CHAPTER 11

Market Power, Collusion, and Oligopoly

Market power is an elusive goal. It is limited everywhere by the threat of entry. Even a firm producing a unique product with no close substitutes might not be able to engage in monopoly pricing, because the profits that it would earn by doing so would lure entrants and destroy its market position.

But market power can be highly profitable to those who achieve it, and is therefore avidly pursued. In this chapter, we will look first at some of the strategies that firms employ in their quest for a monopoly position. These can include mergers, predatory pricing, and *fair trade* agreements. We will examine each strategy and each strategy's limits. We will also see that activities that appear to be attempts either to gain or to exploit monopoly power are not always what they seem.

Collusion among existing firms is one of the most straightforward and common methods of trying to monopolize a market. It is important enough that we devote an entire section to it, Section 11.2. Using tools from the theory of games, we will see why collusion is often doomed to fail.

We will then see that a collusive arrangement among firms that would ordinarily collapse under its own weight can at times be supported by various forms of regulation. This discussion occupies Section 11.3. Although regulation sometimes plays this role, it also plays a variety of others, and there are a great number of theories of the regulatory process. We will survey a few ideas from this large body of thought.

Finally, we will turn from the pursuit of market power to its exercise. We already have (from Chapter 10) a simple model of monopoly behavior, which ignores the firm's need to respond to other firms' actions. In Section 11.4, we will survey some theories of oligopoly that provide a starting point for thinking about industries with small numbers of firms, each enjoying some monopoly power but each affected by the others' behavior. Under this heading, we will consider some classical models of oligopoly and the contemporary theory of contestable markets. In Section 11.5, we will look at the related theory of monopolistic competition, which also tries to model firms that exercise some degree of monopoly power while simultaneously competing with other firms.

11.1 Acquiring Market Power

In this section, we will explore some methods that firms either use or are alleged to use in their attempts to acquire and exploit market power. We will explore the limits of these methods, and we will learn that they are not always what they seem.

Mergers

Horizontal integration
A merger of firms that produce the same product.

Vertical integration
A merger between a firm that produces an input and a firm that uses that input.

The issue of monopoly power arises whenever two firms merge to form a larger firm. Mergers can be roughly classified into two types. **Horizontal integration** combines two or more producers of the same product. An example would be the combination of three computer manufacturers like Dell, Gateway, and IBM into a single company. **Vertical integration** combines firms one of which produces inputs for the other's production processes. An example would be the merger of a computer manufacturer (like Dell) with a chip manufacturer (like Intel).

Horizontal Integration

There are essentially two different reasons why firms might want to merge horizontally. First, there may be economies of scale or other increased efficiencies associated with size so that a larger firm can produce output at a lower average cost. Second, there may be an opportunity for the larger firm to exercise some monopoly power. Of course, both motives may be present in a single merger.

From a welfare point of view, mergers are desirable insofar as they reduce costs, and they are undesirable insofar as they create monopoly power. Exhibit 11.1 illustrates the trade-off. We assume that the industry is initially competitive, with marginal cost curve MC. (The marginal cost curve is drawn horizontally in order to simplify the diagram; nothing of importance depends on this simplification.) If the firms in the industry merge, technical efficiencies will lower the marginal cost curve to MC', but they will also enable the new, larger firm to exercise monopoly power, producing the monopoly quantity Q', where MC' crosses the marginal revenue curve MR.

The welfare consequences of the merger are ambiguous. There is a gain of $F + G$, representing the cost savings due to greater efficiency (the rectangle $F + G$ has area equal to Q' times the cost savings per unit). There is also a loss of E, due to the reduction in output. Which of these is greater will vary from one individual case to another.

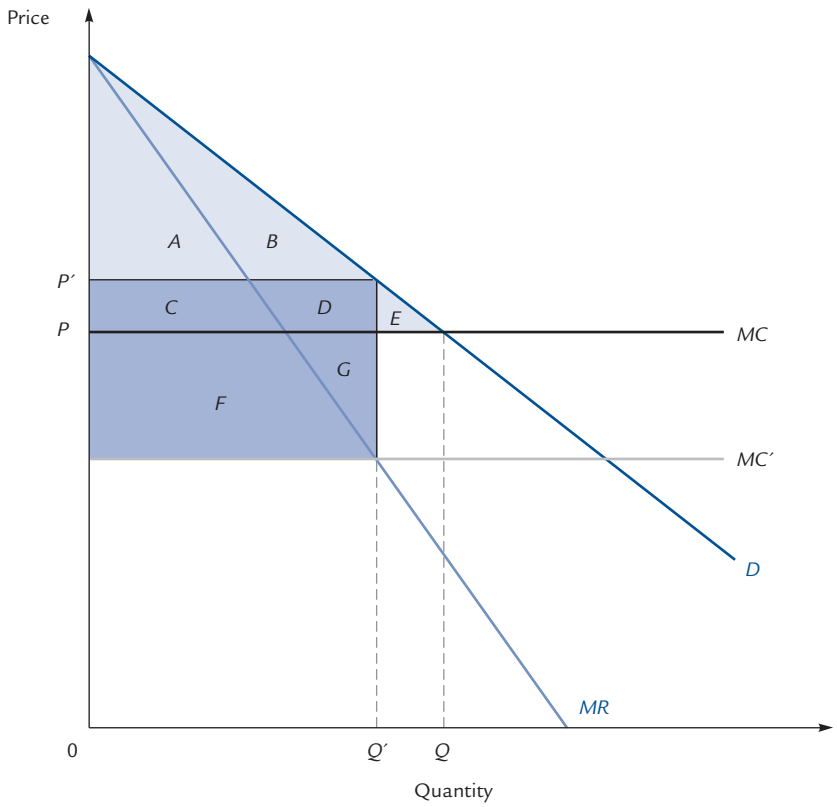

Dangerous Curve

The analysis here is incomplete if it is possible for another firm to enter the market. Even if the new entrant has the relatively high marginal cost curve MC, it can undercut the price P'. Sufficiently many such new entrants—or even just the threat of new entrants—will drive the market price back down to P.

If MC' is very much lower than MC, then the picture looks like Exhibit 11.2. In this case, the monopoly price P' is actually lower than the competitive price P, and both consumers and producers gain from the merger.

Exercise 11.1 Suppose that the merger does not reduce costs at all, so that $MC = MC'$. Draw the appropriate graph. In this case does the merger have an unambiguous effect on social welfare?

EXHIBIT **11.1** **A Horizontal Merger**



	Before Merger	After Merger
Consumers' Surplus	$A + B + C + D + E$	$A + B$
Producers' Surplus	—	$C + D + F + G$
Social Gain	$A + B + C + D + E$	$A + B + C + D + F + G$

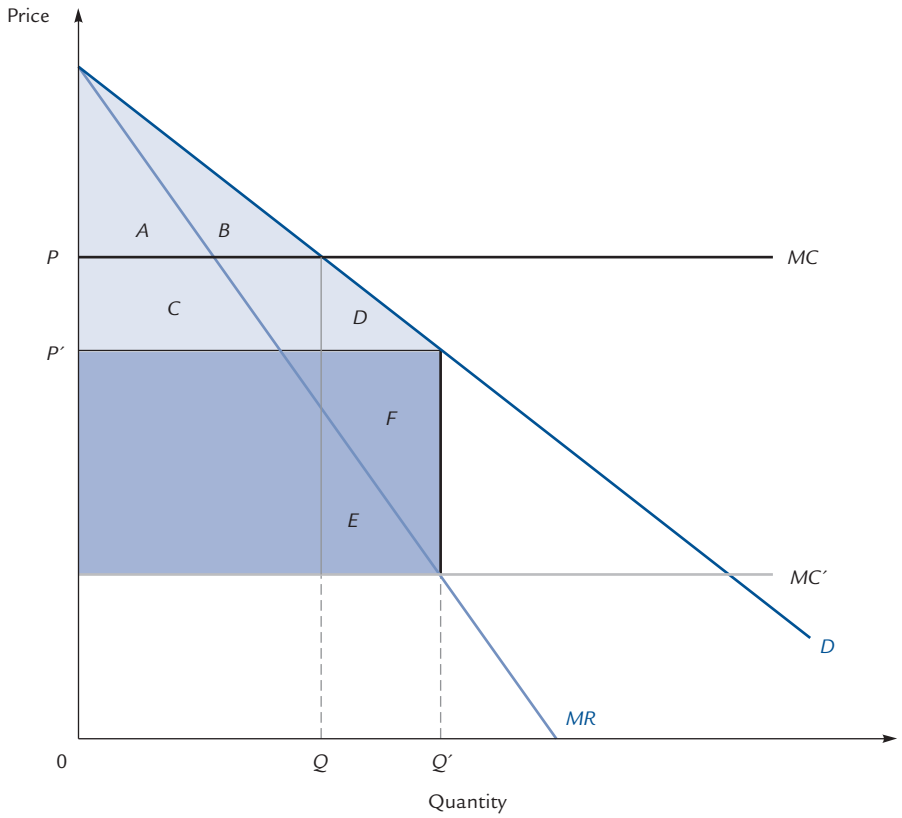
Initially, the industry's marginal cost (= supply) curve is *MC*. If the industry is competitive, it produces the equilibrium output *Q* at the price *P*. Because the *MC* curve is horizontal, there is no producers' surplus.

Following a merger, marginal cost is reduced to *MC′*, but the newly created firm has monopoly power and so produces the quantity *Q′*, where *MC′* crosses the marginal revenue curve *MR*. The monopoly price is *P′*. The table above computes welfare before and after the merger.

The Great American Merger Wave

In the years 1895–1904, a great wave of mergers swept through America's manufacturing industries. Many of the country's largest corporations—U.S. Steel, American Tobacco, Dupont, Eastman Kodak, General Electric, and dozens more—were formed at this time. The resulting megacorporations often controlled 70, 80, or even 90% of their markets, leading to the widespread assumption that the purpose of the mergers was to create monopoly power.

EXHIBIT **11.2** **A Horizontal Merger Leading to a Large Cost Reduction**



	Before Merger	After Merger
Consumers' Surplus	$A + B$	$A + B + C + D$
Producers' Surplus	—	$E + F$
Social Gain	$A + B$	$A + B + C + D + E + F$

If the competitive industry's marginal cost curve is *MC,* and if a merger converts the industry into a monopoly with the much lower marginal cost curve *MC′*, then price will fall from *P* to *P′*, benefiting both consumers and producers.

But Professors Ajeyo Banerjee and Woodrow Eckard object to this assumption.[1] Here's why: Mergers that create monopoly power—and therefore raise prices—are good for every firm in the industry, whether or not they're part of the merger. If American Tobacco, with its 90% market share, was able to significantly raise prices, then small tobacco firms should have rejoiced, and their share prices should have risen. But that didn't happen. In general, firms that were left out of the mergers saw their share prices fall.

Banerjee and Eckard point out that this would all make sense if the mergers were designed not so much to create monopoly power as to lower production costs. In that

[1] A. Banerjee and E.W. Eckard, "Are Mega-Mergers Anti-Competitive? Evidence from the First Great Merger Wave," *Rand Journal of Economics* 29(4), Winter 1998, 803–827.

case, the firms that were left out would have found it difficult to compete with the more efficient megacorporations, which would explain why their stock prices fell.

Antitrust Policies

The Sherman Act of 1890 and the Clayton Act of 1914 give the courts jurisdiction to prevent mergers that tend to reduce competition. There has been much controversy about exactly what criteria the courts should apply in determining whether a particular merger is illegal.

One viewpoint is that mergers should be prohibited only when they reduce economic efficiency. According to this viewpoint, the court should compare areas in Exhibit 11.1 before deciding whether or not to allow a particular merger. If a merger reduces costs by enough to make the graph look like Exhibit 11.2, then according to this viewpoint the merger should certainly be allowed.

In a series of decisions beginning with *Brown Shoe v. the United States* (1962), the Supreme Court under Chief Justice Earl Warren explicitly rejected this viewpoint. Instead, the Court placed particular emphasis on the welfare of small firms that are not involved in the merger. The Court held that the Sherman and Clayton acts should be interpreted so as to protect such firms by disallowing mergers that would make it difficult for them to compete. In these cases, the Court took the position that a merger could be illegal precisely *because* it would lead to a reduction in costs, lower prices, and increased economic efficiency. The reason is that smaller, less efficient firms would not be able to survive in the new environment, and the Court considered the interests of those firms to be protected by the law.

