's demand curve. Note that the horizontal distance between Artesia's demand curve and the market demand curve is Utopia's output. As you can see, Utopia's output becomes smaller as Artesia increases output along its demand curve. In fact, for each one-unit increase in output by Artesia, Utopia reduces output by one-half

unit (as can be seen from Utopia's reaction curve), so the two firms' total output increases by only half as much as the increase by Artesia. That is why price declines less rapidly along Artesia's residual demand curve than along the market demand curve (Artesia's demand curve is flatter; in fact, the slope is exactly half the slope of the market demand curve).

With knowledge of its demand curve, profit maximization by Artesia is straightforward. With demand curve *SCD*, the marginal revenue curve is *MR*, intersecting the marginal cost curve halfway between zero output and the output where marginal cost intersects demand. Therefore, Artesia's profit-maximizing output is 48 units, with price shown by the height of point *E*. Utopia is producing 24 units, so total industry output is 72 units.

Because we are using the same demand and cost conditions as we did with the Cournot model, it is instructive to compare the outcomes. Note that total output is higher with the Stackelberg model (72 versus 64), so price to consumers is lower. Output is closer to the competitive result than in the Cournot model, but still lies between the competitive and monopoly outputs. In addition, Artesia is making a larger profit and Utopia a smaller profit than in the case of a Cournot equilibrium. (This is not shown in the graphs but is easily verified.) This outcome is to be expected: Artesia is exploiting its superior knowledge of how Utopia will respond to make a larger profit at Utopia's expense.

Our discussion highlights a key point: namely, that the conjectures a firm makes in an oligopoly market about how its rivals will respond can affect firms' outputs and profits as well as total industry output. For example, total industry output is higher in a Stackelberg model than in a Cournot model. And the firm that is a Stackelberg leader can take advantage of its leadership position to set a larger (firm) output, thereby enhancing its profit at the expense of firms that follow its lead in naive Cournot fashion.

Whether the Stackelberg or Cournot model better describes an oligopoly depends on the particular market being examined. Where an oligopoly is composed of roughly equal-sized firms, none with superior knowledge or exercising a leadership position, the Cournot model is likely to be more apt. However, when one firm is more sophisticated about how rival firms will react and uses this information to operate as a leader in terms of output, pricing, and/or the introduction of new products, the Stackelberg model is more appropriate. The leadership role played by Intel in terms of setting price and introducing new products in the computer chip market over the last decade provides a possible example of the latter case.

The Dominant Firm Model

In the Stackelberg model, the leader firm assumes that rivals display Cournot behavior and plans its output and price accordingly. We now will examine an alternative model in which the leader firm makes a different conjecture about the behavior of rival follower firms. In this model, known as the **dominant firm model**, the leader or dominant firm assumes that its rivals behave as competitive firms in determining their output. (Sometimes this model is referred to as the *dominant firm with a competitive fringe* model because the competitive firms are on the fringe.) The dominant firm model has been used by economists to analyze the performance of many industries.

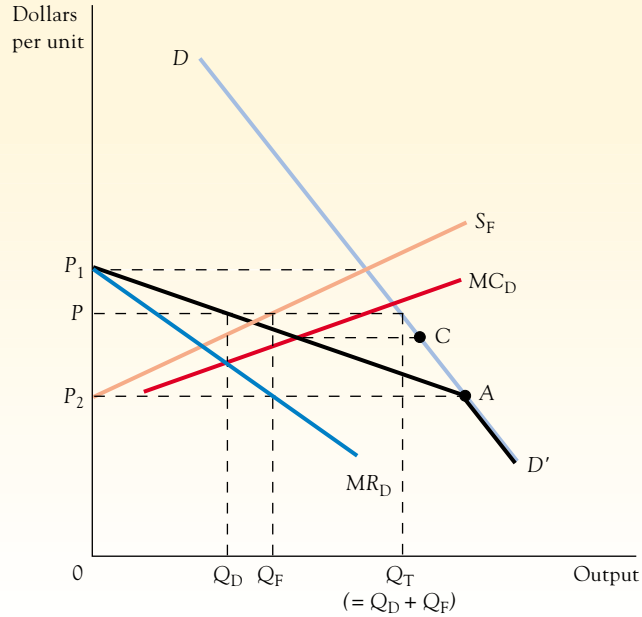
DOMINANT FIRM MODEL
a model of oligopoly in which the leader or dominant firm assumes its rivals behave like competitive firms in determining their output

Figure 13.6 shows how this market structure operates. To determine what price will maximize its profit, the dominant firm must know its demand curve. As with the Stackelberg model, the dominant firm's demand curve is a residual demand curve that shows what it can sell after accounting for other firms' output. In this case, the other firms in the market are assumed to behave as competitive firms: they will accept whatever price is set by the dominant firm and produce an output where their marginal cost equals that price. The output of the competitive fringe firms can therefore be determined from their supply curve because they collectively behave as a competitive industry. This supply curve is shown as S_F in the diagram. The market demand curve is *DD′*.

FIGURE 13.6

The Dominant Firm Model

With the supply curve of fringe firms shown as S_F, the residual demand curve of the dominant firm is derived by subtracting the quantity supplied by fringe firms at each price from total quantity demanded at that price; the result is curve P_1AD'. The dominant firm maximizes profit by producing Q_D and charging price P; fringe firms produce Q_F at that price, so total output is Q_T.



At any price, the dominant firm can sell an amount equal to the total quantity demanded at that price (as shown by DD′*) minus the quantity the fringe firms produce (as shown by* S_F*).* For example, the dominant firm's demand curve begins at P_1 because at that price the fringe firms will supply as much as consumers wish to purchase, and the dominant firm could sell nothing. At the other extreme, if the dominant firm charges a price less than P_2, it faces the entire market demand curve because the fringe firms will produce nothing at such a low price. Between P_1 and P_2, the dominant firm's residual demand curve is P_1A, where the horizontal distance between this demand curve and the market demand curve at each possible price shows the output of the fringe firms.

Armed with a knowledge of its demand curve, P_1AD', the dominant firm also knows its marginal revenue curve, MR_D, and maximizes profit by producing where marginal revenue equals marginal cost. With a marginal cost curve shown as MC_D, the dominant firm's profit-maximizing output equals Q_D. The price is P, the height of the dominant firm's residual demand curve (not the market demand curve) at output Q_D. At price P, other firms produce Q_F as shown by their supply curve, and total output, Q_T, is the sum of their output and the dominant firm's output. At price P, consumers wish to purchase an output of Q_T, and so the market is in equilibrium. At the equilibrium, note that price is above marginal cost for the dominant firm, but it is equal to marginal cost for the fringe firms, S_F. This implies that total output is less than if the industry were competitive. The competitive output for the dominant firm is where MC_D intersects the residual demand curve; at that point both it and the other firms are producing where marginal cost equals price. Total output and price under competitive conditions are indicated by point C on the market demand curve.

