Chapter IX

Market Structure and Antitrust

This chapter introduces industrial organization economics — the study of firm behavior under various market conditions ranging from perfect competition to monopoly. Earlier chapters have provided the foundation for the material covered here. Chapter II showed how the basic supply-and-demand model can provide valuable insights regarding market reactions to changes in legal rules and regulations. Chapter I introduced standard assumptions regarding individual utility maximization and firm profit maximization. In particular, the underlying assumptions of consumer choice were examined. This chapter examines in detail some of the factors that determine an individual firm's response to changes in market conditions as well as changes in legal rules and regulations.

Section A examines firm production decisions, with an emphasis on the underlying relationships that determine costs of production. Section B examines how firms behave under various competitive market conditions, first considering the benchmark cases of perfect competition and monopoly, and later more realistic models of competition that involve product differentiation and oligopolistic interdependence.

A. Costs & Production

Producers of any good or service are faced with the constant task of adjusting their output in response to changes in market conditions. The effectiveness of a firm in responding to changes in market demand is in part determined by its ability to alter the quantity and mix of inputs used in its production process. In analyzing the ability of firms to alter their input use, economists distinguish between the short-run and the long-run. We discuss each of the scenarios below.

1. Short-Run Production

In the **short-run**, at least one input used in the production process cannot be altered. Inputs that can be altered within the short-run are called **variable resources**. Inputs that cannot be altered within the short-run are called **fixed resources**. Thus, in the short-run, firm output can be adjusted by changing the variable resources used, but some of the necessary resources are fixed and cannot be altered in response to changes in the desired output level. Traditionally, this distinction has led to a separation of the costs of production into two components: fixed and variable costs. **Fixed costs** are those that do not vary with

the amount of output produced and must be paid even if output is zero. **Variable costs** are those that change based upon the amount of output produced and are equal to zero if output is zero.

A factory — and the costs associated with owning or renting a factory — is a traditional example of a fixed resource or fixed cost. In the short-run, a firm cannot build, expand, sell, or purchase a factory to meet changes in market demand. Furthermore, a firm cannot stop paying rent or loan payments in the short-run without breaching contracts. Fixed costs also include inputs purchased on a long-term supply contracts.

Suppose GM uses four factories to manufacture one million Chevrolet Volts, an electric car. In the first year, GM sells one million Volts. In January of the second year, Tesla drops the price of its electric sedan by 50 percent and the demand for the Volt drops by 90 percent. In the short run, GM must incur the costs of owning all four factories even though they are not being used. Suppose instead that fuel prices quadruple causing demand for the Volt to increase to four million cars. GM instantly begins using all four of the factories, but it still needs many more factories to meet the increased demand. However, in the short-run, GM cannot expand its existing factories, or build new factories quickly enough to meet the demand.

Firms, however, can adjust their use of workers and materials in the short run. A factory worker with a terminable at-will employment contract, for example, is a traditional example of a variable resource or variable cost. When Tesla decreases the price of its sedan by 50 percent and demand for the Volt decreases, GM immediately terminates enough employment contracts so that it does not pay unneeded employees. Once the fuel prices quadruple and the demand for the Volt increases, GM immediately hires enough factory workers for all four factories to operate at capacity.

a. The Law of Diminishing Marginal Returns

Consider the hypothetical home construction company Sam's Shacks, owned and operated by Sam. In its production of new homes, Sam's Shacks uses several different fixed resources. Examples of Sam's fixed resources include heavy equipment (trucks and tractors), an office, and tools. For Sam, the most important variable input — and focus of this example — is labor, though other variable resources include lumber, wire, and nails. In the short-run, as the market demand for new homes changes, Sam is able to alter his firm's output — as in the GM example above — by increasing or decreasing the number of workers he employs. The most important decision that Sam must make is how many workers he should hire with his given level of fixed inputs.