More recently, U.S. courts have largely retreated from this position and placed considerable emphasis on economic efficiency as a criterion for allowing mergers. Most European courts, however, continue to disallow mergers that create or strengthen dominant market positions, even when they are economically efficient. In the European Court of Justice, "Efficiencies are often seen as evidence of market power, rather than as benefits which may outweigh the anti-competitive consequences of mergers."[2]

Vertical Integration

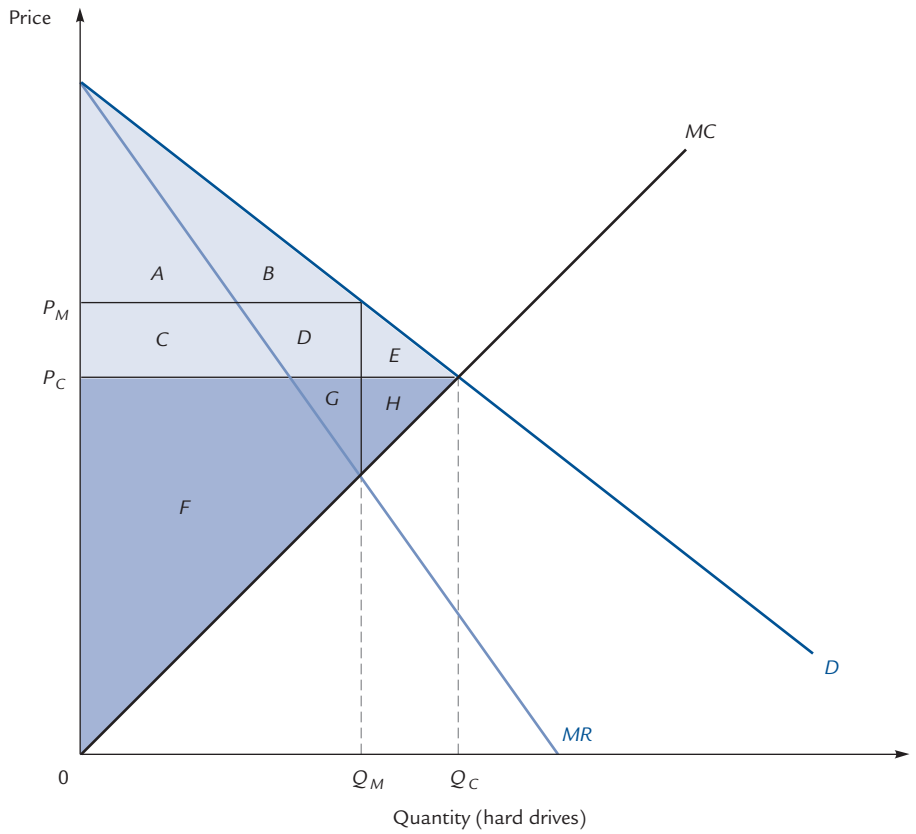
If there were only one computer manufacturer (say, Dell), you'd pay a monopoly price for your computer. If there were only one computer manufacturer *and* only one hard drive manufacturer (say, Seagate), you'd pay even more. That's because Seagate would charge Dell a monopoly price for hard drives, and Seagate's monopoly price would become part of Dell's marginal cost. When a monopolist's marginal cost curve rises, so does the price of his product.

Now suppose the two monopolies combine into a single company; say, for example, that the monopolist Dell acquires the monopolist Seagate. Suddenly, Dell isn't paying a monopoly price for hard drives anymore. That lowers Dell's marginal cost, which leads to a lower price for Dell's computers.

The moral of this fable is that vertical integration can eliminate monopoly power and benefit consumers. Exhibit 11.3 shows the argument in more detail. The graph represents the market for hard drives. Initially, Seagate charges Dell the price P_M, earning a producer's surplus of $C + D + F + G$ and leaving a consumer's surplus of

[2] P. Cayseele and R. Van den Bergh, "The Economics of Antitrust Laws," in: Bouckaert, B., and G. DeGeest (eds.), *Encyclopedia of Law and Economics,* Kluwer (2000).

EXHIBIT **11.3** **Vertical Integration**



A monopoly hard drive manufacturer (Seagate) produces Q_M hard drives for sale to a monopoly computer manufacturer (Dell). This maximizes producer's surplus at $C + D + F + G$ while restricting consumers' surplus to $A + B$.

If Dell acquires ownership of Seagate, it will earn both the producer's and the consumers' surpluses and will therefore want to maximize the sum of the two. This is accomplished by producing the quantity Q_C of hard drives, creating a gain equal to the sum of all the lettered areas. Social gain is increased by $E + H$. More hard drives are produced, more computers are produced, and the price of computers goes down.

$A + B$ for Dell. (Note that although Dell is the *producer* in the market for computers, it is the *consumer* in the market for hard drives.)

But when Dell acquires Seagate, it is essentially in the position of selling hard drives to itself, which means that Dell collects both the producer's and consumer's surpluses. To maximize the sum of the surpluses, Dell increases production from the quantity Q_M to Q_C, where the total surplus is $A + B + C + D + E + F + G + H$. More hard drives means more computers, and more computers means lower computer prices.

That shows that a vertical merger is attractive to consumers. Is it also attractive to Dell and Seagate? The answer is yes. Dell's total surplus after the merger is greater than the sum of the two companies' surpluses before the merger. Therefore, both companies' owners can come out ahead, provided Dell buys Seagate for an appropriate price.

Exercise 11.2 In terms of the areas in Exhibit 11.3, what is an appropriate price for Dell's purchase of Seagate?

This example shows that when a monopolist integrates vertically with a monopolist, the net effect is to benefit everyone, including consumers. But there are other types of vertical integration. You could, for example, imagine a merger that combines a competitive computer manufacturer with a monopoly disk drive manufacturer, or a competitive disk drive manufacturer with a competitive computer manufacturer. Each case needs a separate analysis, and some cases are very complicated. In those cases, vertical integration can be either good or bad for consumers, depending on the specifics of market structure and the shapes of the demand and cost curves.

Predatory Pricing

Predatory pricing occurs when a firm sets prices so low as to incur losses, forcing its rivals to do the same. If the firm can outlast the competition in the resulting "price war," it may hope to be the only survivor. Conceivably, a firm could engage in predatory pricing in some markets while continuing to charge normally in others. In this case, predatory pricing becomes a form of price discrimination.

Predatory pricing
Setting an artificially low price so as to damage rival firms.

Economists disagree about how widespread this practice really is. There are a number of reasons for skepticism. First, there is nothing to prevent the reemergence of rival firms as soon as the would-be monopolist raises its prices. Second, during the period of price warfare, all sides are losing money. The predator's losses, however, are greater: It is the predator who is attempting to expand market share and therefore selling greater quantities at the artificially low price. Indeed, if the other firms "lay low" by producing very little (or even nothing) for a while, they can force the predator to take losses that are enormous compared with their own. Finally, a firm being preyed upon, if it is capable of competing successfully in the long run, can usually borrow funds to get through the temporary period of price cutting. Thus, even a predator whose assets greatly outstrip its rivals' may not have any survival advantage over them.

The United States Supreme Court expressed its own skepticism of predatory pricing as a viable economic strategy when Zenith and other U.S. firms accused Matsushita and other Japanese firms of using predatory pricing to monopolize U.S. markets for consumer electronics. The court found it implausible that predatory pricing would be a profitable strategy, and concluded that the Japanese firms offered low prices because they were competing for business rather than implementing an "economically senseless conspiracy."

Despite all of these arguments, there are still reasons to think that predation might sometimes be profitable. The most significant of these is that predation can serve as a warning to future entrants. By driving one rival from the marketplace, the predator can prevent many additional rivals from entering in the first place. This can make predation a sensible strategy, even when the predator's losses from underpricing far exceed its gains from the first rival's elimination.

Even so, firms can sometimes protect themselves against predation. One recent case involved a company called Empire Gas, which sold liquid petroleum and competed against several smaller, more localized companies. By cutting prices below wholesale in just a few markets at a time, Empire tried to send a message about its willingness to punish competitors. But several competitors responded by offering

their customers long-term contracts at competitive prices. Even though Empire's prices were lower, many customers realized that the low prices were unlikely to last very long, and preferred to pay a bit more in exchange for the long-term assurance of a reasonable price. Eventually, the Court of Appeals ruled that Empire Gas surely did engage in predatory pricing, but no remedy was necessary because no harm had been done.

Example: The Case Against Wal-Mart

In 1991, three pharmacies in Arkansas sued Wal-Mart for predatory pricing of prescription drugs. The three pharmacies maintained that Wal-Mart had deliberately set low prices to drive them out of business and establish a monopoly; Wal-Mart responded that it offered lower prices because it was more efficient than the other pharmacies. In essence, the plaintiffs were arguing that Wal-Mart priced below marginal cost, whereas Wal-Mart argued that both its prices and its marginal costs were low. A trial court agreed with the plaintiffs, but the Arkansas Supreme Court (in a 4–3 decision) overturned the trial court and ruled in Wal-Mart's favor.

Wal-Mart was helped at trial when one of the plaintiffs admitted that competition from Wal-Mart had provoked him to greater efficiency, which suggests that before Wal-Mart's arrival, prices had in fact been higher than necessary.

Example: The Standard Oil Company

Historians have traditionally attributed much of the success of the Standard Oil Company to predatory price cutting. Founded in 1870 by John D. Rockefeller, Standard Oil was estimated to supply 75% of the oil sold in the United States by the 1890s. In 1911 Standard Oil (by now reorganized and called Standard Oil of New Jersey) was dissolved by order of the U.S. Supreme Court.

The role of predatory pricing in the Standard Oil case was reexamined by John McGee of the University of Washington in 1958.[3] In a widely quoted article, he argued that no historical evidence supports the assertion that predatory pricing played a major role in Rockefeller's success. Instead, McGee argued, this success could be attributed primarily to a successful policy of buying out rivals. The one-time cost of such buyouts was substantially less than the cost of predation.

Buyouts also have the advantage of allowing the would-be monopolist to acquire the rival firm's physical plant and equipment, which at least delays the rival's ability to reconstitute itself. A firm that stops producing in response to predatory price cutting still has its factories, ready to go back into production the instant prices are raised.

On the other hand, buyouts have the disadvantage of actually encouraging new entrants, who may be hoping to be bought out at a favorable price. And a firm that has been "bought" may soon reappear under a new name. It is said that more than a few nineteenth-century businessmen made lifetime careers out of being bought out by John D. Rockefeller.

[3] John McGee, "Predatory Price Cutting: The Standard Oil (N.J.) Case," *Journal of Law and Economics* 1 (1958): 137–169.

The Robinson–Patman Act

Because of the potentially predatory nature of price discrimination, the Robinson–Patman Act of 1938 forbids price discrimination in cases where it tends to "create a monopoly, lessen competition, or injure competitors." This language is sufficiently imprecise as to invite controversy over exactly when price discrimination should be considered predatory. The most widely accepted standard (but by no means the only one) was offered in 1975 by Phillip Areeda and Donald Turner of the Harvard Law School.[4] They argue, among other things, that no price can be considered predatory unless it is below marginal cost. As long as the firm is pricing at or above marginal cost, those rivals who are more efficient (i.e., have even lower costs) should be able to survive. Only when the firm prices below marginal cost is there a risk of its driving out a more efficient rival.

The Supreme Court gave its interpretation of the Robinson–Patman Act in the 1967 case *Utah Pie v. Continental Baking Company.* Utah Pie was a small, local company with 18 employees marketing frozen pies in the Salt Lake City area. Continental Baking, Carnation, and Pet were large national producers of a wide variety of food products. Utah Pie alleged that these three giants price-discriminated in an injurious way by selling frozen pies at a lower price in Salt Lake City than they did elsewhere. The Supreme Court agreed.

All parties to the *Utah Pie* case were in agreement that the defendants charged lower prices in Utah Pie's marketing territory than they did outside it. However, this could have resulted from the fact that elasticity of demand for Continental pies was greater in areas where Utah Pie's products were sold. In other words, Continental's actions could have been a simple case of ordinary third-degree price discrimination.

According to the Areeda–Turner rule, the price discrimination could have been considered predatory only if the defendants had priced below marginal cost in the Salt Lake City area. No evidence was offered that they had done so. Thus, the Supreme Court's decision makes deviation from marginal cost an irrelevant criterion in deciding whether a pricing policy can be considered predatory. For this reason economists generally regard *Utah Pie* as a bad decision. By forbidding Continental et al. to undercut Utah Pie's prices, the Court is as likely to have created a local monopoly (in the hands of Utah Pie) as to have prevented one.

In fact, the Supreme Court essentially took the position that the mere fact that the price of pies decreased in Salt Lake City constituted a violation of the Robinson–Patman Act![5] This reinforced the Court's interpretation of the Sherman and Clayton acts, by reaffirming that benefits to consumers are not considered a defense against the charge of injury to other firms.

Resale Price Maintenance

I (the author of your textbook) recently decided to buy a digital camcorder. So I drove to Best Buy, a major electronic retailing chain, where an extremely knowledgeable and helpful salesperson educated me about the available features and the pros and cons of each brand. After taking a half hour of his time, I knew which camera I wanted—a Panasonic. Best Buy's price was $900. I went home, found the identical camera on the World Wide Web for $600, and bought it online.

[4] P. Areeda and D. Turner, "Predatory Pricing and Related Practices Under Section 2 of the Sherman Act," *Harvard Law Review* 88 (1975): 689–733.
[5] For more on this point, see Bork, *The Antitrust Paradox,* pp. 386–387.

Obviously, this practice is a disaster for Best Buy. A little less obviously, it can be a disaster for Panasonic as well. If there are enough customers like me, Best Buy will stop offering its excellent service—which means that customers like me will be less likely to learn about the advantages of a Panasonic camera.

By supplying cameras to online discounters, Panasonic attracts additional customers (namely those who won't pay Best Buy prices) while risking the loss of Best Buy's promotional services. Apparently, they've decided that the benefits of dealing with discounters outweigh the costs. But not every firm in similar circumstances has reached the same conclusion. The Schwinn bicycle company used to require all sellers of Schwinn bicycles to charge a full retail price. If a seller was caught discounting, Schwinn would cut off that seller's supply. This practice—when a monopoly seller prohibits retailers from offering discounts—is called **resale price maintenance** or **fair trade.**

Resale price maintenance or fair trade

A practice by which the producer of a product sets a retail price and forbids any retailer to sell at a discount.

Resale price maintenance is sometimes misinterpreted as an attempt by the manufacturer to keep prices high. But the price consumers will pay for Schwinn bicycles is determined by the quantity of bicycles Schwinn chooses to produce. If Schwinn had a monopoly and wanted to raise prices, all it would have to do is restrict output. And conversely, unless Schwinn restricts output, no fair trade arrangement could have enabled it to sell its bicycles at a price higher than demanders were willing to pay.

It is most plausible, then, that Schwinn engaged in retail price maintenance in order to ensure that retailers would continue to offer a high level of service—displaying bicycles in showrooms and educating customers about their features. As with cameras, if some retailers offered cut-rate prices, customers would first go to the stores with the fancy showrooms and knowledgeable salesforces, ask their questions, make their decisions, and then buy from the discounters. Eventually, those retailers who offered quality service would find that there are no rewards in that activity, and so they would eliminate all of the costly forms of assistance that customers find valuable. Consumers could find themselves worse off, and so could Schwinn, as buyers would now have a greatly reduced incentive to purchase Schwinn bicycles.