One interesting implication of this model is that the share of total industry output produced by the dominant firm may not indicate how close output comes to the competitive result. For example, suppose that the supply curve of the fringe firms is perfectly elastic (as with a constant-cost competitive industry) at price P: the supply curve coincides with the horizontal dotted line in the graph. Then the dominant firm's residual demand curve is also given by this horizontal dotted line, and marginal revenue will equal P out to output Q_T.

The dominant firm will produce where its marginal cost curve intersects this horizontal line. Industry output will be the same, Q_T, but now the dominant firm is producing about 90 percent of it. Furthermore, price is equal to marginal cost, as under competition, even though one firm is contributing 90 percent of total output. This example illustrates the critical importance of the elasticity of the competitive firms' supply curve for the functioning of this sort of market structure.

Recall that this model differs from the Stackelberg model only in what the leader, or dominant, firm assumes about rival firms' output. In the Stackelberg model, the leader firm assumes Cournot behavior on the part of rivals; in this model, it assumes competitive behavior. The dominant firm model is more appropriate when there are a sufficiently large number of fringe firms for the assumption of competitive behavior to be plausible.

The Elasticity of the Dominant Firm's Demand Curve

Based on the fact that the dominant firm's output is equal to the total market output minus the quantity the fringe firms' supply, we can derive the dominant firm's elasticity of demand as follows:

$$\eta_D = \eta_M\left(\frac{1}{MS}\right) + \varepsilon_{SF}\left(\frac{1}{MS} - 1\right); \tag{1}$$

where η_D is the elasticity of the dominant firm's demand; η_M is the elasticity of the market demand; MS is the dominant firm's market share; and ε_{SF} is the elasticity of supply of the fringe firms.[7] To see how to apply the formula, consider the case of the pharmaceutical firm Hoffman-La Roche (Roche for short), whose brand-name product Valium is the market-leading anti-anxiety drug. Suppose that Roche can be taken to be the dominant firm in the anti-anxiety market, that it has a 25 percent market share, that it faces a competitive fringe of firms that produce the generic equivalent of Valium, that the elasticity of supply by the competitive fringe is equal to 2, and that the elasticity of market demand for anti-anxiety drugs is equal to 1. Using these assumptions, we can calculate the elasticity of demand for the dominant firm's product, Valium, as follows:

$$\eta_D = 1\left(\frac{1}{0.25}\right) + 2\left(\frac{1}{0.25} - 1\right) = 10.$$

Even though its output is equal to one-fourth the entire market output, Roche faces a residual demand with an elasticity of 10. Thus, if the company raises Valium's price by just 5 percent, it will lose half its sales.

The formula for the dominant firm's demand elasticity shows that the demand elasticity becomes larger when (1) the dominant firm's market share becomes smaller, (2) the

[7]To derive this formula, we start with the fact that the dominant firm's output (Q_D) equals the market output (Q_M) minus the output of the competitive fringe (Q_{SF}):

$$Q_D = Q_M - Q_{SF}.$$

This relationship also holds for a given change in output that results from a price change:

$$\Delta Q_D = \Delta Q_M - \Delta Q_{SF}.$$

Now divide by Q_D and multiply the two terms on the right by Q_M/Q_M and Q_{SF}/Q_{SF}, respectively:

$$\frac{\Delta Q_D}{Q_D} = \left(\frac{\Delta Q_M}{Q_M}\right)\left(\frac{Q_M}{Q_D}\right) - \left(\frac{\Delta Q_{SF}}{Q_{SF}}\right)\left(\frac{Q_{SF}}{Q_D}\right).$$

Dividing this expression by $\Delta P/P$ yields the formula in the text. Note that the minus sign on the right-hand side became a plus sign because we are treating the elasticity of demand as a positive number; Q_M/Q_D equals $1/MS$; and Q_{SF}/Q_D equals $(Q_M - Q_D)/Q_D$ or $(1/MS) - 1$.

elasticity of the market demand becomes greater, and (3) the elasticity of supply by the competitive fringe becomes greater. For example, if Roche's market share was 10 percent instead of 25 percent, the elasticity of demand for Valium would be greater: $28 = [1(1/0.1) + 2((1/0.1) - 1)]$ versus the 10 already calculated. If the elasticity of the demand for anti-anxiety drugs was 5 instead of 1, the elasticity of demand for Valium would be $26 = [5(1/0.25) + 2((1/0.25) - 1)]$ instead of 10. And if the elasticity of the fringe supply was 5 instead of 2, the elasticity of demand for Valium would be $19 = [1(1/0.25) + 5((1/0.25) - 1)]$ instead of 10.

APPLICATION 13.4 THE DYNAMICS OF THE DOMINANT FIRM MODEL IN PHARMACEUTICAL MARKETS

As the patent on a brand-name pharmaceutical expires, the producer of the drug typically confronts competition from generic manufacturers. Generic manufacturers do little research of their own; rather, they specialize in copying brand-name products after their patents expire. Generic manufacturers tend to become both more numerous and more capable of expanding their output capacity the longer that a brand-name drug is "off-patent." In such a setting, therefore, the brand-name drug producer can be taken to be the dominant firm, with its market share decreasing and the competitive fringe's elasticity increasing the more years the brand-name drug is off-patent.

What does the dominant firm model predict about the price charged by a brand-name drug maker and the sensitivity of consumers to the brand-name drug's price as the number of years that the drug has been off-patent increases? As the fringe supply curve tends to shift rightward (see Figure 13.6) as the time since the brand-name drug maker's patent expired increases, it works to shift the dominant firm's residual demand curve leftward and put downward pressure on the price charged by the brand-name drug maker. Moreover, as both the brand-name maker's market share decreases and the elasticity of the fringe supply increases, equation (1) indicates that the demand elasticity facing the brand-name maker tends to increase with the time since patent expiration. The available empirical evidence bears out these theoretical predictions generated by the dominant firm model.[8]

The empirical evidence also suggests that after patent expiration, brand-name makers pursue a market segmentation strategy—charging a lower price to hospitals and health maintenance organizations than to retail pharmacies. Hospitals and health maintenance organizations are more sensitive to price owing to their large volume of purchases and their greater knowledge about the (characteristically small) risks of substituting a generic drug for a brand-name product. Market segmentation, however, has not been without its costs. Retail pharmacists have sued pharmaceutical companies for conspiring to deny them the same price discounts offered to hospitals and health maintenance organizations. Brand-name drug makers opted to settle one such suit in 1996 for $551 million.[9]

[8]Richard Caves, Michael Whinston, and Mark Horwitz, "Patent Expiration, Entry, and Competition in the U.S. Pharmaceutical Industry," *Brookings Papers on Economic Activity: Microeconomics*, (1991), pp. 1–48.

[9]"Judge Agrees to Settlement in Drug Case," *New York Times*, June 22, 1996, p. 17.