When the construction season starts, Sam approaches the hiring of workers conservatively — he always starts by hiring one worker. As demand for new homes picks up, Sam realizes that this one worker cannot complete all of the necessary work. As a result, he hires a second worker. After Sam hires the second worker, he always notices something unusual about the company's total level of output. When only one worker was employed, Sam's output was 100 square feet per week. However, by hiring a second worker, the total output of the company jumps to 210 square feet per week. In other words, the addition of a second worker more than doubles the total output of his company. Yet, Sam knows this increase in production has nothing to do with the inherent abilities of the second worker — both are equally skilled and efficient. Having taken an economics course in college, Sam wisely attributes this result to the division of labor and specialization (recall Leonard E. Read's story *I, Pencil*, *supra* Chapter 2). No longer does the single worker have to perform all of the tasks necessary to build a new home. Now, one worker spends the

Table IX-1

Number of Workers	Total Product	Marginal Product
0	0	
1	100	100
2	210	110
3	330	120
4	460	130
5	570	110
6	650	80
7	690	40
8	680	−10

day doing carpentry while the other does electrical work. With each worker focused only on a single task, they become better and faster at what they do because they no longer have to take a break between each project to work on something else.
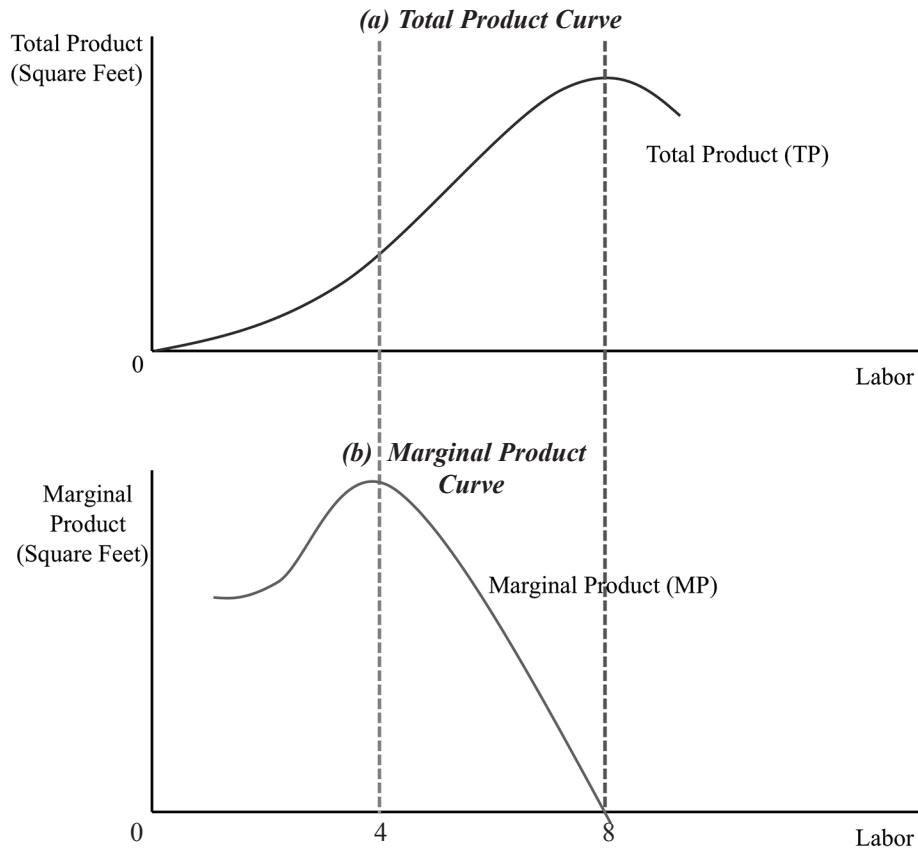
As the building season progresses, Sam continues to notice unusual output results as he hires additional workers. When Sam hires a third and fourth worker, the company's output increases to 330 and then 460 square feet per week—continuing the trend of each worker increasing total output by more than the previous worker. This is summarized in Table IX-1. The increases in output for the third and fourth workers can also be attributed to continued specialization. Now two workers do the carpentry work and two workers do the electrical work, each specializing more within each field. One carpenter measures and cuts all of the lumber while the other puts it together. Furthermore, one worker drills the holes and runs the wire through the walls while the other wires the individual outlets and fixtures.

When total output increases by greater amounts for each additional variable input added to the firm's fixed resources, the firm is said to be experiencing **increasing marginal returns**. This is what Sam witnesses with the first four workers—each of these workers increased the total number of square feet constructed by more than the previous worker. Increasing marginal returns are, as Sam noted, caused by the division of labor and specialization.

However, Sam always seems to hit a point where each additional worker hired adds less to the company's total output than when the previous worker was hired. As summarized in Table IX-1, total output is 570, 650, 690, and 680 for the fifth through eighth worker, respectively. Once again, all the workers are equally skilled and efficient. Sam is very confused by this. It seems as if he hits a point where the division of labor and specialization just does not add as much to the company's output.