Through resale price maintenance, Schwinn ensures that its dealers, who cannot compete with each other by offering lower prices, will instead compete with each other by attempting to offer higher-quality service. Thus, according to this theory, a practice that at first seems designed to establish monopoly power at the expense of consumers can actually be more plausibly explained as a practice designed to make the product more desirable by providing consumers with services that they value.

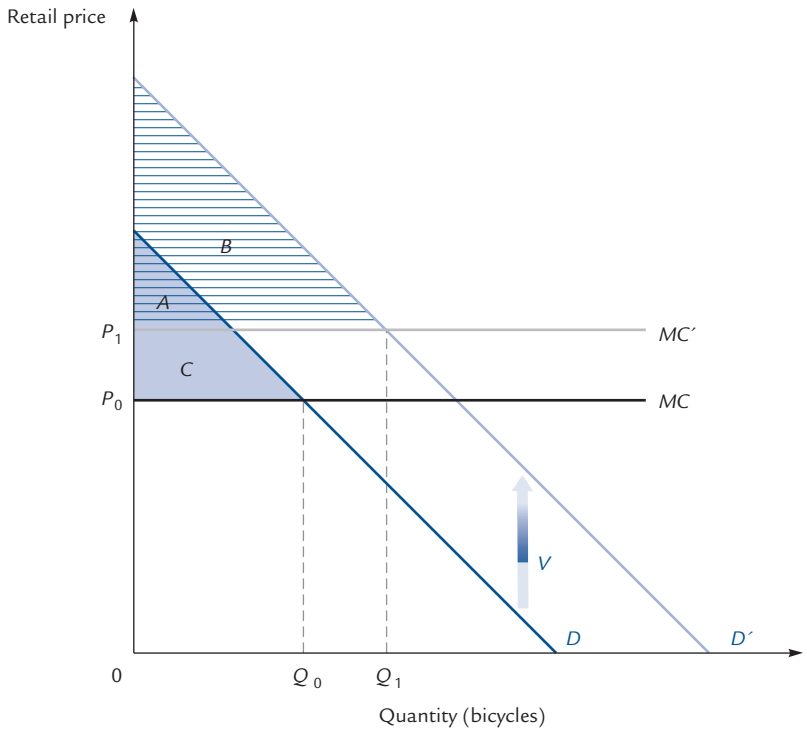
Exhibit 11.4 illustrates the theory. Suppose that P_0 is the wholesale price at which Schwinn sells its bicycles, and suppose, for simplicity, that retailers have no costs other than purchasing the bicycles from Schwinn. The retailers' marginal cost curve MC is flat at P_0, and if the retail market is competitive, they sell Q_0 bicycles, where MC meets the demand curve D. Now suppose that Schwinn sets a retail price of P_1 and requires all dealers to adhere to this price. Dealers will then compete for customers by providing additional services up to the point where the cost of providing these services is $P_1 - P_0$. This raises their marginal cost curve to MC'.

Exercise 11.3 Explain why dealers provide services exactly up to the point where the cost of providing them is $P_1 - P_0$.

We assume that the dealer services add some quantity V to the value of each bicycle; thus, the demand curve moves vertically upward a distance V to D'. The new quantity sold is Q_1, where MC' meets D'.

Notice that Schwinn would engage in this practice only if Q_1 is greater than Q_0; Schwinn wants to maximize the number of bicycles it can sell at a given wholesale

EXHIBIT **11.4** **Resale Price Maintenance**



Suppose that Schwinn provides bicycles at a wholesale price of P_0 and that this is the only cost that retailers have. If the demand curve is *D,* then under competition the quantity sold is Q_0 and consumers' surplus is $A + C$.

If Schwinn maintains a retail price of P_1, dealers compete with each other by offering services that cost $P_1 - P_0$ per bicycle to provide. The value of these services to consumers is some amount V, so that the demand curve moves vertically upward a distance V to D'. The new quantity sold is Q_1.

Because Schwinn chooses to engage in the practice, we can assume that $Q_1 > Q_0$. Elementary geometry now reveals that $V > P_1 - P_0$ (the value of the dealer services exceeds the cost of producing them) and $A + B > A + C$ (consumers' surplus is increased).

price. It is an easy exercise in geometry to check that if $Q_1 > Q_0$, then $V > P_1 - P_0$; that is, the value of the dealer services to consumers exceeds the cost of providing those services. This, in turn, by another easy exercise in geometry, implies that area B is greater than area C, so that, for a given wholesale price P_0, the consumers' surplus with resale price maintenance $(A + B)$ is greater than the consumers' surplus without resale price maintenance $(A + C)$.

Exercise 11.4 Perform the easy exercises in geometry.

Do not confuse the demand curves in Exhibit 11.4, which are the demand curves facing retailers, with the demand curve facing Schwinn. The demand curve facing Schwinn passes through the point (P_0, Q_0) without resale price maintenance, and it moves out to pass through the point (P_0, Q_1) when resale price maintenance is allowed.


Dangerous Curve


Dangerous Curve

The analysis (in Exhibit 11.4) is incomplete, because it takes the price P_0 as given. In fact, when resale price maintenance makes bicycles more attractive to consumers, the demand curve facing Schwinn moves out, leading Schwinn to set a new, higher price for bicycles. As a result, consumers keep only some of the increase in social welfare, and Schwinn gets the rest. Nevertheless, with the assumptions made here, it is possible to show that even after the price rises, consumers' surplus is still greater with resale price maintenance than without.

The theory that resale price maintenance exists to ensure a high level of service to customers is by no means the only one possible. A variety of other explanations have been offered. Indeed, in the same article where Professor Lester Telser first proposed the "service" argument, he went on to contend that it did *not* apply to resale price maintenance in the lightbulb industry, which was the special case that he was attempting to explain.[6] A recent study examined the evidence from a number of legal actions and found that the dealer service argument appears to correctly explain resale price maintenance approximately 65% of the time.[7]

The U.S. antitrust laws, as interpreted by the federal courts, severely limit the exercise of resale price maintenance. In May 1988, the Supreme Court issued a ruling that substantially relaxed these restrictions and made it easier for manufacturers to prevent retailers from offering discounts. In their decision, the justices called explicit attention to the role of resale price maintenance in maintaining high levels of dealer service. Later that week, the *New York Times* editorial page called for new legislation to overturn the effects of the ruling. The editorial called for giving manufacturers the right to "set high standards for service and refuse to supply retailers who don't meet them," while denying manufacturers the right to set prices.[8]

What the *Times* apparently failed to understand is that in the presence of competition among dealers, there is no difference between setting a standard for service and setting a retail price. Given a service standard, the price must rise until it just covers the cost of meeting the standard; given a price, the standard must rise until the cost of meeting it drives profits to zero. To allow manufacturers to set one but not the other is like allowing bathers to select the water level in the left half of the tub while disallowing them to select the water level in the right half. No matter how scrupulously you tried to obey such a law, you'd probably have trouble forcing yourself to forget that when you choose one level, you are automatically determining the other one.

Example: Barnes and Noble versus Amazon

Barnes and Noble is a large chain bookstore that offers a comfortable atmosphere for browsing. You can sit in comfortable chairs, sip coffee, and listen to music while you contemplate your selections. These amenities are costly to provide, in some ways that are obvious and other ways that are not so obvious. Barnes and Noble rents large amounts of space to give its customers elbow room. It keeps the shelves well-stocked, which not only invites damage and theft but also requires a substantial financial investment and hence a forgone opportunity to earn interest.

[6] This point is reinforced in L. Telser, "Why Should Manufacturers Want Fair Trade II?" *Journal of Law and Economics* 33 (1990): 409–417.
[7] P. M. Ippolito, "Resale Price Maintenance: Economic Evidence from Litigation", *Journal of Law and Economics* (1988).
[8] "Let the Retail Price Be Right," *New York Times* editorial, May 6, 1988.

Amazon.com is a Web-based virtual bookstore that offers the convenience of shopping at home. Amazon has fewer expenses than Barnes and Noble: Rather than providing you with elbow room, Amazon invites you to keep your elbows on your desktop. Rather than keeping a large number of books in stock, Amazon orders many books from suppliers only after they have been requested by customers.

Amazon passes some of its cost savings on to the customer. Many popular hardcovers are about 20% cheaper at Amazon. This means you have two choices: Shop in comfort at Barnes and Noble, where you can look at the books before you buy them, or shop at Amazon and save a few dollars.

Unfortunately for Barnes and Noble—and for the people who like to shop there—there's also a third option: Browse at Barnes and Noble and then buy from Amazon. Consumers who behave this way raise Barnes and Noble's costs and therefore reduce the amount of space and comfortable chairs that Barnes and Noble is willing to provide.

Under these circumstances, it is plausible that book publishers would want to engage in retail price maintenance—essentially forbidding Amazon to offer discounts, so that the service at Barnes and Noble is not diminished. (Publishers care about the quality of service at Barnes and Noble because it entices people to buy books.)

However, the issue in book publishing is less clear-cut than in the case of bicycles or stereo equipment. A discount bike shop or a discount stereo store offers nothing special except discounts. By contrast, Amazon offers a service that many customers value highly: The opportunity to shop without leaving home.

Therefore, publishers probably have mixed emotions about Amazon. On the one hand, it threatens Barnes and Noble and so drives away those readers who will only buy books in comfortable surroundings; on the other hand, it brings in a different class of readers who might never have shopped at Barnes and Noble. Thus, it's not clear whether publishers should want to stifle Amazon's business practices.

11.2 Collusion and the Prisoner's Dilemma: An Introduction to Game Theory

Collusion takes place when the firms in an industry join together to set prices and outputs. The firms participating in such an arrangement are said to form a **cartel.** By restricting each firm's production, the cartel attempts to restrict industry output to the monopoly level, allowing all firms to charge a monopoly price. This maximizes the total producers' surplus of all firms in the industry. If necessary, the resulting profits can then be redistributed among firms so that each gets a bigger "piece of the pie" than it had under competition.

Collusion

An agreement among firms to set prices and outputs.

Cartel

A group of firms engaged in collusion.

Collusion is an ancient phenomenon. In the tenth century B.C. the Queen of Sheba (near what is now Yemen) held a monopoly position in the shipment of spices, myrrh, and frankincense to the Mediterranean. When Solomon, the king of Israel, entered the same market, "she came to Jerusalem, with a very great train, with camels that bear spices, and very much gold, and precious stones," which could indicate how much she valued the prospect of an amicable agreement to divide the market.[9] More recently, Adam Smith observed:

[9] 1 Kings 10:2.

> *People of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices.*[10]

A more contemporary example dates from the year 2000, when the world's two largest auction houses, Christie's and Sotheby's, paid hundreds of millions of dollars in fines after conspiring to fix the commissions they charged sellers. In yet another example, the Justice Department charged the eight Ivy League universities with illegally colluding to coordinate their financial aid offers. At an annual meeting called Overlap, the Ivy League schools (and fifteen others) negotiated agreements on both a general formula for determining aid offers and the specific amounts that would be offered to individual students. Because of the universities' agreement not to bid against each other, many students paid more for their educations than they would have under competition. The Justice Department argued that this made the Overlap group an illegal cartel.

According to the *Wall Street Journal,* the colleges defended their practices as a way of ensuring that students would not be influenced by financial considerations in choosing a college.[11] This defense was at least novel: If the major auto manufacturers had been caught colluding to fix high prices, they might not have thought to argue that they were performing a public service by ensuring that consumers would not be influenced by financial considerations in choosing a car. But the Justice Department was unimpressed, and the Ivy League schools, without admitting wrongdoing, agreed to cancel Overlap and not to collude in the future.

Game Theory and the Prisoner's Dilemma

Cartels require cooperation. In order to understand the difficulties facing those who would cooperate, we will digress briefly into a topic from the theory of games.[12] The particular "game" we will analyze is called the *Prisoner's Dilemma*.

A crime has been committed and two suspects have been arrested. The suspects are taken to the police station and the district attorney meets with each one separately. To each she makes the following offer: "If you each confess, I'll send you both to jail for 5 years. If neither of you confesses, I can still get you on a lesser charge and send you to jail for 2 years each. If your buddy confesses and you don't, you'll get 10 years and he'll get 1. But if *you* are the only one to confess, you'll get off with 1 year while I put *him* away for 10. Now do you confess or don't you?" Each prisoner has to decide without conferring with the other.

Exhibit 11.5 will help you keep track of the district attorney's offer. Prisoner A, by choosing to confess or not confess, selects one of the columns in the table. Prisoner B selects one of the rows.

Let's evaluate the choices available to Prisoner A. What if B has confessed, thereby choosing the first row? Then A's choices are to confess and get 5 years, or to not confess and get 10 years. He should confess.

On the other hand, what if B has not confessed, thereby choosing the second row? Then A's choices are to confess and get 1 year, or to not confess and get 2 years. He should confess.

---

[10] Adam Smith, *The Wealth of Nations.*

[11] "U.S. Charges Eight Ivy League Universities and MIT with Fixing Financial Aid," *Wall Street Journal,* May 23, 1991.

[12] This theory was developed in the late 1940s by the mathematician John von Neumann and the economist Oscar Morgenstern. It has had a great deal of influence in economics and political science.

EXHIBIT **11.5** **The Prisoner's Dilemma**

		Action of Prisoner A: Confess	Action of Prisoner A: Not Confess
Action of Prisoner B: Confess		5 years each	A gets 10 years B gets 1 year
Action of Prisoner B: Not Confess		A gets 1 year B gets 10 years	2 years each

Each prisoner must decide whether to confess or not to confess. Prisoner A reasons that there are two possibilities: Either B confesses, in which case A is better off confessing (so that he gets 5 years instead of 10), or B does not confess, in which case A is better off confessing (so that he gets 1 year instead of 2). Regardless of B's action, A should confess, and regardless of A's action, B should confess. As a result, they each go to jail for 5 years, whereas if neither had confessed they would only have gone to jail for 2 years.

Needless to say, Prisoner A confesses. Following the same logic, so does Prisoner B. They both end up with 5 years in jail, even though they would have both been better off if neither had confessed.