13.4 CARTELS AND COLLUSION

In the oligopoly models we have examined so far, individual firms were assumed to behave independently. Each firm makes a specific conjecture regarding how other firms will respond to its actions and then maximizes its own profit accordingly, without any concern for how it affects other firms' profits. An alternative class of oligopoly models is based on various types

of cooperation among firms. The firms coordinate their pricing and output decisions in an attempt to increase their combined profit, thereby increasing their individual profits as well.

CARTEL
an agreement among independent producers to coordinate their decisions so each of them will earn monopoly profit

The most important cooperative model of oligopoly is the cartel model. A **cartel** is an agreement among independent producers to coordinate their decisions so each of them will earn monopoly profit. Because cartels are illegal under the antitrust laws in the United States (though, surprisingly, not in many other countries), they are not common here. There have been a number of international cartels, however, and we examine one of the most famous, OPEC, later in this section. Familiarity with the cartel model is useful, because collusive practices that fall short of outright cartel agreements can be investigated with it. We begin by considering what happens if firms in a competitive industry form a cartel, and then extend the results to oligopolistic markets.

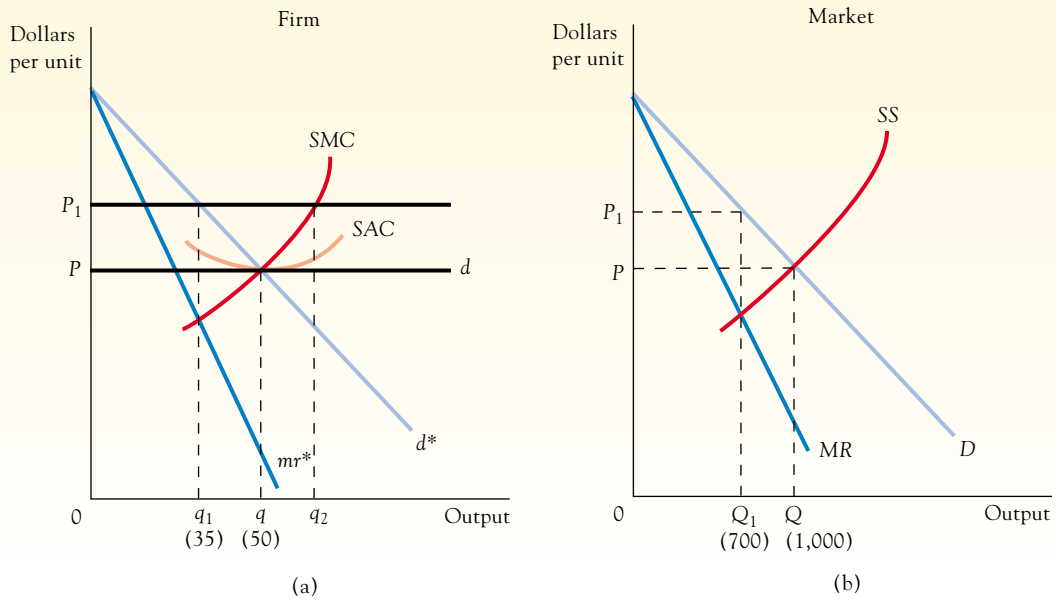
Cartelization of a Competitive Industry

Let's see how a group of firms in a competitive market can earn monopoly profits by coordinating their activities. We assume that the industry is initially in long-run equilibrium, and then we identify the short-run adjustments (with existing plants) that the industry's firms can make to reap monopoly profits for themselves. Figure 13.7b shows the industry equilibrium with a price of P and an output of 1,000 units. Figure 13.7a shows the competitive equilibrium for one of the firms in the industry. Note that initially, the firm faces the horizontal demand curve d at the market-determined price and produces an output of 50 units. To simplify matters, suppose that there are 20 identical firms in the industry, each producing 50 units of output.

FIGURE 13.7

A Cartel

Under competitive conditions industry output is Q and price is P. If the firms in the industry form a cartel, output is restricted to Q_1 in order to charge price P_1, the monopoly outcome. Each firm produces q_1 and makes a profit at price P_1.



Next, the firms form a cartel and agree to restrict output to attain a higher price. Each firm agrees to produce an identical level of output, equal to one-twentieth of total industry output because there are 20 firms. The cartel agreement has the effect of changing the demand curve facing each firm. Before the agreement, if one firm alone reduced output, its action would not appreciably affect price, as shown by the firm's horizontal demand curve *d*. Now, however, other firms match a restriction in output by one firm, so when one firm cuts output by 15 units, all firms match the reduction, industry output falls by 300 units, and price rises significantly. The demand curve showing how price varies with output when firms' output decisions are coordinated in this way is the downward-sloping curve *d** in Figure 13.7a. At any price, the quantity on the *d** curve is 1/20 that on the industry demand curve.

Faced with this downward-sloping demand curve, the firm's profit-maximizing output occurs where its short-run marginal cost curve *SMC* intersects the new marginal revenue curve *mr**. Output is 35 units, and because all 20 firms reduce production to the same output level, total output falls to 700 units and the price rises to P_1. Each firm is now making an economic profit. Indeed, *the idealized cartel result is just the same as if the industry were supplied by a monopoly that controlled the 20 firms*. Figure 13.7b illustrates the result, with the short-run supply curve *SS* (the sum of the *SMC* curves of the firms) intersecting the industry marginal revenue curve *MR* at an output of 700 units and a monopoly price of P_1. By forming a cartel and restricting output to achieve the monopoly equilibrium, the firms maximize their combined profit. Figure 13.7b shows the total market effect of the coordinated output reduction by the 20 firms; Figure 13.7a shows the effects on each firm individually.

Firms can always make a larger profit by colluding rather than by competing. Acting alone, competitive firms are unable to raise price by restricting output, but when they act jointly to limit the amount supplied, price will increase. As we will see in the next subsection, however, achieving a successful cartel in practice is not as simple as it may seem.

APPLICATION 13.5 WILL THE INTERNET PROMOTE COMPETITION OR CARTELIZATION?

The common wisdom is that the Internet serves to promote competition among suppliers, thereby creating bargains for surfing shoppers.[10] Indeed, a survey by Erik Brynjolfsson and Michael Smith of MIT finds that prices on the Internet are 9 to 16 percent lower than in retail outlets.

Although the Internet lowers the cost of search and thus makes it easier for buyers to shop around for a lower price, Hal Varian of U.C. Berkeley cautions that there is a good reason why the Web might actually result in *higher* prices for consumers. Namely, if there are only a few sellers, the availability of low-cost information about the prices they are charging could make it easier for these sellers to coordinate their pricing through the Web. As a historical example, Varian points to the Joint Executive Committee set up by the major U.S. railroads in the 1880s prior to the enactment of antitrust laws. The Committee served as a cartel by collecting and publishing information about the prices individual railroads charged and their weekly shipments. The railroads often cheated on the published prices by offering lower rates to shippers in secret in exchange for more business. Such cheating would have been mitigated by a public Internet exchange, according to Varian, since each railroad could readily monitor the prices charged by others. Any attempt to offer a lower price could quickly be countered, thereby making the cartel less vulnerable to cheating.