The phenomenon that Sam witnesses as he hires these additional workers is known as the law of diminishing marginal returns. The **law of diminishing marginal returns** states that, as additional quantities of a variable resource are added to a given amount of fixed resources, a point is eventually reached where each additional variable resource will add smaller and smaller amounts to total output. For workers five through eight, Sam experiences the law of diminishing marginal returns. This occurs because after the fourth worker, the jobsite becomes too crowded and limits the ability of each worker to complete as much work as she could have if there were fewer workers. Aside from generally bumping into each other, once there are more workers, the carpenters will have to take turns using saws and other equipment instead of just being able to use tools when they are needed.

**Figure IX-1. The Law of Diminishing Marginal Returns Product Curves
for Sam's Shack**

(a) Total Product Curve

Total Product
(Square Feet)

Total Product (TP)

0   Labor

*(b) Marginal Product
Curve*

Marginal
Product
(Square Feet)

Marginal Product (MP)

0   4   8   Labor

For the workers doing electrical work, adding workers means that one will have to wait
for another to finish drilling before he can use the drill to run wires through a wall.

The law of diminishing marginal returns is shown graphically in Figure IX-1. **Total
product** (TP) is the total output of goods or services produced by a firm. For Sam's Shacks,
TP is measured in square feet of construction (column two in Table IX-1). **Marginal
product** (MP) is the change in TP that results from a one unit change in variable re-
sources—in this case each additional worker hired (column three in Table IX-1). Figure
IX-1 displays these two curves for Sam's Shacks. The MP curve is graphed in panel (b).
Each of the first four workers that Sam hires adds more to the TP than the worker hired
previously. Thus, Sam's MP of labor increases for the first four units hired. This is reflected
in panel (b) with an MP curve that increases over the first four units of labor. For each
of the next four workers that Sam hires, MP becomes progressively smaller because of
the law of diminishing marginal returns—each additional worker hired adds a smaller—
and eventually negative—amount to the TP. As a result, the MP curve declines over this
range. When Sam hires the eighth worker, TP actually decreases and the MP curve drops
below zero, indicating that MP is negative.

The TP curve is graphed in panel (a). Note that when MP is increasing, TP grows at
an increasing rate. However, as MP begins to decline because of the law of diminishing

marginal returns, the TP curve begins to grow at a decreasing rate. Finally, as MP becomes negative with the eighth worker, the TP curve begins to decline.

b. Short-Run Costs

As will be discussed later in this chapter, variable costs of production affect production decisions. Variable costs are determined by the price of the variable resource *and its marginal product* (MP). Let us return to the example of Sam's Shacks in order to examine the relationship between costs and marginal productivity. For Sam, MC first decreases and then increases. This pattern is a function of the law of diminishing marginal returns. Remember that for each of the first four workers hired, MP increases. Thus, fewer and fewer units of labor are required to produce an additional unit of output. This greater productivity causes MC to decrease as these four workers are employed. Furthermore, remember that marginal product decreases as Sam hires workers five through eight. In this range, more and more workers are required to produce an additional unit of output. This drop in marginal productivity causes the MC of production to start increasing. In other words, when diminishing marginal returns set in, the MC of production begins to increase. The relationship between marginal productivity and MC for Sam's Shacks is included in Table IX-2 and can be seen in Figure IX-2(b). Notice that the MC curve is U-shaped. MC decreases because of increasing marginal returns and then increases because of diminishing marginal returns.

Short-run cost curves describe the relationship between costs and output when there is at least one fixed input. Seven different measures of short-run costs are of interest.

1.   **Marginal cost** (MC) is the cost of producing each additional unit of output.

2.   **Fixed cost** (FC) is the price paid for inputs that cannot be varied in the short-run.

3.   **Variable cost** (VC) is equal to the price of the variable input multiplied by the quantity of the variable input used.

4.   **Total cost** (TC) is equal to the sum of fixed and variable costs: $TC = FC + VC$.

5.   **Average fixed cost** (AFC) is fixed cost divided by the firm's total output (Q): $AFC = FC \div Q$.

6.   **Average variable cost** (AVC) is equal to variable costs divided by the firm's total output: $AVC = VC \div Q$.

7.   **Average total cost** (ATC) is the total cost of production divided by the firm's total output: $ATC = TC \div Q$.

Table IX-2 provides example values for all of these cost curves for Sam's Shacks where Q is the total construction conducted in square feet per week. It is assumed that Sam's fixed costs are $50 and that he pays each worker he hires $20.