It is easy to misunderstand the point of this example. Students sometimes think that Prisoner A confesses because he is afraid that Prisoner B will confess. In fact, A confesses for a much deeper reason. He confesses because it is his best strategy *regardless* of what B does. Prisoner A would want to confess if he knew that B had confessed and would also want to confess if he knew that B had not confessed. The same is true for B.


Dangerous Curve

The Prisoner's Dilemma and the Invisible Hand

The Prisoner's Dilemma is an interesting case in which the invisible hand theorem is not true. When each party acts in his own self-interest, the outcome is not Pareto-optimal. If neither confessed, both would be better off. We saw in Chapter 8 that in competitive markets, by contrast, the equilibrium outcome is always Pareto-optimal. The fact that the invisible hand can fail in a simple example like the Prisoner's Dilemma makes its success in competitive markets all the more remarkable.

Solving the Prisoner's Dilemma

How can the Prisoner's Dilemma be solved? Suppose that the prisoners of Exhibit 11.5 are members of a crime syndicate that can credibly threaten to impose severe penalties on anyone who confesses. Then the individual prisoners can be induced not to confess, and both will be better off. Contrary to what your intuition may tell you, they both benefit by being "victims" of coercion. (More precisely, each benefits from the coercion applied to the other, and this benefit exceeds the cost of the coercion applied to himself.)

Therefore, it is possible that people will prefer to have their options limited in situations that resemble the Prisoner's Dilemma. In China before World War II, goods were commonly transported on barges drawn by teams of about six men. If the barge

reached its destination on time (often after a journey of several days), the men were rewarded handsomely. On such a team any given member has an incentive to shirk, in the sense of working less hard than is optimal from the team's point of view. This incentive exists regardless of whether he believes that the others are shirking. Thus, the situation is similar to the Prisoner's Dilemma, with the choices "Confess" and "Not Confess" replaced by "Shirk" and "Don't Shirk." As in the Prisoner's Dilemma, an outside enforcer commanding everyone not to shirk can make everyone better off. In recognition of this, it was apparently common for the bargemen themselves to hire a seventh man to whip them when they slacked off!

The Repeated Prisoner's Dilemma

The Prisoner's Dilemma becomes a far richer problem when the two players expect to meet each other repeatedly in similar situations. Even though Prisoner A can always do better in the current game by confessing, he must also worry about whether his actions today will influence Prisoner B's actions tomorrow.

Suppose that A and B plan to play the Prisoner's Dilemma on three separate occasions: Monday, Tuesday, and Wednesday. You might think that each prisoner would have some incentive not to confess on Monday, so that he develops a reputation for reliability. Let us see whether this is true.

We begin by imagining the situation on Wednesday, which is the easiest day to think about. Because Wednesday is the last day, there are no future games to consider, and the game is just like an ordinary Prisoner's Dilemma. Regardless of what has gone before, each prisoner has the usual incentive to confess.

Now let us imagine the situation on Tuesday. Suppose that on Tuesday Prisoner A does not confess in order to convince Prisoner B that he won't confess on Wednesday. Will Prisoner B believe him? No, because Prisoner B realizes that once Wednesday arrives, Prisoner A will surely want to confess. Because he cannot convince Prisoner B of his goodwill anyway, Prisoner A confesses on Tuesday as well. By the same logic, so does Prisoner B.

Finally, how will the prisoners behave on Monday? Each one knows, by the logic of the preceding paragraph, that the other will confess on Tuesday. Thus, there is no credibility to be gained by not confessing on Monday. Both, therefore, confess on Monday as well.

The same reasoning applies to any repeated Prisoner's Dilemma with a definite ending date. By reasoning backward from that ending date, we see that there is never any incentive to establish a good reputation, because no such attempt can ever be credible. When there is no definite ending date, the analysis of the repeated Prisoner's Dilemma becomes a subtle and difficult problem.

Tit-for-Tat

In 1984, Professor Robert Axelrod of the University of Michigan announced the results of a remarkable experiment.[13] Axelrod had invited various experts in the fields of psychology, economics, political science, mathematics, and sociology to submit strategies for the repeated Prisoner's Dilemma. Using a computer, he invented one imaginary prisoner with each strategy, and he had each prisoner play against each other prisoner in a 200-round repeated game. Each prisoner also played one 200-round game against a

[13] His results are reported in a fascinating book, *The Evolution of Cooperation* (New York: Basic Books, 1984).

carbon copy of himself, and one 200-round game against a prisoner who always played randomly. The jail sentences from Exhibit 11.5 were translated into points as follows:

Sentence	Points
1 year	5
2 years	3
5 years	1
10 years	0

One of the strategies submitted was called *Tit-for-Tat*. According to the Tit-for-Tat strategy, the prisoner does not confess in the first round. In future rounds he continues not confessing, except that if the opponent confesses, then the Tit-for-Tat player punishes him by confessing in the next round. In subsequent rounds, he returns to not confessing, confessing only once as punishment each time his opponent confesses.

Tit-for-Tat won the tournament decisively. Thereupon, Axelrod organized a new and much larger tournament with 62 entrants. In the second tournament the lengths of games were determined randomly, rather than making them all 200 rounds. Also, all participants in the second tournament were provided with detailed analysis of the outcome of the first tournament, so that they could use these lessons in designing their strategies. Once again, Tit-for-Tat, the simplest strategy submitted, was the decisive winner.

In a final experiment, Axelrod used his computer to simulate future repetitions of the tournament. He assumed that the strategies that did well would be more widely submitted as time went on. Thus, a strategy that did well in the first tournament, like Tit-for-Tat, was replicated many times in the second tournament, whereas strategies that did less well were replicated fewer times. This was intended to mimic evolutionary biology, where those animals that succeed in competition have more offspring in future generations. As the tournament was repeated, one could observe the evolution of various strategies. The chief result was that Tit-for-Tat never lost its dominance.

The success of Tit-for-Tat has a paradoxical flavor, in view of the fact that the backward reasoning of the preceding subsection suggests that there is no gain to acquiring a reputation for playing "reasonably" in a repeated Prisoner's Dilemma. The success of Tit-for-Tat seems to rely on just such reputational effects. Thus, we have a puzzle. Economists don't always have all the answers.

The Prisoner's Dilemma and the Breakdown of Cartels

We now return to the topic of cartels. In a cartelized industry, price is set above marginal cost. In order to maintain this price, industry output must be held below the competitive level, and each firm is assigned a share of this production. Because price exceeds marginal cost, any given firm can increase its profits by selling a few more items at a slightly lower price. Of course, this increased output will tend to lower the price and to reduce industry-wide profits. For this reason, a monopolist would resist the temptation to increase output. However, a member of the cartel who "cheats" by increasing its output beyond its allotted share will reap all of the benefits from its action while bearing only some of the costs. It gets all of the additional revenue from the increment to output, whereas everybody shares the losses due to the fall in price.

It follows that a cartel member will be less mindful of the negative consequences of its actions than a single monopolist would be. It tends to cheat when it can get away

EXHIBIT **11.6** **The Breakdown of Cartels**

		Action of Firm A: Cheat	Action of Firm A: Not Cheat
Action of Firm B: Cheat		$5 profit each	A gets $3 profit B gets $12 profit
Action of Firm B: Not Cheat		A gets $12 profit B gets $3 profit	$10 profit each

Each member of the cartel must decide whether to cheat by producing more than the agreed-upon output. Cheating will increase the cheater's profits (because price is higher than marginal cost) and decrease the other firms' profits (by driving down the price of the product). It is in each firm's interest to cheat, whether it believes the other firm is cheating or not.

with it, and so does every other member of the cartel. Eventually, output increases all the way out to the competitive level.

The breakdown of cartels is perfectly analogous to the Prisoner's Dilemma. Imagine two firms, A and B, who have formed a cartel and must decide whether to abide by the agreement or to cheat. They are confronted by the options shown in Exhibit 11.6. Reasoning exactly as in the Prisoner's Dilemma, each firm chooses to cheat, and the cartel breaks down.

If a cartel is to succeed, it needs an enforcement mechanism. That is, it needs a way to monitor members' actions and a way to punish those who cheat. Because price-fixing agreements are illegal in the United States, the enforcement must be carried out in secret. (Indeed, since the *Madison Oil* case of 1940, the courts have held that even an attempt to fix prices is illegal under the Sherman Act, regardless of whether the attempt is successful.) Whenever you hear it asserted that a cartel has been successful, your first question should be: What is the enforcement mechanism?

Example: The NCAA

The nation's colleges are suppliers of intercollegiate sports, and the television networks are demanders. In order to extract high prices from the networks, colleges want to limit the number of teams and the number of games they play each season. But the Prisoner's Dilemma makes this difficult: Each college wants to play additional games to earn additional revenue, regardless of how the other colleges are behaving.

To prevent such "cheating," most colleges have joined the National Collegiate Athletic Association (NCAA) and given it the right to regulate their sports programs. For a long time, the NCAA also negotiated directly with the television networks, but the Supreme Court ruled in 1984 that these negotiations were illegal and that individual colleges could negotiate separately with the networks.

You might think that colleges would benefit from their new negotiating power. The opposite is true. Now that they can negotiate separately, it has become harder to enforce the cartel agreement, as a result of which more games are played and revenues from television have fallen. However, the NCAA still wields considerable power and keeps revenues substantially higher than they would otherwise be.

Example: The Dairy Compact

On its face, dairy farming is a highly competitive industry. However, dairy farmers in the eastern United States maintain artificially high milk prices through a cartel organization that sets and enforces minimum prices. Why is there a successful cartel in dairy farming and not, say, in wheat farming? The simple answer is that dairy farming is, through acts of Congress, exempt from antitrust laws that would make cartelization illegal. This allows the cartel to operate out in the open and to perform effectively.

The next question is: Why have dairy farmers won an exemption from the antitrust laws when wheat farmers have not? The author of your textbook does not know the answer to this question.

Example: Concrete Pouring and Organized Crime

Throughout the 1980s, the concrete-pouring industry in New York City was dominated by a cartel of six firms called "The Concrete Club." Whenever a project was put out for bids, the Concrete Club chose one of its members to handle that project and agreed that no member of the Club would attempt to underbid that firm. As a result, the cost of a cubic yard of concrete rose to $85, the highest in the nation.

Without a strong enforcement mechanism, it would be very difficult for a cartel like the Concrete Club to succeed. Not only would its own members be tempted to cheat but competition from nonmembers would soon drive prices down to the competitive level.

In this case, the enforcement mechanism was provided by New York's organized crime families, who managed the cartel and imposed heavy penalties on cheaters. Competition from outside the cartel was eliminated by the families' control of the Concrete Workers Union, which prevented non-Club members from working on any project involving more than $2 million.[14]

Example: The International Salt Case

To succeed, a cartel must know when its members are cheating. The International Salt Company may have discovered a creative solution to this monitoring problem. The company distributed a patented machine called the Lixator, which was used to dissolve rock salt. In some areas of the country, Lixators were sold outright; in others, they were leased subject to a requirement that the lessee agree to purchase all of its salt from International. In 1947 the Supreme Court ruled, in effect, that International Salt had attempted to create monopoly power in the market for salt. According to the analysis of two-part tariffs in Section 10.3, this explanation is unlikely to be correct. Instead, that analysis suggests that International was price discriminating by effectively charging heavier users more for a Lixator.

In 1985, John Peterman of the Federal Trade Commission reviewed the evidence and found that the economists' explanation was also suspect.[15] He discovered a clause in the

[14] The information in this section is taken partly from J. Cummings and E. Volkman, *Goombata* (Little Brown, 1990) and partly from P. Maas, *Underboss* (HarperCollins, 1997).

[15] John Peterman, "*The International Salt* Case," *Journal of Law and Economics* 22 (1985): 351–364.

Lixator rental contract that allowed any firm to buy its salt elsewhere if it could find it at a price lower than International's. Thus, International could not have charged more than the going market price for salt; if it had, it wouldn't have sold any.

What, then, could account for the structure of the Lixator contract? Here is one intriguing possibility. Suppose that salt suppliers were colluding. In that case, they would have needed a way to gather information on which suppliers were undercutting the agreement, so that the cheaters could be punished. The Lixator contract, with the clause that Peterman discovered, gave International's own customers an incentive to report low salt prices to International. In this way International could be continually informed of who the price cutters were and how much they were charging.

The Government as Enforcer

When cartels have been successful, the outside enforcer has often been the government. The most candid example in U.S. history is the National Industrial Recovery Act of 1933, under the provisions of which government and industry leaders met together to plan output levels with the explicit purpose of keeping prices artificially high. The act was unanimously declared unconstitutional by the U.S. Supreme Court two years after its inception.

A more subtle channel through which government plays the role of enforcer is the apparatus of the various federal regulatory agencies. You may be surprised to learn that many industries welcome regulation. A firm that wants to be told how much to produce seems as unlikely as a bargeman who wants to be whipped. Yet, like the bargeman, the firm can find itself in a Prisoner's Dilemma where it benefits from having its actions restricted. In the next section we will explore some of the more common forms of regulatory activity.

Monopolies as Enforcers

In Section 11.2, we saw that Wal-Mart has been accused of predatory pricing—charging artificially low prices for prescription drugs in order to drive competitors out of business.

If that was in fact Wal-Mart's intention, how would drug manufacturers like Merck and Pfizer respond? Two thoughtful economists[16] observe that a Wal-Mart monopoly, like any monopoly, would maintain high retail prices by restricting quantities, which is bad for the manufacturers. Therefore, the economists argue, the manufacturers would attempt to thwart Wal-Mart's predatory pricing through practices like resale price maintenance, requiring Wal-Mart to charge as much as its competitors. Ironically then, the laws against one "monopolistic" practice (namely resale price maintenance) make it harder for manufacturers to combat another monopolistic practice (namely price discrimination).