A modern parallel to the activities of the Joint Executive Committee is the manner in which airlines post fares

[10]This application is based on Hal R. Varian, "Online Commerce Creates Strange Competition," *New York Times*, August 24, 2000, p. C2.

online: an airline will announce its rates and associated terms and then watch to see how competitors respond. In the late 1980s airlines began to use online reservation systems to signal their pricing intentions to each other. For example, United would post its intended fare changes at 2 A.M. on Thursday. If rival airlines followed suit by 6 A.M., United's price remained in effect. If not, United would return its price back to the pre-2A.M. level. In 1992, the Department of Justice brought an antitrust suit against several large airlines in an effort to limit the extent to which online reservations systems serve as a signaling device and thereby promote cartelization.

In the long run, the key will be the number of sellers in any online category of goods or services. If there are many sellers, the extra flow of information through the Internet is likely to work to the benefit of buyers, pushing prices down. But in Web-based exchanges where there are only a few sellers and many buyers, the availability of more timely price information may serve to promote cartelization, thereby increasing prices to consumers. The Federal Trade Commission is attempting to set some standards of behavior for online exchange to ensure that they promote competition rather than cartelization.

Why Cartels Fail

If cartels are profitable for the members, why aren't there many more? One reason is that in the United States they are illegal. But even before there were laws against collusive agreements, cartels were rare except when actually supported by government; when they did exist, they were short-lived. Three important factors appear to contribute to cartel instability.

1. *Each firm has a strong incentive to cheat on the cartel agreement.* A cartel achieves monopoly profit through its members restricting output below the levels that each would individually choose, a reduction that results in a higher price. Once a higher price is achieved, individual cartel members could earn even more profit by expanding output. Each firm would like to enjoy the cartel's benefit—a higher price—without incurring the cost—lower output. If only one firm expands output, price will not fall appreciably, but the additional sales at the monopoly price will add significantly to that firm's profit. It is thus in each firm's self-interest to violate the cartel agreement to restrict output.

Figure 13.7 illustrates the incentive to cheat on the cartel agreement. If the firm in Figure 13.7a adheres to the cartel agreement, it will produce q_1 and sell at price P_1. However, note what happens if the firm expands output beyond q_1 while the other firms continue to abide by the cartel agreement's restrictions. In this event the firm faces a horizontal demand curve at price P_1; that is, one firm expanding output alone will not affect price. Remember that the downward-sloping demand curve d^* is relevant only for simultaneous expansion and contraction of output by all firms. The firm acting alone can increase its profit significantly by expanding output, since marginal revenue (equal to price with the horizontal demand curve) is above marginal cost at q_1. Profit will be maximized if the firm increases sales to q_2 at the price of P_1.

Every firm has the same incentive to expand output and cheat on the cartel agreement. Yet if many firms do so, industry output increases significantly, and price falls below the monopoly level. It is in each firm's interest to have other firms restrict their output while it increases its own. Every firm's self-interest is therefore a threat to the cartel's survival. To be successful, the cartel must have some means of monitoring and enforcing its agreement.

The foregoing suggests why government backing generally is so essential to ensuring a cartel's stability. Government provides the means of monitoring and enforcing a cartel

agreement. The caviar cartel provides a good example of this.[11] Prior to the collapse of the Soviet Union, the Ministry of Fisheries in Moscow set stringent quotas for the annual sturgeon catch, the source of caviar, one of the world's most expensive delicacies. Close government monitoring limited poaching and illegal dealing in caviar. However, as the Soviet Union disintegrated, four new independent states and two autonomous regions appeared around the Caspian Sea—the location of over 90 percent of the world's sturgeon stocks. Central authority evaporated, and the independent actions of numerous caviar poachers and illegal traders ripped the formerly tightly regulated cartel wide open. Caviar prices plummeted.

2. *Members of the cartel will disagree over appropriate cartel policy* regarding pricing, output, allowable market shares, and profit sharing. In Figure 13.7, we assumed that the firms have identical cost curves, making agreement on the profit-maximizing cartel output and price relatively easy. But when firms differ in size, cost conditions, and other respects, agreement will not come as easily since the firms will have different goals. If, for example, the cartel members' costs differ, they will disagree on what price the cartel should set. The problems become even more acute when the firms must make long-run investment decisions. Every cartel member will want to expand its capacity and share of total output and profit, but not all can be allowed to do so.

These problems are basically political, and no matter what policy the cartel follows, it will reflect a compromise among divergent views. As happens with any compromise, some firms will be unhappy with the outcome, and those firms are all the more likely to refuse to join the cartel or join but violate any cartel agreement on output.

Agreement will also be more difficult the less homogeneous the product. For example, in the United States there are two primary areas in which oranges are grown: Florida and California–Arizona. Through regulations instituted in 1937, the U.S. government (as an exception to antitrust laws) has allowed growers' cartels to control prices and supplies in the two areas. The organization of a growers' cartel has been much more problematic in Florida than in California–Arizona.[12] This reflects the longer growing season and greater varieties of oranges that can be produced in the climate and soil conditions there. Because there are more product dimensions that must be taken into account, Florida orange growers have been less successful at reaching an effective cartel agreement—despite the United States government's official approval of such an agreement.

3. *Profits of the cartel members will encourage entry into the industry*. If the cartel achieves economic profits by raising the price, new firms have an incentive to enter the market. If the cartel cannot block entry of new firms, price will be driven back down to the competitive level as production from the "outsiders" reaches the market. Indeed, if an increase in the number of firms in the market causes the cartel to break down, then price will temporarily fall below the cost of production, forcing losses on the cartel members. The prospect of entry by new competitors eager to share in the profits is probably the most serious threat to cartel stability.

To be successful, therefore, a cartel must be able to get its members to comply with cartel policy (limiting output) and to restrict entry into the market. These tasks are not easily accomplished, and history is strewn with examples of cartels that flourished for a short time only to disintegrate because of internal and external pressures.

[11]"Bootleggers Thrive, Sturgeons Flounder, as Caviar Cartel Splits," *Washington Post,* June 1, 1992, pp. 1 and 8.

[12]Gary D. Libecap and Elizabeth Hoffman, "The Failure of Government-Sponsored Cartelization and the Development of Federal Farm Policy," *Economic Inquiry*, 33 No. 3 (July 1995), pp. 365–382.