It is important to recognize that the cost curves reflect both the explicit and implicit costs of production. Thus, fixed costs, such as the cost of a building or factory (Sam's heavy equipment and tools in this example), include the rent or return that the building could earn in its next best alternative use. When all costs are covered, the firm is receiving a normal return on the fixed assets, or zero economic profits. Of course, the goal of most firms is to achieve more than a normal return. An above normal return is an economic profit.

Figure IX-2(a) shows the FC curve for Sam's Shacks based on the values in Table IX-2. Note that FC is represented by a horizontal line at the $50 level, demonstrating that FC does not vary with changes in output. VC is also graphed in Figure IX-2(a). For Sam, this is the cost of the workers, who Sam pays $20 each, per quantity produced. This curve

Figure IX-2. Total and Marginal Cost Curves

(a) Total Cost Curves

(b) Marginal Cost
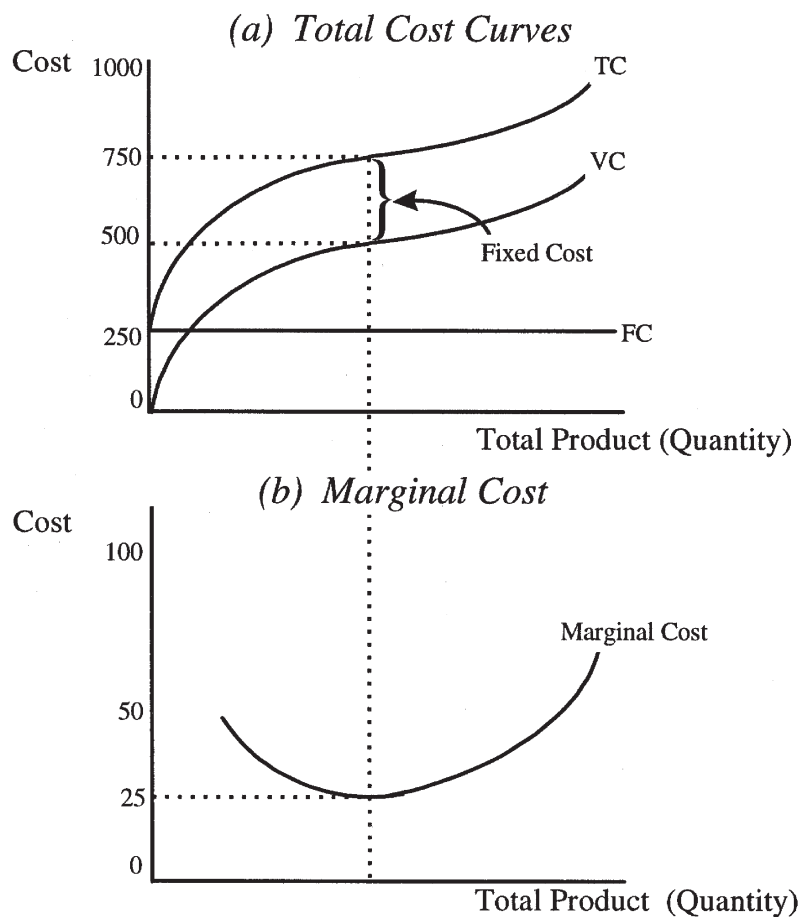
indicates that VC equals zero when output is zero. VC initially increases at a decreasing rate as the firm realizes increasing marginal returns, and then begins to increase at an increasing rate at the point where diminishing marginal returns sets in. The TC curve for Sam's is also shown in Figure IX-2(a). The TC curve is simply the vertical summation of the fixed and variable cost curves (or FC + VC in Table IX-2). Thus, the TC curve is the VC curve shifted upwards by the amount of FC.

Average cost data for Sam's Shacks graphed in Figure IX-3 also reflect the law of diminishing marginal returns. The shape of the AVC curve is determined by the shape

Table IX-2. Cost Measures for Sam's Shacks

Q	FC	VC	TC	MC	AFC	AVC	ATC
0	50		50				
100	50	20	70	20	0.5	0.2	0.7
200	50	37	87	17	0.25	0.185	0.435
300	50	52	102	15	0.167	0.173	0.34
400	50	66	116	14	0.125	0.165	0.29
500	50	81	131	15	0.1	0.162	0.262
600	50	98	148	17	0.083	0.163	0.247
700	50	118	168	20	0.071	0.169	0.24
800	50	142	192	24	0.063	0.178	0.24

of the MC curve. When MC is below AVC, AVC declines. When MC is above AVC, AVC begins to rise. The MC curve intersects the AVC curve at its minimum point—indicated by point A. This basic relationship is known as the **average-marginal rule**. The MC curve has the same relationship with the ATC curve as it did with the AVC curve—the MC curve will intersect the ATC curve at its minimum point—indicated by point B. AFC declines over the entire range of output. This reflects the fact that FC is spread over a larger total output as production expands.