But alternative theories are possible. Suppose that Merck and Pfizer want to form a cartel. Because of Prisoner's Dilemma issues, they need an enforcer. Conceivably, a monopoly retailer could serve as that enforcer, by refusing to sell more than the agreed-upon quantities of any drug. Side payments among Wal-Mart, Merck, and Pfizer could then ensure that everyone shares in the profits from cartelization. Thus drug manufacturers might welcome monopoly power in the retail market.

[16] D. Boudreaux and A. Kleit, "How the Market Self-Polices Against Predatory Pricing," *Antitrust Reform Project* (June 1996).

It has been argued that the United Auto Workers (UAW), which has monopoly power in the market for labor, serves as a cartel enforcer for American auto makers; the idea is that the auto makers implicitly agree to produce restricted quantities of cars and the UAW enforces the cartel by refusing to provide additional labor to any manufacturer who attempts to exceed the agreed-upon quantities. If this theory is correct, car manufacturers should be glad that the UAW has monopoly power. How might you go about testing such a theory?

11.3 Regulation

In the United States, as in most industrialized countries, government regulation touches nearly every aspect of economic activity. Government agencies regulate hiring practices and working conditions, limit entry into professions as diverse as medicine and cosmetology, and dictate environmental standards that affect the design of everything from your car to your showerhead. Regulations are highly varied in their justifications, their effects, and the institutional arrangements through which they are enforced. Many different agencies are empowered to devise and enforce economic regulations. Some of these agencies function independently, while others are subsidiary to an executive department. Also, legislatures often pass specific statutes that are designed with regulatory intent.

Regulation has a wide variety of effects and purposes. Among these are the protection of consumers, the promotion of competition, and even the career interests of the regulators themselves. Another aspect of regulation is that it can sometimes serve to lessen competition in designated industries by introducing the government as the enforcer of a *de facto* cartel.

In the examples that follow, we will emphasize the *cartel enforcement* role of regulation, because that is the aspect of regulation that is relevant to the subject of this chapter. Do not allow this emphasis to mislead you into thinking that other aspects of regulation are less important or less interesting; they are only less germane to this discussion.

Examples of Regulation

Regulating Quantity

In the United States, the Interstate Commerce Commission (ICC) regulates railroads and trucking, and the Federal Aviation Administration (FAA) regulates airlines. No trucking company can operate without authority from the ICC and no airline can operate without authority from the FAA.

It has not always been easy to obtain that authority. For many years, the ICC routinely denied applications to enter the trucking industry and strictly limited the activities of existing firms by specifying the routes they were allowed to serve and the types of freight they were allowed to haul. These strict practices kept the price of trucking services high and were therefore vocally supported by trucking firms. The FAA was comparably strict about controlling entry by new airlines and the routes that existing airlines were allowed to serve.

Over the past two decades, with the encouragement of both parties in Congress, both the ICC and the FAA have significantly curtailed their regulatory activities. One result is that prices in both industries have fallen substantially—in the case of the airline industry, by about 50% over the past two decades.

But regulatory attempts to limit entry into other industries continue. Recently, the U.S. government has taken steps to limit entry into medical specialties, actually going so far as to pay $100 million to 42 New York hospitals in exchange for their *not* training doctors to become specialists. At around the same time, the University of California hospitals agreed to eliminate 452 residencies. The combined effect will be to raise the price of specialized medical care.

Regulating Quality

Regulation often takes the form of minimum *quality* standards. By preventing goods below a prescribed minimum quality from reaching the marketplace, such regulations increase the market power of those suppliers whose output meets the prescribed standards. You might think that consumers always benefit when the average quality of goods increases, but a moment's reflection will convince you that this need not be the case. Few would prefer to live in a world in which every car had the quality (and the price tag) of a Rolls Royce. Many consumers choose goods of lower quality because they would rather devote more income to other things. The poor choose goods of lower quality more frequently, and they are therefore hurt disproportionately when low-quality goods disappear from the marketplace. A poor man who is permitted to purchase steak but not hamburger might have to eat potatoes instead of meat.

In 1989, there were two kinds of bread widely available in Egyptian retail markets. The lower-quality product sold for the equivalent of 0.8¢ U.S. per loaf, while the higher-quality product sold for 2¢. By the middle of 1990, the government forced the cheap bread to be withdrawn from the market. For many Egyptians, the results were disastrous. The *New York Times* reported the plight of a family of six, each of whom ate one loaf per meal.[17] Because they were forced to buy the more expensive bread, the family's food expenses increased by more than $10 per month—a quarter of their income. There is no sense in which this family can be said to have benefited from the new minimum quality standard.

But there are some markets, such as the market for drugs, where low-quality products can be harmful or even fatal. In those markets, many people will instinctively agree that minimum quality standards must be beneficial to consumers. Therefore it can be particularly instructive to investigate such markets to determine the actual effects of regulation.

In the United States, the sale of nonnarcotic drugs was largely unregulated until 1938. In that year, the Food and Drug Administration (FDA) first began requiring consumers to obtain a doctor's prescription before buying drugs. Have mandatory prescriptions improved consumers' health? Professor Sam Peltzman of the University of Chicago investigated this question in two ways: (1) by comparing American death rates before and after 1938; and (2) by comparing American death rates with death rates in other countries where prescriptions are still not mandatory. (Except for Argentina and Uruguay, most Latin American countries do not require prescriptions. Neither does Greece, and neither do many countries in Asia.) Peltzman concluded that, while the available evidence is too weak to support a firm conclusion, it appears that mandatory prescriptions do not save lives or lead to other improvements in health.[18]

In 1962, the U.S. Congress passed the Kefauver Amendments, which required drug manufacturers to prove that their products are safe and effective; the Kefauver

[17] "2 Cent Loaf Is Family Heartbreak in Egypt," *New York Times,* July 9, 1990.
[18] S. Peltzman, "The Health Effects of Mandatory Prescriptions," *Journal of Law and Economics* 30 (1987): 207–238.

Amendments are enforced by the FDA. To investigate the effect of this regulation, Professor Peltzman looked at the rate of new-product development in the drug industry both before and after 1962, and concluded that the Kefauver Amendments have cost more lives than they have saved.[19]

For nearly 40 years, the Kefauver Amendments have saved some lives by protecting consumers from harmful drugs. At the same time, they have cost other lives by delaying the appearance of useful drugs; people have died while drugs that could have saved them were still being tested. Because the cost of testing is a disincentive to innovate, the amendments have probably cost additional lives by reducing the number of new drugs that are developed in the first place. They have also raised the price of existing drugs by reducing the number of substitutes.

Peltzman estimated such costs and benefits by observing the behavior of pharmaceutical companies both before and after 1962. He found that the net effect was overwhelmingly negative. The amendments reduced the number of new drugs entering the marketplace from approximately 41 per year to approximately 16 per year, and they introduced an average delay of two years for a drug to reach the marketplace. In recent years, partly because of studies like Peltzman's and partly in response to the spread of AIDS, the FDA has relaxed its rules substantially, allowing new and important drugs to be fast-tracked into the marketplace.

The FDA regulates not only the quality of drugs but also of medical devices and food additives. A few years ago, the fast-track program was extended to apply to medical devices. In many areas, though, FDA approval continues to take a long time. It was not until December 1997, after many years of delay, that the FDA approved irradiation of meat products for controlling disease- causing microorganisms. The FDA concluded that irradiation is a safe and important tool to protect consumers from food-bornediseases, effectively acknowledging that for several years it had denied consumers access to a safe and effective means of protecting their health. Of course, if irradiation had turned out to be harmful, the years of delay might have been a great blessing to consumers.

Frequently, quality regulations take the form of professional licensing requirements. Your doctor, your lawyer, your cab driver, and your beautician all need licenses to practice. Such requirements can help to establish minimal standards of competence; they can also restrict the number of practitioners and thereby keep prices above the competitive level.

Regulating Information

Another way in which entry to a market can be effectively curtailed is by restricting the ability of consumers to learn about new suppliers. Suppliers who cannot make their existence known are essentially excluded from the market. In practice, this is often accomplished through restrictions on advertising. Professional societies such as the American Medical Association and the American Bar Association have gone to extraordinary lengths to restrict advertising by their members.

Many reasons have been offered to support the idea that advertising raises prices. It is sometimes alleged that buyers must "pay for the advertising as well as the product." On the other hand, advertising saves the consumer the cost of having to search for information about available products. Indeed, a buyer who prefers not to pay for

[19] S. Peltzman, "An Evaluation of Consumer Protection Legislation: The 1962 Drug Amendments," *Journal of Political Economy* 81 (1973): 1049–1091.

advertising always has the option to incur the costs of seeking out a seller who does not advertise and to buy the product at a correspondingly lower price. When buyers do not do this, they reveal that they value the informational content of advertising at a price at least equal to whatever they are paying for it.

In fact, by providing information about a wide array of sellers, advertising can promote competition and might therefore actually *reduce* prices. In 1972, Lee Benham set out to investigate this question in the market for eyeglasses.[20] This market was particularly suitable for study since there is wide variation in advertising restrictions across states. He found that in states where advertising was prohibited, the price of eyeglasses was higher by 25 to 100%. This particularly persuasive empirical study has convinced many economists that the net effect of advertising is often (though surely not always) to lower prices.

Regulating Prices

Instead of setting quality standards, the government sometimes sets minimum prices below which goods cannot be sold. This excludes the producers of low-quality goods from the marketplace, increasing the demand for those high-quality goods that are close substitutes.

By far the most important example is the federal minimum wage law. Although this law is often presented as protective of the unskilled, it is precisely they whom it excludes from the labor market. At a minimum wage of $5.15 per hour, someone who produces $3.00 worth of output per hour will not be hired to work. Overwhelming empirical evidence has convinced most economists that the minimum wage is a significant cause of unemployment, particularly among the unskilled.

Among the beneficiaries of the minimum wage law are the more highly skilled workers who remain employed and who can command higher wages in the absence of less-skilled competition. These more highly skilled workers tend to be represented by labor unions, which, not surprisingly, tend to support increases in the minimum wage.

Minimum wage laws also have other, less obvious effects. When the federal minimum wage was first proposed in the 1930s, it was heavily supported by the northern textile industry. The reason was that wages were lower in the South than in the North, due partly to a lower cost of living in the South. As a result, northern firms found it difficult to compete. By imposing a federally mandated minimum wage, northern producers hoped to eliminate the advantage held by their southern competition and indeed hoped to drive the South out of textile manufacturing altogether.

Regulating Business Practices

Laws that prohibit transactions at certain times of the day or week tend to inhibit competition and raise prices. So-called blue laws in many states prohibit the sale of various goods on Sunday. This solves a Prisoner's Dilemma for suppliers. Any given supplier must choose between the options "Work on Sunday" and "Not Work on Sunday." Each will choose to work on Sunday whether its competitors are doing so or not; but each prefers to have nobody working Sunday than to have everybody working. Blue laws allow the supplier to watch football on Sunday afternoon without losing business to a rival. Of course, this boon to suppliers comes at the expense of consumers, for whom Sunday is a convenient shopping day.

[20] L. Benham, "The Effect of Advertising on the Price of Eyeglasses," *Journal of Law and Economics* 15 (1972): 337–352.

An interesting variant of the blue laws was recently in effect in the city of Chicago. Until quite recently, it was illegal to buy meat in Chicago after 6 P.M. and repeal was opposed by the butchers' union.

The Economics of Polygamy

The laws against polygamy provide an instructive example of the effects of output restrictions. We will consider the effect of a law that forbids any man from marrying more than one woman.

We can view men as suppliers of "husbandships," which are purchased by women at a price.[21] This price has many subtle components, including all of the agreements, spoken and unspoken, that married couples enter into. Choices about where to live, how many children to have, who will do the dishes, and where to go on Saturday nights are all contained in the price of the marriage. When husbandships are scarce, men can require more concessions on such issues as conditions of their marriages. For example, if there were only one marriageable man and many marriageable women, the man would be in a position to insist that any woman he marries must agree to attend professional wrestling with him every weeknight (assuming that this is something he values). If one woman will not agree to this price, he can probably find another woman who will.

Thus, the price of a husbandship is higher when husbandships are scarce, and, similarly, the price of a husbandship is low when husbandships are abundant. If each man wanted to marry four women, the price of husbandships would be bid down (or, equivalently, the price of wifeships would be bid up) to the point where men would have to make considerable concessions in order to attract even one wife. It is in the interests of men as producers to restrict output so that this does not happen. Antipolygamy laws accomplish this. Thus, the analysis suggests that laws against polygamy, like other laws restricting output, benefit producers (in this case men) and hurt consumers (in this case women).

Sometimes students argue that no woman in the modern world would want to be part of a multiwife marriage and that therefore women could not possibly benefit from the legalization of polygamy. But this is incorrect, because even under polygamy those women who wanted to could demand as a condition of marriage that their husbands agree not to take any additional wives. And even if no man took more than one wife, the price of wives would still be higher.

For example, imagine a one-husband–one-wife family where an argument has begun over whose turn it is to do the dishes. If polygamy were legal, the wife could threaten to leave and go marry the couple next door unless the husband concedes that it is his turn. With polygamy outlawed, she does not have this option and might end up with dishpan hands.

Another reason why students are sometimes surprised by this conclusion is that they are aware of polygamous societies in which the status of women is not high. But, of course, the difference in polygamy laws is not the only important difference between those societies and our own. The fact that polygamy is legal in many places where women are otherwise oppressed does not constitute an argument that the oppression is caused by polygamy. Our analysis compares the status of women with and without legalized polygamy on the assumption that other social institutions are held constant.

[21] Because we are examining the market for husbands, men are the producers and women the consumers. It would be equally correct to treat the marriage market as a market for wives, in which women are the producers and men the consumers. Since we are investigating the effects of the law that restricts the supply of husbands, it is more convenient to think of "husbandships" rather than "wifeships" as the commodity being traded.

In view of our analysis, it is interesting that polygamy laws are often alleged to "protect" women. It has been observed that laws prohibiting any man from marrying more than one woman are perfectly analogous to laws preventing any firm from hiring more than one African-American.[22] Surely no one would be so audacious as to claim that the purpose of such a law was to protect African-Americans.