Application 13.6 The Difficulty of Controlling Cheating

The Organization of Petroleum Exporting Countries (OPEC) has relied on such measures as accounting firms to monitor member nations' outputs to eliminate cheating on production quotas.[13] These measures, however, have not been entirely successful. This is perhaps not surprising, given that member states possess differing production costs, petroleum reserves, time horizons, and goals. Enforcing agreements is also difficult when member states are at war with one another, as in the case of Iran and Iraq in the 1980s.

To demonstrate the difficulties facing OPEC—or any other cartel—in fixing prices, Raymond Battalio, a Texas A&M professor, conducted an experiment with a class of 27 introductory economics students. In the experiment, each student was asked to write either a *0* or a *1* on a slip of paper. A *1* indicated a willingness to adhere to a collusive agreement, whereas a *0* signified a desire to cheat. There were real money payoffs associated with the game and they were structured so that if everyone chose *1*, the total payoff to all the students would be maximized. The payoff to any individual student, however, was maximized if all other students opted for *1* and the individual cheated and selected *0*. If more than one student cheated, the payoff for cheating decreased as the number of cheaters increased, but was still greater than the payoff to noncheaters. Significantly, the payoffs associated with all possible outcomes were disseminated to students at the start of the experiment.

On the first round of the experiment, there were 6 noncheaters and 21 cheaters. Allowed to discuss the outcome, the students quickly realized that they could all be better off if they all voted to honor the agreement and place a *1* on their ballots. A student "leader" proposed, "Let's all put down *1*, and nobody cheat." After the students reached an explicit oral agreement to all vote *1* another vote was taken. There were 4 *1* votes and 23 *0* votes. The ringleader muttered, "I'll never trust anyone again as long as I live." Asked how he voted, the ringleader replied, "Oh, I voted *0*." The results suggest that in a cartel setting, self-interest leads to efforts to maximize returns individually even when the risk is great of lowering overall returns both to oneself and to one's "partners."

[13]This application is based on "OPEC May Monitor Members' Output On Site to Reduce Cheating on Quotas," *Wall Street Journal*, November 13, 1987, p. 44; and "All for One . . . One for All? Don't Bet on It," *Wall Street Journal*, December 4, 1986, p. 1.

Application 13.7 The Rolex "Cartel"

Although collusive agreements between different firms typically come to mind when cartels are mentioned, the model has wider applicability. Essentially, any firm with multiple production facilities or distribution channels and some market power faces a cartel management problem. The firm must coordinate the output and pricing decisions of its various plants and distribution channels to ensure that its total economic profit is maximized. To the extent, for example, that Rolex has some market power, it needs to ensure that one of its licensed dealers does not attempt to cheat on the Rolex "cartel" by selling more than a certain number of watches at the agreed-upon price. Although such cheating may benefit the individual dealer, it undermines the overall profitability of the cartel. To limit the undermining, Rolex may set up exclusive territories for its various dealers, allot a set number of watches to each dealer, and specify a manufacturer's suggested retail price.

In the mid-1980s, Rolex became disgruntled with Carl Marcus, a Beverly Hills retailer of its watches.[14]

[14]"The Rolex War Rages on for Beverly Hills Jeweler," *Los Angeles Times*, November 13, 1987, pp. 1 and 17.

Marcus had been buying Rolex watches from other dealers around the country who were unable to sell their allocated number. He then sold these watches in the Beverly Hills area at a discount of 5 to 25 percent below the price most jewelers were charging for the same items.

Rolex launched an advertising campaign against Marcus alleging that he had sold a used watch to a customer while claiming it was new. The advertisements asked Marcus's former customers to have their watches inspected by Rolex to verify that they had not been similarly duped. The underlying motive of the advertising campaign was not so much to protect unsuspecting customers as it was to preserve Rolex's market power and profit. By offering to inspect the watches Marcus sold, Rolex could obtain the registration numbers and identify the dealers who had undermined the cartel by selling some of their allotments to Marcus.

Oligopolies and Collusion

Much of what has been said about the consequences of cartel formation in a competitive industry also applies to an oligopolistic industry. Firms in an oligopolistic industry can increase their profits (they may already be making some pure economic profits to begin with) by colluding. The key to understanding this is our earlier finding that oligopoly output is usually greater than that of pure monopoly. This means that the oligopolistic firms' combined profit can be increased if total output is restricted to the monopoly level and, with an appropriate sharing arrangement, each firm can realize greater profit. The incentive to collude—the prospect of higher profit—also exists in an oligopoly.

We generally expect collusion to be more common in oligopolies than in competitive markets because there are only a few firms in an oligopolistic industry. The limited number of firms means that fewer parties must participate in the collusive agreement, making reaching agreement easier. Additionally, monitoring is simpler when few firms are involved; it is easier to detect cheaters.

Nonetheless, factors that inhibit the formation and maintenance of cartels in competitive markets are also present when oligopolistic firms collude. Each party to the cartel agreement still stands to make more profit by cheating than by abiding by the cartel agreement. Even though one firm does not face a horizontal demand curve if it cheats—as does the competitive firm in Figure 13.7a—it will face a more elastic demand curve when cheating than when complying, so its profit will be greater if it cheats and the other firms abide by the agreement. Furthermore, the higher price achieved by collusion can still prompt entry by new firms, and that also undermines the success of the cartel.

A final point is important. It is not necessary for all a market's firms to participate in the cartel for it to be worthwhile. Several firms producing most of an industry's output can increase their profit by colluding, although not by as much as when the entire industry is cartelized. In this case, the colluding firms will behave as if they were one giant firm and collectively exploit whatever monopoly power they have. The remaining firms then benefit from the higher price set by the colluding firms. In this case of partial market cartelization, the results can be analyzed using the dominant firm model illustrated earlier in Figure 13.6. In that graph, the marginal cost curve of the dominant firm is interpreted now as the summed marginal costs of the colluding firms, and they collectively maximize profit subject to the residual demand curve. Therefore, partial cartelization of an industry produces price and output lying between those found with pure monopoly and competition. We will see an example of this model's application in the next subsection.

Application 13.8 Firm Count, Market Concentration, and Successful Collusion

Even though price fixing is forbidden by the Sherman Act, numerous price-fixing agreements have been documented in the United States. The existence of such agreements corroborates Adam Smith's well-known observation: "People of the same trade seldom meet together, even for merriment or diversion, but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices."[15]

The ability to make a collusive arrangement stick appears to depend on the number of sellers and the degree of seller concentration in the relevant product or geographic market. Specifically, of the price-fixing conspiracies successfully prosecuted by the U.S. Department of Justice between 1963 and 1972, 10 or fewer firms were involved in 79 percent of the cases.[16] In 78 percent of the documented cases, the top 4 firms involved in the conspiracy accounted for more than 50 percent of the total market output. Thus, a small number of firms and a high degree of seller concentration appear to be more conducive to successful price collusion. In the relatively few cases in which a large number of firms have been involved, the conspiracies have been uncovered very rapidly. Indeed, details of some of the large-group organizational meetings sometimes have even been printed in local newspapers and thus have been quite easy for authorities to detect.