Gregory Mankiw compares the average-marginal rule and why the marginal cost curve intersects the AVC and AFC curves at the minimum to a student's GPA. N. Gregory Mankiw, Principles of Microeconomics 279 (3d ed. 2004). He likens MC to the student's grade in one class and AVC or ATC to the overall GPA. If the grade in the class is less than the student's GPA, the GPA will decline; if the grade is higher than the GPA, the GPA will increase. So it is with MC and AVC or ATC. If the MC is less than AVC, AVC is falling; if it is higher, then AVC is rising. This means the intersection will be the minimum. For Sam's Shacks, MC is less than AVC when Sam is producing 0 to 500

Figure IX-3. Average & Marginal Cost Curves
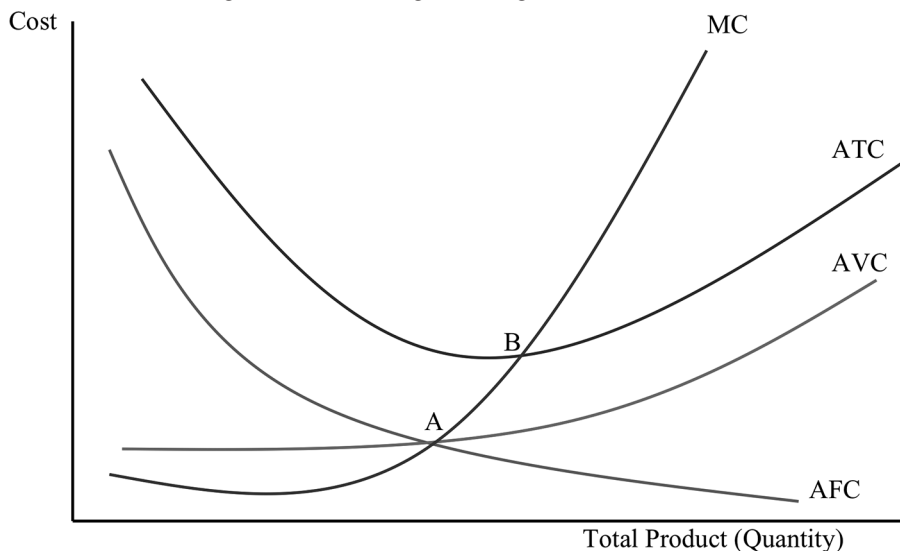
square feet, then MC is greater than AVC. Thus, somewhere between 500 and 600 square feet, MC intersects AVC.

c. Changes in Costs of Production

Changes in the cost of producing any good or service cause the cost curves to shift by an amount reflecting the change in cost. In general, there are two important sources that change the costs of production: (1) changes in the cost of inputs and (2) changes in production technology. Taxes on output also affect costs as do government regulations that impact technology or input costs. When variable costs of production increase, the MC curve shifts to the left; when these costs of production decrease, the MC curve shifts to the right. When either fixed or variable costs of production increase, the ATC curve shifts up and when either of these costs of production decrease, the ATC curve shifts down.

Consider Figure IX-4, which shows two sets of MC and ATC curves for a firm — MC_1 and ATC_1, and MC_2 and ATC_2. If the cost of an input used in the production of this firm's goods or services increases, the MC curve shifts to the left. This can be seen in Figure IX-4 as a shift from MC_1 to MC_2. Furthermore, increased input costs shift the ATC curve up from ATC_1 to ATC_2. An increase in the excise tax shifts the MC and ATC curves in the same way. Changes in regulation that increase the costs of compliance similarly cause the costs of production to increase. Technological improvements cause the MC curve to shift to the right, indicating a decrease in the costs of production — movement from MC_2 to MC_1 in Figure IX-4. Likewise, the ATC curve shifts down from ATC_2 to ATC_1.

d. Economies of Scale

The short run ATC curve is U-shaped, reflecting two competing forces. First, ATC declines as output expands because the reductions in ATC from spreading fixed costs over an increasingly large number of units of output are larger than the increase in ATC from