What Can Regulators Regulate?

In any study of the effects of regulation, it is necessary to ask what regulators actually do. But regulators' own descriptions of their activities should not always be taken at face value.

Economists George Stigler and Claire Friedland examined the effects of regulation in the electric power industry.[23] They examined electric rates in the years, 1912–1937. During these years, some states regulated the price of electricity and others did not. Stigler and Friedland found that the presence of regulation had no observable effect on the actual price of electricity. The evidence suggested that the regulatory commission consistently ended up setting the price that the utilities would have chosen anyway.

Stigler applied a similar analysis to the regulation of the securities industry by the Securities and Exchange Commission (SEC).[24] The SEC requires issuers of securities (e.g., corporate stocks) to make public disclosures of relevant information. If you try to sell stock in a gold mine that has never produced any gold, the SEC will require that this fact be disclosed to potential buyers. Stigler examined the performance of newly issued stocks compared with the performance of the market as a whole before and after the formation of the SEC in 1934. He found that there was no change in the propensity of newly issued stocks to perform well. It appeared that the SEC made no real difference; there is no evidence that the mix of securities that was offered under regulation differed appreciably from the mix of securities that would have been offered in an unregulated market.

These and other studies have convinced a growing number of economists that an industry should not necessarily be considered regulated just because of the existence of an agency with the formal power to regulate it. In many cases, there may be political or other considerations that prevent the agency from ever taking any steps that actually have the effect of altering economic behavior. Whether or not an allegedly "regulated" industry is really regulated in any meaningful sense is an empirical question, one that must be decided on a case-by-case basis.

Creative Response and Unexpected Consequences

Creative response
A response to a regulation that conforms to the letter of the law while undermining its spirit.

Although it can be in the interest of an industry to be regulated, it is almost always in the interest of an individual firm to avoid the effects of regulation when possible. This often leads firms to engage in **creative response,** behaving in ways that conform to the letter of the law while undermining its spirit. For this reason and others, regulations can have unexpected consequences—sometimes directly contrary to the intentions of the regulators.

[22] G. Becker, "A Theory of Marriage," *Journal of Political Economy* 81 (1973): 813–846.
[23] G. Stigler and C. Friedland, "What Can Regulators Regulate? The Case of Electricity," *Journal of Law and Economics* 82 (1974): S11–S26.
[24] G. Stigler, "Public Regulation of the Securities Market," *Journal of Business* 37 (1964).

Until a few years ago, parents traveling on airplanes were allowed to hold infants on their laps. More recently, parents have been required to buy a separate seat for the infant. This regulation, apparently motivated by a desire to make infants safer, has had exactly the opposite effect as many parents, unwilling to pay for the additional seats, have opted to travel by car instead of by airplane. Because the death rate per mile is about 70 times greater in a car, economists have estimated that the net effect of the regulation has been an increase in the number of infant deaths.

Another striking example concerns the use of pesticides. Certain pesticides are banned because of potential health hazards. But a side effect is to raise the cost of growing fruits and vegetables, thereby raising their price and lowering the quantity demanded. The prominent biologist Bruce Ames has pointed out that the fall in fruit and vegetable consumption is likely to be more damaging to health than the pesticides were.

Sometimes the unexpected consequences of regulation can be unexpectedly delightful. In renaissance Europe, regulations forbade unlicensed actors to speak on stage. According to some historians, the result was the advent of modern pantomime.

Here are some further examples from recent history, to demonstrate how creative responses can undermine the apparent intent of a regulation.

Example: Affirmative Action Laws

Affirmative action laws provide an example where a creative response may have led to consequences directly contradictory to the intent of the original legislation. These laws and regulations arose from the observation that African-American workers were systematically paid less than white workers. They required employers to remedy this imbalance by paying higher wages to African-American workers.

However, wages are only part of the compensation that a worker receives. Typically, workers receive a variety of valuable fringe benefits as well. One of the most important fringe benefits, especially in entry-level positions, is on-the-job training. Such training enables employees to acquire basic skills that will raise their income later in life. Its value often represents a substantial portion of the employee's total compensation.

Since on-the-job training is largely unobservable to outsiders, employers can adjust its quantity without being found guilty of violating those laws that regulate workers' compensation. Thus, some employers were able to comply with the affirmative action regulations without actually changing the total value of the compensation that they offered to African-Americans. They simply paid a higher wage, satisfying the regulator, while compensating by offering less on-the-job training. Between the years 1966 and 1974 the *observable* wage differences between African-Americans and non-Hispanic caucasians were essentially eliminated, but they were partially replaced by *unobservable* differences. For African-American workers, this meant higher starting salaries, less on-the-job training, and lower future wages than before affirmative action.

The net effect of all this on the economic status of African-Americans could be either positive or negative. In one study Professor Edward Lazear found that the relative economic status of African-Americans (taking account of all their expected future earnings) was not improved by the affirmative action laws.[25] In fact, his evidence supported just the opposite conclusion—that during the period 1966–1972, the gap between African-American and white compensation, inclusive of the value of on-the-job training, actually widened.

[25] E. Lazear, "The Narrowing of Black-White Wage Differentials Is Illusory," *American Economic Review* 69 (1979): 553–564.

Example: Reasonable Quantities of Sale Items

In the late 1970s, the Federal Trade Commission (FTC), which regulates (among other things) against false and deceptive advertising, discovered that one of its regulations led to responses that were counterproductive. The FTC periodically receives complaints about the unavailability of advertised specials. Consumers travel to stores that are advertising items at unusually low prices, only to find that those items are sold out shortly after the commencement of the sale. Understandably, these consumers are annoyed. The FTC responded to these complaints in the mid-1970s by issuing a series of regulations requiring stores to have on hand a "reasonable quantity" of any item that was advertised at a sale price.

To understand the effect of these regulations, it is necessary first to understand the reasons for sales. In many (though certainly not all) cases, a store will decide to discontinue stocking a certain item and will want to dispose quickly of its remaining stock. In such cases, ordering sufficient additional inventory to have a "reasonable quantity" on hand would contravene the very purpose of the sale. Therefore, one effect of the regulations was that sales of this type were discontinued. In view of this effect, fewer items were offered at sale prices. At the same time, it meant that when there *were* sales, the sale items were usually available.

Throughout the late 1970s, the FTC interviewed consumers about their feelings regarding the new rules. On the basis of these interviews, the FTC decided that the rules tended to benefit people with higher incomes at the expense of the poor. People with high incomes have a high value of time; they find it very costly to drive to a store only to discover that the item they are shopping for is out of stock. To them the cost of these fruitless shopping trips outweighs the benefit of having more sales to choose from. People with low incomes have a lower value of time and place greater value on being able to buy at sale prices. They prefer there to be more sales, even if the stores sometimes run out before they get there.

On the basis of this analysis, the FTC rescinded its rules on advertised specials.

Positive Theories of Regulation

Throughout this section we have examined some of the consequences of certain existing regulations. However, we have made no attempt to address the question of why some industries are regulated and others are not. We have focused primarily on ways in which regulation might act to limit competition. But we have made no attempt to formulate a general principle concerning when regulations will limit competition and when they will serve some other function, such as promoting economic efficiency.

Many economists think that there is a need for a positive theory of regulation, to predict the circumstances under which various types of regulations arise and what their effects will be. Such a theory would have to explain why the trucking industry is more heavily regulated than the airline industry, why some occupations require professional licenses while others don't, and why electricity prices seem to have been unaffected by regulation. A complete theory would begin with an explicit account of what it is that regulators are trying to accomplish. For example, regulators might be motivated by a desire to redistribute wealth in certain ways, or by a desire to protect consumers from major disasters, or even by a desire to maximize their own power. From such assumptions, one could derive conclusions about when, where, and what types of regulations are most likely to occur.

A theory of this sort might also be used to explain why regulations are selectively enforced. For example, radar detectors are legal in 48 states, despite the fact that their only purpose is to facilitate breaking the law. Why are people permitted to purchase the opportunity to violate speed limits with a reduced probability of punishment? Various theories are consistent with this observation. If the goal of regulators is to increase economic efficiency, they might want to allow speeding by those whose time is sufficiently valuable. These would be primarily those who find it worthwhile to invest in a radar detector. An alternative theory is that regulators prefer not to antagonize the politically powerful and that those who are wealthy enough to want radar detectors are also powerful enough to keep the regulators at bay.

Which theory seems more sensible to you? Can you think of other examples that would tend to confirm or refute one of these theories? What alternative theories can you propose?

11.4 Oligopoly

Oligopoly
An industry in which individual firms can influence market conditions.

An **oligopoly** is an industry in which the number of firms is sufficiently small that any one firm's actions can affect market conditions. Thus, in an oligopoly each firm has a certain degree of monopoly power. The behavior of such firms depends on many things, including whether they are threatened by potential entry. We will first consider markets in which entry is costless (and therefore an ever-present threat) and then markets in which the number of firms is fixed.

Contestable Markets

Contestable market
A market in which firms can enter and exit costlessly.

A market in which firms can enter and exit costlessly is called a **contestable market.**[26] A commonly cited example is the market for airplane service on a particular route, say, from Houston to Dallas. The owner of an airplane that is currently flying back and forth between Houston and San Antonio can easily move into the Houston-to-Dallas market if there is a profit opportunity, and can easily return to the Houston-to-San Antonio market at any time.

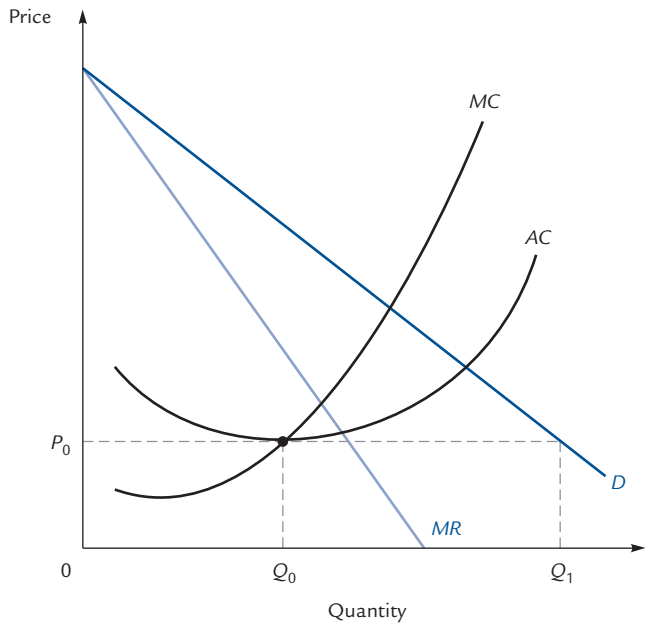
In a contestable market, even a single firm producing a unique product with no close substitutes might not be able to engage in monopoly pricing, because the profits that it would earn by doing so would lure entrants and destroy its market position. Exhibit 11.7 illustrates the position of a monopolist threatened by potential entry. Assuming all firms are identical, their entry price will be P_0.

Exercise 11.5 Explain why firms would enter if the market price of output were P_0 but would not enter at any lower price.

It follows that the market price cannot be higher than P_0, since any higher price will attract entry. At this price the firm will produce the quantity Q_0. The market will demand Q_1, which may be several times Q_0. If, for example, Q_1 is twice Q_0, there will be room for a second firm to imitate exactly the actions of the first firm without exhausting market demand. If Q_1 is seven times Q_0, there will be room for seven firms

[26] The theory of contestable markets is surveyed by its founders in W. Baumol, S. Panzard, and R. Willig, *Contestable Markets and the Theory of Industry Structure* (San Diego: Harcourt Brace Jovanovich, 1982).

EXHIBIT **11.7** **A Contestable Market**



If the market is contestable, firms will enter at any price above P_0. Therefore, the market price cannot be higher than P_0, because any higher price would attract entry. At this price the firm supplies Q_0 units of output and the market demands Q_1. Thus, there is room in the industry for Q_1/Q_0 firms.

altogether. In general, the number of firms that actually enter will be equal to Q_1/Q_0, each producing Q_0 items at a price of P_0, which equals both average and marginal cost.[27] In other words, potential entry will force firms to behave as competitors, even if there are very few firms.

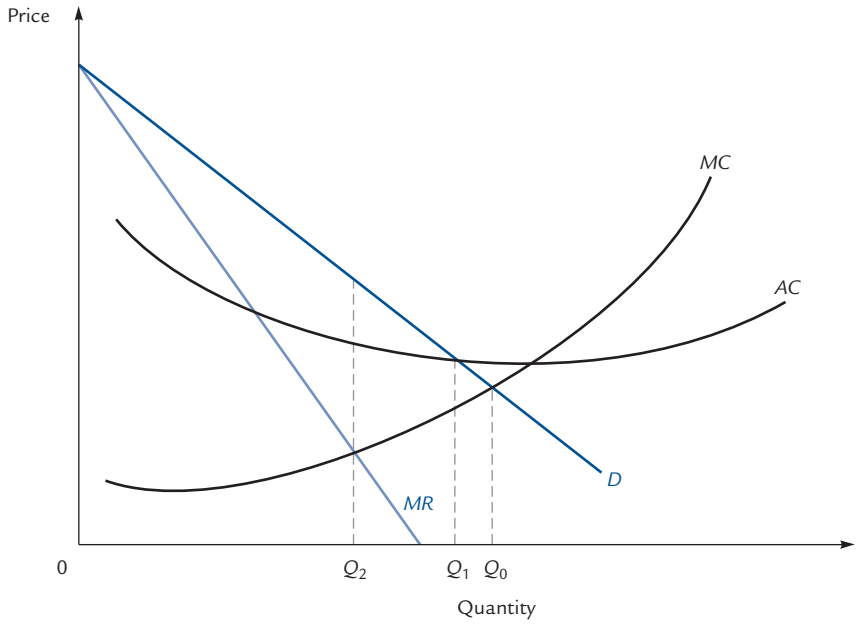
In a contestable market with identical firms whose average cost curves cross the industry demand curve in the region where they are upward sloping, price, average cost, and marginal cost are all equal.