[15]Adam Smith, *The Wealth of Nations* (New York: The Modern Library, 1937), p. 128.

[16]George A. Hay and Daniel Kelly, "An Empirical Survey of Price-Fixing Conspiracies," *Journal of Law and Economics*, 17 No. 1 (April 1974), pp. 13–38.

The Case of OPEC

Few Americans took notice when OPEC was formed in 1960. Originally containing just 5 member nations, the cartel grew to an ominous 13 countries by 1973. During these years, however, OPEC could not be judged successful. World oil prices actually declined slightly in the 1960s. OPEC had not yet learned to use its potential market power.

All that changed with the 1973 Arab–Israeli War. During the War, OPEC's Arab members temporarily cut off oil exports. The result was dramatic: oil prices nearly quadrupled in a matter of months. The price of a barrel of oil on the world market was $3 in 1973 but jumped to $11 in 1974, providing a graphic demonstration of what an output restriction could accomplish. OPEC continued to hold down output after resuming exports; by accident, OPEC had learned how to run a cartel! After the War's end, oil prices remained relatively stable (in real terms) until they received another jolt in 1978, when revolution swept Iran. Iranian exports, which accounted for 20 percent of all OPEC exports, fell to almost zero. The world oil price rose to an average of $20 per barrel in 1979. Soon, thereafter, the Iran–Iraq War, which began in 1980, resulted in the widespread destruction of oil-producing facilities in both countries and reduced oil exports further. The world oil price jumped again to over $30 per barrel in 1980 and averaged $35 in 1981.

Before continuing the historical saga of developments in the world oil market, let's examine the market as an example of a cartel. Specifically, because OPEC does not control all supply sources (in 1972, OPEC production was 64 percent of noncommunist world output), we will use the dominant firm model. In view of the history of OPEC, this is a bit of an oversimplification—OPEC has not always been monolithic enough to act as a single profit-

maximizing firm. Nonetheless, we can gain insight into the way this market has functioned by treating OPEC as a dominant firm.

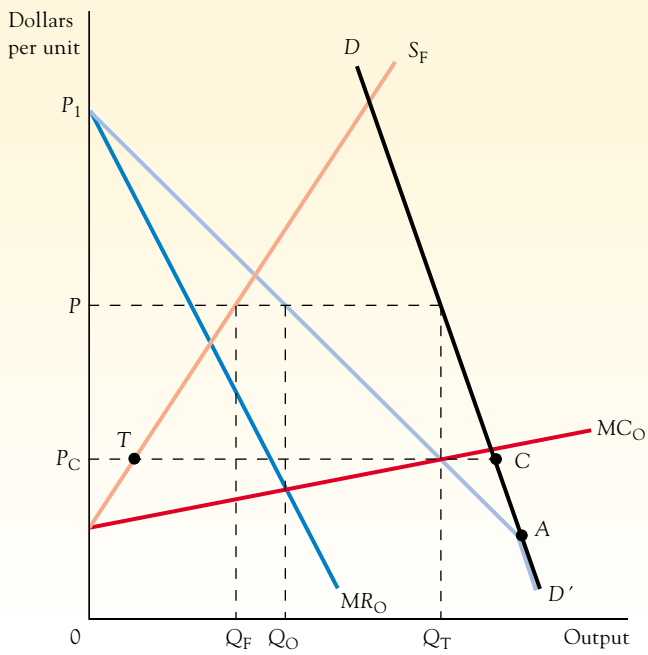
In Figure 13.8 we show world demand for oil as DD' and non-OPEC supply as S_F. These relationships should be interpreted as short-run, where the elasticities are relatively low. As we will see, low elasticities of demand and supply from fringe firms have a significant effect on the outcome. The marginal cost curve for OPEC is shown as MC_O, exhibiting substantially lower costs than other sources of supply, another characteristic of this market. OPEC's residual demand and marginal revenue curves are derived as before. Under competitive conditions (which prevailed before 1973), when OPEC and non-OPEC producers are producing where marginal cost equals price, price is P_C and total output is shown at point C on the demand curve. When OPEC behaves as a dominant firm, however, it produces where marginal cost equals its residual marginal revenue, an output of Q_O with a price of P. Note that price has risen sharply compared with the competitive price. This result should be contrasted with that shown earlier in Figure 13.6. In that representation of the dominant firm model, we assumed that demand and fringe supply were more elastic. A comparison of these graphs indicates why low elasticities of demand and fringe supply imply a higher cartel price. Equation (1) on page 371 also suggests this: decreases in the elasticity of the market demand and the fringe supply both work to lower the elasticity of demand faced by the dominant firm.

Because much of OPEC's early success was dependent on low elasticities, we need to be sure we understand why this was a characteristic of the world oil market.

The price elasticity of demand for oil is quite low in the short run. When oil prices rose, consumers were caught with a stock of energy-consuming capital designed for cheap oil. Houses, office buildings, and gas-guzzling automobiles could not be replaced overnight; it took time

FIGURE 13.8

OPEC Cartel as a Dominant Firm
World oil demand is shown by DD' and non-OPEC supply as S_F. The residual demand curve confronting OPEC is then P_1AD'. With its marginal cost curve MC_O, OPEC produces Q_O at price P. Non-OPEC output at that price is Q_F, so total output is Q_T.



for higher energy prices to have a substantial effect on energy consumption. A low price elasticity means, of course, that moderate output restrictions will produce a large price increase, just what we see in Figure 13.8.

The price elasticity of supply of oil from non-OPEC suppliers is also quite low in the short run. The biggest threat to a cartel is increased production by noncartel members or entry of new firms. OPEC's ability to raise oil prices successfully depended on total (OPEC and non-OPEC) output; if OPEC output restrictions were matched by substantial increases by others (as would happen with highly elastic non-OPEC supply), price would be largely unaffected. How much non-OPEC suppliers could increase output depended on their price elasticity of supply. Because non-OPEC producers were already producing at near-capacity levels, their immediate ability to increase output was limited. Furthermore, new oil fields could be brought into production only after a lengthy process of seismic exploration, drilling, and installation of pipelines requiring several years. Thus, non-OPEC producers were unable to respond quickly to the higher prices produced by the cartel.

These characteristics of the world oil market help explain OPEC's early success. In addition, OPEC also enjoyed another advantage: oil-importing nations frequently adopted policies that strengthened OPEC's position. In the United States, for example, price controls on oil kept the price received by domestic oil producers artificially low and discouraged production and exploration. Price controls on natural gas discouraged production of this alternative energy source. Similarly, tough environmental restrictions on the mining and use of coal slowed the transition to coal as another energy alternative. Finally, price controls on oil products such as gasoline and heating oil kept the prices of such products below world market levels and encouraged consumption. Encouraging domestic consumption and discouraging domestic production implies an increase in demand for oil from OPEC, and the United States inevitably became more dependent on imported oil during the 1970s.