Contestable Markets and Natural Monopoly

There is also the possibility of natural monopoly in a contestable market. That is, the firm's average cost curve might still be downward sloping where it crosses industry demand. This is shown in Exhibit 11.8. In this case, a monopoly producer cannot operate at the "competitive" point Q_0, because its profits there would be negative. On the other hand, if it follows the usual monopoly pricing rule of setting marginal cost equal to marginal revenue (producing Q_2), it may earn positive profits and lure other firms into the industry. The threat of entry forces the producer to operate at the zero profits point Q_1.

[27] There is a slight problem related to the fact that Q_1/Q_0 may not be exactly equal to an integer, in which case we expect the number of firms to be either the integer just above or just below Q_1/Q_0.

EXHIBIT **11.8** **Natural Monopoly in a Contestable Market**



If the market is contestable, a natural monopolist must set output at Q_1 so that it earns zero profits and avoids attracting entry.

Oligopoly with a Fixed Number of Firms

When there is no threat of entry, the behavior of an oligopoly is more difficult to predict. One possibility is the formation of a cartel. As we have seen, the Prisoner's Dilemma guarantees that there are forces tending to undermine the success of cartels. On the other hand, cartels are really *repeated* Prisoner's Dilemmas, since firms produce output every day. We have also seen that the outcome of repeated Prisoner's Dilemmas is hard to predict.

When there is no collusion, each firm's actions depend on the actions that it expects the other firms to take. Therefore, the way in which firms form their expectations about each other's behavior is a crucial ingredient in modeling oligopoly. We will examine two different models that proceed from different assumptions about expectation formation. In one, the **Cournot model,**[28] firms take their rivals' output as given. In the other, the **Bertrand model,**[29] firms take their rivals' prices as given.

Cournot model
A model of oligopoly in which firms take their rivals' output as given.

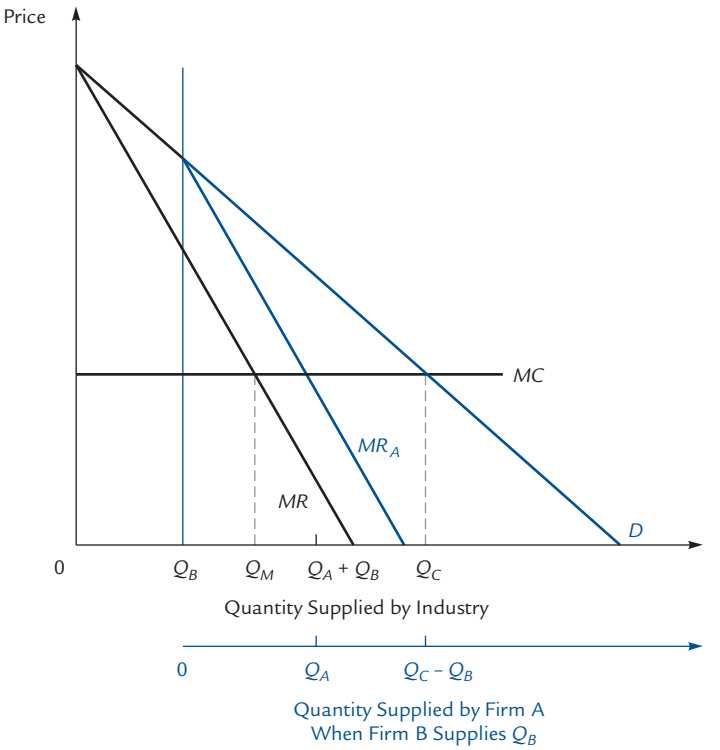
Bertrand model
A model of oligopoly in which firms take their rivals' prices as given.

The Cournot Model

To simplify the analysis, we will assume an industry with exactly two identical firms having the flat marginal cost curve shown in Exhibit 11.9. We will also assume a straight-line demand curve, so that marginal revenue has exactly twice the slope of

[28] For the nineteenth-century French mathematician Augustin Cournot.
[29] For the nineteenth-century French economist Joseph Bertrand.

EXHIBIT **11.9** **The Cournot Model of Oligopoly**



We assume that two identical firms have the flat marginal cost curve *MC* and face a market demand curve *D*. A competitive industry would produce the quantity Q_C. A monopolist would produce the quantity $Q_M = \frac{1}{2}Q_C$, where *MC* crosses the marginal revenue curve *MR*.

If Firm A assumes that Firm B will always produce quantity Q_B, then Firm A views itself as a monopolist in the market for the remaining quantity. The demand curve in that market is the colored part of the market demand curve, measured along the colored axis. The marginal revenue curve is the colored curve MR_A. Firm A produces the monopoly quantity Q_A, which is half the competitive quantity ($Q_C - Q_B$). Combining this fact with the equation $Q_A = Q_B$ (which follows from the fact that the firms are identical), we compute that $Q_A = Q_B = \frac{1}{3}Q_C$. Thus, the industry output is $\frac{2}{3}Q_C$, less than the competitive output but more than the monopoly output.

demand. A monopoly would produce the quantity Q_M and a competitive industry would produce the quantity Q_C. Because of what we have just said about the slopes of the curves, we must have:

$$Q_M = \frac{1}{2} Q_C$$

Now let us see what the two firms will produce. Suppose that Firm B produces the quantity Q_B and Firm A makes the assumption that this quantity will never change. Then Firm A views itself as a monopolist in the market for the remaining quantity. That is, Firm A is a monopolist in a market where the zero quantity axis is the colored vertical line in Exhibit 11.9 and the demand curve is the color part of the industry demand curve. In such a market, the marginal revenue curve is the color curve MR_A parallel to the industry marginal revenue curve *MR*. Firm A produces the quantity Q_A, where

$M_C = MR_A$. Since this is the monopoly quantity, it must lie halfway between Firm A's zero quantity axis at Q_B and the competitive point $Q_C - Q_B$. That is,

$$Q_A = \frac{1}{2}(Q_C - Q_B)$$

We can also write one additional equation. Because it is assumed that Firms A and B are identical, it is reasonable to expect that they will produce equal quantities of output. This gives us the equation;

$$Q_A = Q_B$$

Putting the two equations together, we get:

$$Q_A = \frac{1}{2}(Q_C - Q_A)$$

which can be solved for Q_A, giving:

$$Q_A = \frac{1}{3} Q_C$$

In other words, each firm produces 1/3 of the competitive quantity, so that between them they produce 2/3 of the competitive quantity. This is more than the monopoly output, which is only 1/2 of the competitive quantity.

The Bertrand Model

The Bertrand model has the same flavor as the Cournot model. In the Cournot model, each firm assumes that its rivals will never change quantity. In the Bertrand model, each firm assumes that its rivals will never change price.

As long as price exceeds marginal cost, an oligopolist in the Bertrand model will always want to undercut its rivals by offering a slightly lower price. Since it assumes that its rivals will not meet this price cut, it follows that the oligopolist will be able to capture the entire market for itself. This is a profitable strategy. The tiniest of price cuts leads to a sizable increase in sales, and all of these sales are at a price that exceeds marginal cost.

Bertrand oligopolists will continue to undercut one another until price falls to marginal cost. Thus, according to Bertrand, price and output will be the same under oligopoly as they are under competition.

Criticism of the Cournot and Bertrand Models

Many economists are uncomfortable with both the Cournot and the Bertrand models of oligopoly, because each model posits that firms make incorrect assumptions about their rivals' behavior. In the Cournot model, firms assume that their own choice of output will not affect their rivals' choices, despite the fact that they know that their rivals' choices are affecting their own. The same is true in the Bertrand model regarding prices instead of quantities.

This criticism highlights the major difficulty that economists face when they attempt to model oligopoly behavior. The assumptions that firms make about one another's behavior are crucial elements in the determination of their own behavior, and the economist must therefore presume to know something about those assumptions. If the assumptions turn out to be incorrect, firms should become aware of this fact over time, invalidating the model. In the real world we expect that oligopolists have at least reasonably accurate information about how their rivals behave, and we would like our models to reflect that fact. Unfortunately, satisfactory models with this property have proven difficult to construct. In much recent research, game theory has proved to be an increasingly effective tool.

11.5 Monopolistic Competition and Product Differentiation

Product differentiation
The production of a product that is unique but has many close substitutes.

One strategy for acquiring some degree of monopoly power in a market that is basically competitive is called **product differentiation.** As its name implies, this strategy involves producing a product that differs sufficiently from the output of other producers that some consumers will have a distinct preference for it. Crest and Colgate both produce toothpaste, but they do not produce identical products. The two products are close substitutes, and neither can be priced very differently from the other without a substantial loss of market share. At the same time, there are some consumers with a very strong preference for one or the other brand, so that each firm faces a demand curve that is at least slightly downward sloping.

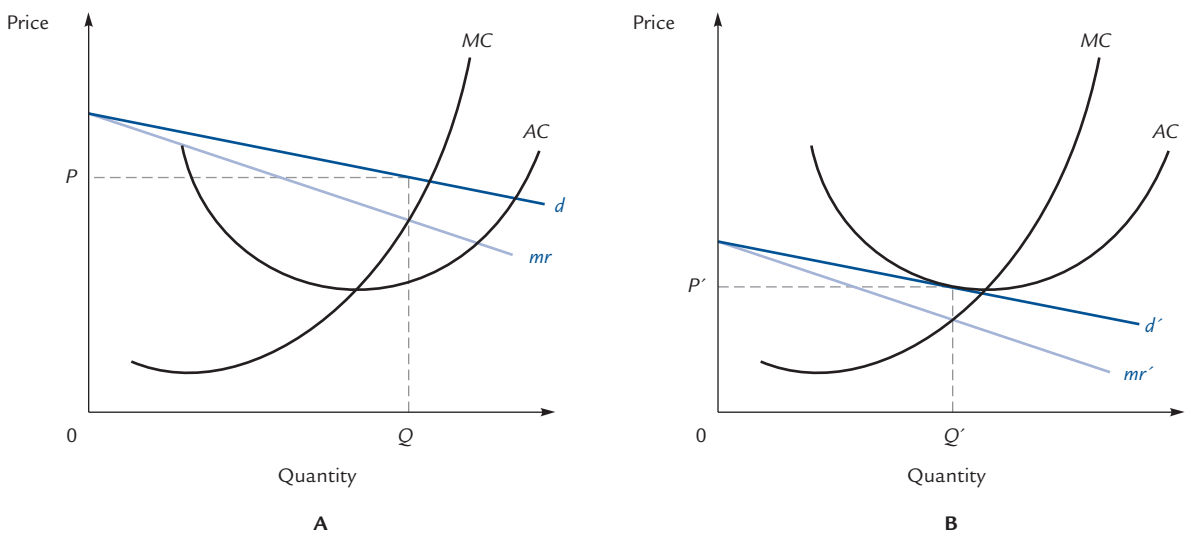
Products with brand names are product differentiated simply by virtue of having different brand names. But other characteristics can differentiate them as well. The location at which a product is sold can differentiate it from others. A 7-Eleven two blocks from your house is not the same to you as a 7-Eleven a mile and a half away, although they are probably close substitutes.

Monopolistic competition
The theory of markets in which there are many similar but differentiated products.

Monopolistic Competition

The theory of markets in which there are many similar but differentiated products is called the theory of **monopolistic competition.** The first panel of Exhibit 11.10 illustrates the conditions facing a monopolistically competitive firm. Suppose that the firm

EXHIBIT **11.10** **Monopolistic Competition**



Panel A shows a short-run equilibrium in which the firm sells quantity Q at price P. Here price exceeds average cost, so the firm earns positive profits. In the long run, entry drives the demand curve facing this firm down to d' in panel B, where the firm is just able to earn zero profits by selling quantity Q' at price P'.

is currently charging price P and selling quantity Q. The demand curve d shows how much the firm can sell at any given price on the assumption that all other firms continue to charge the original price P.

Exercise 11.6 Explain why you might expect the curve d to be quite elastic compared with the demand curve facing an ordinary monopolist.

The quantity Q is determined by the condition that $MC = mr$, where MC is the firm's marginal cost curve and mr is the marginal revenue curve associated with d. In panel A of the exhibit, the firm is earning positive profits, since the price P exceeds average cost at quantity Q.

In the long run, these profits will attract entry by other firms selling similar products. As a result, the demand curve facing the firm will shift downward, to d' in panel B of Exhibit 11.10. The firm produces quantity Q' and charges price P'. At this point price and average cost are equal, so that profits are zero and there is no further entry.


Dangerous Curve

At the long-run equilibrium quantity Q', the demand curve must touch the average cost curve to give zero profits. You might wonder why we have drawn the curves tangent rather than crossing. The reason is that if the curves crossed, the firm could earn positive profits by producing a quantity slightly less than Q'. But we know that Q', the zero-profits point, is also the point of maximum profits, since it is the point where $MC = mr$. Thus, it cannot be correct to draw the average cost curve actually crossing demand.

Welfare Aspects of Monopolistic Competition

In Exhibit 11.10 we can see that price is set above marginal cost by a monopolistic competitor, so that the level of output is suboptimal. On the other hand, since we expect monopolistic competitors to face quite elastic demand curves, the deviation of output from the competitive level might not be too great.

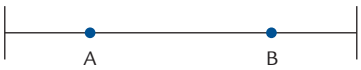
A related issue is that a monopolistic competitor, as shown in Exhibit 11.10, does not produce at the minimum point of its average cost curve. Indeed, it cannot do so, since in long-run equilibrium it produces at a point of tangency between its average cost curve and its downward-sloping demand curve. It follows that if a monopolistically competitive industry were replaced by a competitive one, the same output could be produced at lower cost.

It is sometimes argued that monopolistically competitive firms tend to invest more in advertising and other methods of luring each other's customers than is socially optimal. Insofar as such practices simply shift customers from one firm to another without changing the nature of the products that are sold, their costs represent unnecessary social losses.

Balanced against all of this is the observation that monopolistically competitive industries do provide consumers with something that competitive industries do not, namely, differentiated products. Although Burger King and McDonald's are already similar, many people would be unhappy if one of them became exactly like the other.