Our story, however, does not end with the $35-per-barrel price of 1981. As economic theory tells us, long-run elasticities are greater than short-run elasticities. As time passed consumers responded to higher energy prices by switching to more energy-efficient homes, appliances, and cars, while non-OPEC suppliers increased output. Both responses put increased pressure on OPEC in the early 1980s. More specifically, world oil consumption fell from a 1979 high of 51 million barrels per day to less than 45 million in 1985 (particularly impressive since the prior long-run trend had been a 3 percent annual rate of increase in consumption). At the same time, non-OPEC production increased sharply, from 15 million barrels per day in 1977 to 24 million barrels in 1985. With total consumption down and non-OPEC production up, OPEC had to restrict its output sharply to try to hold the price up. Even though OPEC cut production from 31 million barrels per day in 1979 to barely 16 million barrels in 1985, the price of oil still fell below its previous level. The per-barrel price dropped to $32 in 1982 and to $27 in 1984, and it continued falling gradually until it stabilized at around $18 in 1987.

The experience of OPEC after 1981 was almost a textbook case of the operation of economic forces that undermine cartels. According to *Business Week*:[17] "The OPEC cartel is facing strains that it has never experienced before. . . . OPEC is so divided by cutthroat competition and internal political bickering that the organization is unlikely to find a way to agree anytime soon . . . Under-the-table discounts . . . are common."

[17]"The Leverage of Lower Oil Prices: On World Economies, on OPEC Countries, on the Oil Industry," *Business Week*, March 22, 1982, p. 69.

External pressures that diminished OPEC's power were also at work:[18] "Demand for OPEC oil . . . collapsed under the combined impact of . . . surprisingly high conservation, expanded use of alternative energy sources such as coal and gas, and rising non-OPEC production of oil from Mexico, the North Sea, and the North Slope of Alaska." (Long-run price elasticities of demand are higher than short-run price elasticities and were being felt as consumers switched to substitutes. Entry at the high cartel price was taking place.)

To illustrate in analytic terms how the passage of time affected the market, long-run demand and non-OPEC supply responses can be incorporated into Figure 13.8. To avoid cluttering the diagram, we have not done so, but we recommend that you work this out. Specifically, assume that the original competitive equilibrium was a long-run equilibrium. Then a more elastic long-run demand curve will pass through point C on the short-run demand curve, and a more elastic supply curve will pass through point *T* on the non-OPEC supply curve. With these curves, where will the long-run cartel equilibrium lie? You will find that in the long run, price will be lower, non-OPEC output will be greater, and OPEC output will be smaller (all compared to the short-run cartel equilibrium shown in our graph)—just the events we saw emerge in the 1980s.

Most observers concluded that OPEC's power had largely eroded by the late 1980s. In August 1990, however, Iraq invaded Kuwait and supply was disrupted again. The price of oil quickly rose to $40 per barrel, before falling back to about $20 per barrel after the United States defeated Iraq in the Gulf War in early 1991. These events do not mean that OPEC was exercising control, of course. Small changes in supply can have pronounced effects on price in a competitive market in the short run when short-run demand is very inelastic. As long as political turmoil pervades the Mideast, a major source of world oil, volatility is likely to characterize this market.

By 1998, the price of oil had fallen below $12 a barrel, the lowest level in a decade, because of increased global production and a decrease in OPEC's share of total output to 40 percent (from more than 50 percent in the early 1970s). The decline in the price of oil was arrested in 1999, due to a pact negotiated between the three largest exporters of oil to the United States: Saudi Arabia, Venezuela, and Mexico. The first two countries belong to OPEC, while Mexico does not. The three countries respectively account for 15, 18, and 14 percent of U.S. imports. The agreement, known as the *Riyadh Pact*, consisted of pledges to reduce output by 1.5 to 2.0 million barrels a day, about 2 to 3 percent of world production of 73 million barrels a day. The price of crude oil rose above $30 in the pact's wake.

Mexican officials who were key to brokering the pact, however, didn't hold any illusions about its long-run prospects. Then-president of Mexico, Ernesto Zedillo, recalled that when he was an economics student at Yale University, one of his professors engaged his class in a game to test the limits of oligopolistic behavior (see Application 13.6).[19] "Within a few minutes, somebody would start cheating; it never failed," Zedillo observed. "No market with more than two participants can sustain a cartel."

[18]Ibid., p. 69.

[19]"'Big 3' Exporters' Pact to Cut Oil Output Signals Seismic Shifts," *Wall Street Journal,* June 23, 1998, pp. A1 and A10.

Summary

- A monopolistically competitive market is one in which there are a large number of competing firms, but each firm produces a differentiated product.
- Each firm in a monopolistically competitive market confronts a demand curve that is fairly, but not perfectly, elastic.
- In long-run equilibrium, firms in a monopolistically competitive market make zero economic profits. They are not, however, producing at the minimum point on their long-run average cost curves since the LAC curve is tangent to a downward-sloping demand curve.
- Oligopoly is characterized by a few firms that together produce all or most of the total output of some product. A pronounced mutual interdependence among the decisions of firms in the industry results.
- What one firm does has a decided impact on other firms in an oligopoly, but the way that other firms will react is uncertain.
- There are three common oligopoly models: Cournot, Stackelberg, and dominant firm.
- Oligopoly firms can collude and operate as a cartel.
- The variety of models for studying oligopoly and the differences in their implications make generalizing about the effects of oligopoly difficult. In some cases, prices are predicted to be near the monopoly level. In other cases, they hover near the competitive level.
- Oligopolistic outcomes lie somewhere between the monopoly and competitive results and differ from one industry to another.

Review Questions and Problems

Questions and problems marked with an asterisk have solutions given in Answers to Selected Problems at the back of the book (page 577).

13.1. What are the assumptions of the theory of monopolistic competition? In what way do these assumptions differ from those of the perfectly competitive model?

13.2. Explain the relationship between the demand elasticity and the excess capacity that occurs for a monopolistic competitor.

13.3. Explain how, in the Cournot model, the output of one firm depends on the output of other firms. Specifically, in Figure 13.3, what will be the output of Artesia if Utopia produces 32 units? If Utopia produces 48? If Utopia produces 64?

13.4. Starting from the Figure 13.3 Cournot equilibrium, suppose that the marginal and average total cost curves (which are the same for both firms) shift downward. Explain how the firms adjust to a new Cournot equilibrium.

13.5. How does the Stackelberg model differ from the Cournot model? Which model predicts that output will be higher?

13.6. Explain how the residual demand curve confronting the dominant firm in the dominant firm model is derived. In this derivation, what is assumed regarding how the output of other firms is determined? How does it differ from the Cournot assumption?

13.7. Explain how equilibrium is determined in the dominant firm model. If market demand increases, how will a new equilibrium be determined?

***13.8.** Suppose that the supply curve of the "follower" firms in the dominant firm model is perfectly horizontal. Does the dominant firm still have the power to set the industry price?

13.9. "Because firms in any industry can always make greater profits by colluding there is an inevitable tendency for competitive industries to become cartels over time." Is the first part of this statement correct? Is the second part? Explain.

13.10. What problems usually make cartels collapse? How was OPEC able to avoid this fate, at least through the mid-1980s?

13.11. Consider the dominant firm model and treat OPEC as the dominant firm. Explain how OPEC would determine the price of oil and the level of output produced by the cartel. How would OPEC's price and output be affected by new discoveries of oil that shift the supply curve of oil for non-OPEC members to the right?

13.12. "There is no general theory of oligopoly." Explain this statement.

13.13. Under which oligopoly model does the outcome most nearly resemble that obtained with pure monopoly? Under which oligopoly model does the outcome most nearly resemble that obtained with perfect competition?

13.14. In an analysis of the automobile industry, what factors would you consider in determining whether to use the competitive model, the monopoly model, or one of the oligopoly models?

13.15. Suppose that there were three identical firms instead of only two under the cost and market demand conditions outlined in Section 13.2. What would be the Cournot equilibrium in terms of each firm's output as well as the total market output? If there were four identical firms sharing the market?

13.16. In a two-firm Stackelberg model of oligopoly, can both firms be "leaders"? Explain why or why not.

13.17. Suppose that Iran and Iraq are Cournot duopolists in the crude oil market and face the following market demand function:

$$P = 100 - (q_{1+}q_2);$$

where q_i represents the output levels of the two countries with Iran being "1" and Iraq being "2" and P is the per-barrel price. The marginal revenue schedules facing the two countries are:

$$MR_1 = 100 - 2q_1 - q_2$$

$$MR_2 = 100 - 2q_2 - q_1.$$

Each country has a marginal cost curve of the form

$$MC_i = q_i;$$

where $i = 1, 2$.

a. Determine each country's reaction function.

b. Does a Cournot equilibrium exist? If so, find the outputs and prices of crude oil in the two countries.

c. Suppose that the two countries collude and become a cartel. What will be the resulting price and outputs for crude oil for the two countries? [Note that the market marginal revenue is $100 - 2(q_1 + q_2)$.]

d. Can it be said that because collusive profits are strictly greater, it is true that these countries should necessarily collude? Are there any potential pitfalls in such a collusive arrangement?

13.18. Suppose that Iraq is the Stackelberg leader in the preceding problem. What will be each country's reaction function? How much will each country produce and what will its profits be?

13.19. Repeat the preceding problem but assume that Iran is the Stackelberg leader.

13.20. It is the not too distant future and all important human needs have been eliminated. Thanks to the firm Bioeconotek, a miracle drug, Needless, has been invented that genetically suppresses a patient's desire to think in terms of having needs. The drug stimulates the patient's cognitive ability to perceive substitute products as being ever more attractive the greater the price of any item.

Needless does not come cheap. Given its curative powers, Bioeconotek charges $1,633 per dose and sells 62,020,000 doses annually despite the presence of a competitive fringe of generic suppliers, which collectively sell 51,650,000 doses annually at the same $1,633 price.

Bioeconotek managers estimate that the fringe supply curve is described by the equation:

$$P = 600 + \frac{Q_{FS}}{50{,}000};$$

and that the market demand is given by:

$$P = 13{,}000 - \frac{Q_M}{10{,}000}.$$

In the preceding equations P is the price per dose, Q_{FS} is the quantity supplied by the competitive fringe, and Q_M is quantity consumed in the total market.

The marginal cost associated with producing and distributing Needless is a constant $600 per dose. Bioeconotek initially invested $400 million in developing the drug but has faced no additional fixed costs since the initial investment.

a. At the current price of $1,633, a Bioeconotek manager estimates the demand elasticity for Needless as:

$$\text{Elasticity} = \left|\frac{\Delta Q}{\Delta P}\frac{P}{Q}\right| = \left|(-10{,}000)\left(\frac{1{,}633}{113{,}670{,}000}\right)\right| = 0.14.$$

Since this value represents an inelastic demand, he argues that Needless's price should be raised. Is there something wrong with this reasoning? Briefly explain.

b. Write an equation that gives the demand for Bioeconotek's own output—that is, demand as seen by Bioeconotek given market demand and the fringe supply.

c. What is the profit-maximizing price for Bioeconotek to charge, given the current market demand and the current fringe supply? (Recall that if a demand curve can be written as $P = a - bQ$, then $MR = a - 2bQ$.)

d. If Bioeconotek wants to implement a "limit-pricing" strategy that successfully eliminates fringe suppliers from the market, what annual output level does it need to select for its own product, Needless?

e. Another Bioeconotek manager worries about future growth of the fringe supply and argues that the company should set the price determined in part d so as to drive the current fringe out of business and eliminate any incentives for further entry. As briefly as possible, explain why this strategy cannot be more profitable than the pricing strategy you gave in part c.

13.21. Suppose that the market for Web search engines can best be characterized as monopolistically competitive. If this is the case, should firms that operate in the market such as Yahoo and Alta Vista be prosecuted by antitrust authorities on account

of the potential inefficiency of the market outcome? Explain why or why not.

13.22. College textbooks have been increasing in price at a rate greater than the general price level over the past several decades. The reason for this phenomenon appears to be that publishers have realized that by having authors revise their texts more frequently, competition from used texts is reduced. Using the dominant firm model, explain how such a more-frequent-revision strategy leads to a higher price for a given textbook.

13.23. Historically, officials from 23 elite northeastern colleges with selective admissions policies and high tuition met each spring to compare financial aid packages for more than 10,000 common applicants. The meetings, known as "overlap," were designed to eliminate any differences in the financial aid packages offered by the various colleges: What would you predict would be the effect on the net price (tuition less financial aid) paid by applicants to the colleges participating in the overlap meetings? Are there any factors that would work to undermine the ongoing viability of the overlap practice? If so, what might these factors be?

13.24. In 2002, drug manufacturer Schering-Plough's monopoly (through patent protection) on the huge-selling allergy medication Claritin was set to expire. To protect its $3 billion in annual sales from Claritin, Schering-Plough sued 10 prospective manufacturers of generic substitutes for Claritin. The suits alleged that anyone who swallowed a generic version would produce in their livers a separate chemical compound or metabolite on which Schering-Plough had patent protection through April 2004. Using the dominant firm model, explain what Schering-Plough's actions were intended to accomplish in the market for allergy medications.

13.25. Assume that Intel can be treated as the dominant firm in the market for computer chips. What three basic factors determine the elasticity of demand confronted by Intel? Explain whether increases or decreases in these three factors will increase Intel's elasticity of demand, and provide an intuitive explanation why.