How can we weigh the inefficiencies associated with monopolistic competition on the one hand against the benefits of product differentiation on the other? Although many economists have strong beliefs about the relative importance of these phenomena, there is not yet any general theory available that allows us to answer such a question in a definitive way.

EXHIBIT **11.11** **Ice Cream Vendors on a Beach**



If the vendors start out in the locations shown, each will move toward the center in an attempt to gain more customers. The equilibrium is reached when they are located right next to each other and can move no farther.

The Economics of Location

Depending on market conditions, firms may choose either to exaggerate or to minimize their differences. An amusing example involves two ice cream vendors on a beach. Suppose that the beach is a straight line one mile long and that bathers are distributed evenly along it. There are two ice cream vendors, indistinguishable except for location, and each bather will patronize the nearest vendor. Where will the vendors locate?

Exhibit 11.11 shows the initial positions of the vendors. Given these positions, vendor A will soon realize that she can have more customers if she moves to the right. As long as she stays to the left of vendor B, she will retain all of the customers to her own left and she can acquire more by moving a bit to the right. Similarly, vendor B has much to gain and nothing to lose by shifting to the left. The only possible equilibrium is for the two vendors to locate right next to each other, exactly at the half-mile mark!

Exercise 11.7 What would happen if the vendors started out next to each other but somewhere other than at the halfway point?

Perhaps this example provides a metaphor for the behavior of the two major U.S. political parties. With voters distributed on a continuum from left to right, and voting for the party "closest" to themselves, the parties will behave just as the ice cream vendors do. Do you believe that this metaphor captures a significant feature of reality?

Summary

This chapter surveys a number of examples and models in which firms exercise or attempt to exercise various degrees of monopoly power.

Horizontal mergers can both reduce production costs and create monopoly power, and therefore they have ambiguous welfare consequences. Vertical mergers can have the effect of reducing the exercise of monopoly power, since no monopolist would want to extract monopoly profits from one of its own subsidiaries.

In order to eliminate rivals, a firm might engage in the practice of predatory pricing, or it might attempt a strategy of buying out its rivals. Each of these strategies is severely limited. In the case of predatory pricing, there is the threat that rivals will resurface after prices are raised. In the case of buy outs, new rivals are attracted to the industry by the prospect of being bought out.

When the firms in an industry can collude, they increase producers' surplus and thus can improve each firm's welfare through a system of side payments. However,

as in the Prisoner's Dilemma, each individual firm has an incentive to cheat. The reason is that a cartel sets price higher than marginal cost, so that each firm will want to sell more than it is supposed to under the cartel agreement. Therefore, cartels tend to break down unless there is a good enforcement mechanism.

In addition to its other purposes, government regulation can serve as an enforcement mechanism for a cartel. Regulations restrict output in many ways. Professional licensing, minimum quality standards, minimum prices, advertising restrictions, and blue laws can all serve to restrict output and keep prices high. However, there is some evidence that the power of regulators to alter market conditions is sometimes less than it seems.

In contestable markets, entry and exit are costless. Even when there is only one firm in a contestable market, that firm must earn zero profits because of the threat of entry.

The Cournot and Bertrand models apply to oligopolies with a fixed number of firms. In the Cournot model, firms take their rivals' output as given and end up producing more than the monopoly quantity but less than the competitive quantity. In the Bertrand model, firms take their rivals' prices as given and end up producing the competitive quantity.

Under monopolistic competition, firms produce differentiated products. Each firm's product is unique but is similar to those of other firms. Thus, each firm faces a downward-sloping but nevertheless quite elastic demand curve. In the long run, entry forces profits to zero, which implies that firms must *not* be operating at the point of minimum average cost. The negative welfare consequences of this must be balanced against the gains to consumers from having a wide variety of product options, but economists have developed no good general theory of the welfare consequences of monopolistic competition.

Author Commentary

www.cengage.com/economics/landsburg

AC1. For more information on Barnes & Noble *v.* Amazon, read this article.

AC2. For more information on Microsoft *v.* Netscape, read this article.

AC3. Just as sellers can cartelize to keep prices up, so buyers can cartelize to keep prices down. Ordinarily, keeping prices down is an end in itself. However, the National Football League attempts to control players' salaries, not just for its own sake, but also to prevent a few teams from dominating the league. Does it work?

AC4. Regulators have an incentive to create the need for more regulators. For an application of the same principle to judges, read this article.

Review Questions

R1. What is the distinction between a horizontal merger and a vertical merger?

R2. Under what circumstances is a horizontal merger welfare-improving?

R3. What are some of the advantages and disadvantages to a firm in engaging in predatory pricing? In a strategy of buying out rivals?

R4. Explain why resale price maintenance might be expected to benefit consumers.

R5. Why do both prisoners confess in the Prisoner's Dilemma? In what sense is the outcome not Pareto-optimal? How could both prisoners be made better off?

R6. Explain the analogy between the Prisoner's Dilemma and the breakdown of cartels.

R7. Why might the firms in an industry want to be regulated?

R8. What determines the number of firms in a contestable market?

R9. Explain carefully how output is determined in a Cournot oligopoly.

R10. Explain carefully how price is determined in a Bertrand oligopoly.

R11. What disturbing feature do the Bertrand and Cournot models have in common?

R12. Explain carefully how price and output are determined under monopolistic competition.

R13. Explain why, in a long-run monopolistically competitive equilibrium, average cost is never minimized.

Numerical Exercises

N1. Suppose that a monopoly steel producer produces steel at zero marginal costs and sells to a monopoly automaker at a price P_{steel}. The automaker has no costs other than the cost of steel, which is converted into cars at the rate of one ton of steel to one car. There is no other way to produce a car than to use a ton of steel. The demand for cars is given by $Q_{cars} = 100 - P_{cars}$.

- **a.** For a given price of steel, what quantity of cars will the automaker produce in order to maximize profits? (*Hint:* The function $-Q_2 + kQ$, with k constant, is maximized at $Q = k/2$.)
- **b.** What is the equation for the automaker's demand curve for steel?
- **c.** How much steel is produced? At what price? How many cars are produced? At what price?
- **d.** If the steel producer acquires ownership of the automaker, how many cars are produced? At what price?

N2. Suppose that Microsoft is the only producer of operating systems and Netscape is the only producer of Web browsers. Suppose also that nobody wants an operating system without a Web browser and nobody wants a Web browser without an operating system. Suppose that both firms produce at zero marginal cost and that the demand for a package consisting of an operating system and a browser is given by $Q = 100 - P$.

- **a.** Suppose that Microsoft and Netscape take each others' prices as given. What is the price of an operating system? What is the price of a browser?
- **b.** Suppose instead that Microsoft first announces a price for its operating system; then Netscape takes this price as given and sets a price for its browser. Now what is the price of an operating system? What is the price of a browser?

c. Suppose that Microsoft merges with Netscape. Now what is the price for a package consisting of an operating system and a browser?

d. Suppose instead that Microsoft sells consumers a package consisting of a operating system and a Netscape browser and pays Netscape a royalty for each package that it sells. What royalty does Netscape charge and what price do consumers pay for the package?

N3. Dr. Miles is a monopolist who sells a type of patent medicine through competitive retailers. The demand curve for this patent medicine is given by $Q = 100 - 2P$, where P is the price and Q is the number of bottles sold.

a. If Dr. Miles has zero marginal cost, how many bottles of medicine will she sell and at what price? Calculate the consumers' surplus. Calculate Dr. Miles's producer's surplus.

b. Now suppose that retailers are able to provide their customers with valuable services by explaining how the medicine is to be used, what ailments it is effective against, and so on. By incurring a cost of C in time and effort per bottle sold, the retailer can provide services that consumers value at V per bottle sold, where V is given by $V = 5C - C^2$. What is the socially optimal amount of service per bottle for retailers to offer? What is the cost of this service?

c. Now suppose that retailers who offer services do not sell any additional medicine, because customers accept the services and then shop elsewhere, buying from a cut-rate supplier who offers no services. To combat this, Dr. Miles institutes a fair trade agreement under which she will sell at a wholesale price of P_0 but retailers must charge a retail price of P_1. Retailers have no other costs. Explain why retailers will incur costs of service equal to $C = P_1 - P_0$. What is the socially optimal value for C?

d. Taking C as given, write the equation of the new demand curve that retailers face after Dr. Miles institutes fair trade. Write the equation of the new demand curve Dr. Miles faces. In view of her wanting to face the highest possible demand curve, what value will Dr. Miles choose for C?

e. Using your answers to part (d), calculate the new price P_0 that Dr. Miles will charge, the new quantity sold, the new consumers' surplus, and the new producer's surplus.

N4. (This is a challenging problem which requires some calculus.) Only one road goes from Hereville to Thereville, and along that road you must cross two toll bridges. The number of travelers from Hereville to Thereville is given by $Q = 100 - P$ where P is the price of travel; that is, P is the sum of the two toll bridges.

a. If one monopolist owns both bridges, how much does he charge to cross?

b. If each bridge is owned by a separate monopolist, how much does each one charge?

c. Is it better for the consumer if the bridges are owned by a single monopolist or by competing monopolists?

d. Would you rather buy your computer operating system and your word processing software from a single monopolist or from competing monopolists? Why?

Problem Set

1. Consider a competitive industry where the demand and supply curves are straight lines of equal absolute value and the supply curve goes through the origin. If all of the firms in the industry merge into one, the new firm will be able to produce at zero marginal cost. On efficiency grounds, should the merger be allowed?

2. Suppose that a monopoly supplier selling in two distinct markets wants to price discriminate. How might the monopolist benefit from a vertical merger?

3. Candy makers sometimes print retail prices directly on the wrappers. Is this a form of resale price maintenance? If so, what are its benefits? If not, what is the reason for the practice?

4. Suppose that a monopoly firm introduces a policy of resale price maintenance. Under the "special services" theory of resale price maintenance, would the firm's output increase or decrease? Conversely, suppose that the purpose of the resale price maintenance is to enforce a cartel among the dealers. Now would the firm's output increase or decrease?

5. Many firms employ salespeople who are assigned exclusive territories. No salesperson may enter another's territory and attempt to sell the manufacturer's product there. Construct a theory to explain why firms adopt this practice. Does your theory suggest what kinds of products will be sold in this way and what kinds will not be?

6. Suppose that bicycle dealers serve their customers by providing fancy showrooms and knowledgeable salesforces to answer questions, but that only a small number of customers value these services. Show that in this case, resale price maintenance can cause an increase in bicycle sales but a *decrease* in social welfare.[30]

7. Suppose that airplane flights are provided at a constant marginal cost P_C. (That is, the marginal cost curve in the airline industry is flat at the price P_C.) If there were a single monopoly airline, it would sell tickets at the higher price P_M. Suppose that the government requires all airlines to charge the price P_M and forbids new entry into the airline industry.

- **a.** Show the consumers' surplus, the producers' surplus, and the deadweight loss.
- **b.** Now suppose the airlines discover that they can make themselves more attractive to customers by offering costly "extras" ranging from in-flight movies to the scheduling of frequent flights that better accommodate travelers' schedules. By how much does the marginal cost curve rise and why?
- **c.** In part (b), what happens to the demand for airline flights? Recalculate the consumers' and producers' surpluses.
- **d.** In part (c), is it possible that the net social gain could be greater than it is under competition? (*Hint:* Which additional services would be offered under competition and which would not?)

8. **True or False:** Resale price maintenance can be good for consumers because it means there will be more dealer services. Thus, if the marginal value of dealer services decreases rapidly, then the benefits of resale price maintenance are reduced.

[30] This is a hard problem. It is based on an analysis by W. S. Comanor in "Vertical Price Fixing, Vertical Market Restrictions, and the New Antitrust Policy," *Harvard Law Review* 98 (1985): 984–1002.

9. Offer some alternative theories to explain why manufacturers want fair trade. How might you go about testing your theories vis-á-vis the one outlined in the text? Do they have different implications about what sorts of products might be sold under these conditions, or about what industry structures are most conducive to fair trade?

10. Can you think of a reason why some manufacturers might want to set a *maximum* retail price for their products, and forbid sellers from charging more than the preset maximum?

11. The firms that sell personal computers have never banded together to form a cartel. **True or False:** We may infer from this that at least one firm would fail to benefit from a successful cartel.

12. In many industries workers are required to belong to a union and to pay union dues, even if they would prefer not to. **True or False:** Workers would be better off if each one could choose whether to belong to the union.

13. **True or False:** When all firms in an industry charge the same price, this is evidence of collusion.

14. In the example of Exhibit 11.1, suppose that the firms merge, but the market is contestable. What quantity does the merged firm produce, and at what price? Do any new firms actually enter?

15. Suppose that there are exactly N identical firms in an industry, all with flat marginal cost curves. Industry demand is linear. How much does each firm produce, compared with the competitive quantity, under the Cournot assumption that each takes its rivals' outputs as given? How much does the industry produce? What happens to industry output as N gets large? (*Hint:* Follow carefully the argument that is given in the text for the case $N = 2$.)

16. Consider an industry where there are two firms having identical flat marginal cost curves. Price and output in the industry are determined as follows: First Firm 1 announces how much it will produce, then Firm 2 decides how much to produce, then the industry's output is sold at a price read off the industry demand curve.

- **a.** Is the industry output greater or less than it would be under Cournot behavior?
- **b.** Which firm is better off: Firm 1 or Firm 2?

17. Suppose there are three ice cream vendors on the beach depicted in Exhibit 11.11. How will they locate themselves in equilibrium?

18. Suppose there are four ice cream vendors on the beach depicted in Exhibit 11.11. How will they locate themselves in equilibrium? What can you say if there are five vendors? What if there are more than five?

19.

- **a.** Suppose that two ice cream vendors are located on a circular beach that goes all the way around a lake. How will they locate themselves in equilibrium?
- **b.** Suppose instead that there are three ice cream vendors on the same circular beach. How will they locate themselves in equilibrium?