Figure IX-4. The Effect of an Increase in Resource Prices on Costs
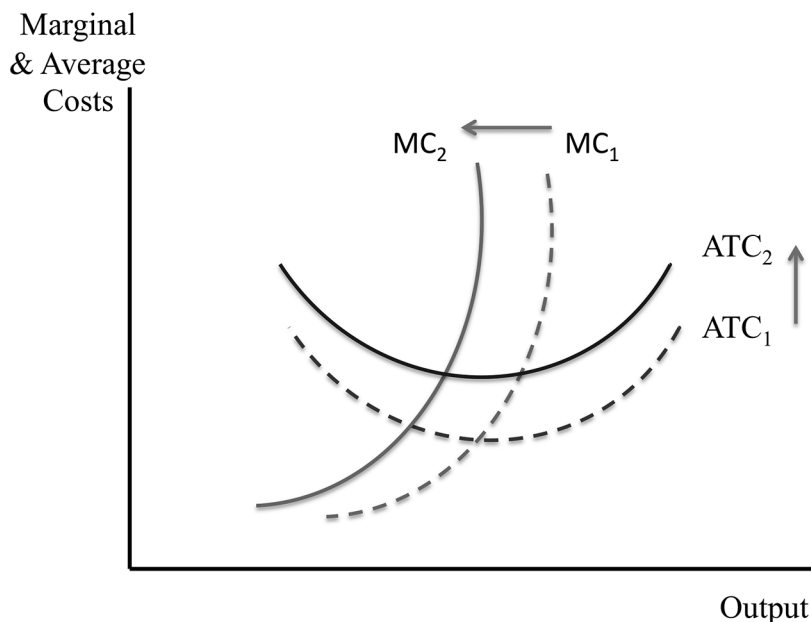
diminishing marginal returns. The firm is said to enjoy **economies of scale** over this range of declining ATC. At some point however, the increase in variable costs due to diminishing marginal returns swamp the reductions in fixed costs, causing ATC to increase. Firms suffer **diseconomies of scale** over the range of increasing ATC. **Constant returns to scale** occur when long-run average cost does not change as output expands over some range. For example, there may be a long range of output over which ATC is at its minimum.

Economies of scope exist when the joint production of multiple outputs is more cost efficient than producing each output separately. For example, the provision of both savings and checking accounts by the same bank is more cost efficient than having bank A provide checking and bank B provide savings. Imagine the amount of duplication that is avoided by having one bank provide both savings and checking accounts. When such economies arise, one firm would produce both outputs rather than two firms each producing only one. Moreover, it would be a better use of resources for a single firm to produce both outputs because to do so would minimize the total opportunity cost of the resources used in production.

2. Long-Run Production

In the **long-run**, all inputs used in the production process are variable. Thus, in the long-run, there are no fixed resources—all inputs can be varied in response to a change in the desired level of output. Unlike in the short-run, firms can build, expand, or shut down factories. From the GM example above, in the long run GM is able to sell unused factories or build new ones as demand changes in the long-run—making even the number of factories variable. Thus, one of the most important long-run decisions that the firm must make is the selection of the optimal plant size—the plant size that will minimize the long-run costs of producing the profit maximizing output.

Assume that a particular manufacturer has five choices regarding plant size, and that the ATC curve for each of these plants is depicted in Figure IX-5. The plant size that is best for a particular manufacturer depends on the level of output to be produced. For example, if output less than Q_1 needs to be produced, then the plant size that corresponds to ATC_1 is the optimal choice. For output choices between Q_1 and Q_2, ATC_2 is optimal. Likewise, for output between Q_2 and Q_3, ATC_3 is optimal. Thus, the optimal plant size for any given output can be found along the continuous portion of the five ATC curves in Figure IX-5. The continuous portions of the five ATC curves in Figure IX-5 are the basis for a new curve—the long-run average cost curve.

Long-run average cost (LRAC) is the lowest per unit cost of producing any level of output when the firm can choose from all possible plant sizes. For the firm in Figure IX-5, the continuous portion of the ATC curves represent the LRAC curve where only five possible plant sizes are available. Figure IX-6 shows a smooth LRAC curve. This curve represents the lowest per unit cost of producing any level of output when a firm can choose from an infinite number of possible plant sizes. Notice that this smooth LRAC curve is tangent to an infinite number of short-run ATC curves. As the LRAC curve slopes downward, it is tangent to the short-run ATC curves before the point of minimum cost for each plant size; as the LRAC curve slopes upward, it is tangent to the short-run ATC curves past the point of minimum cost. At the lowest point on the LRAC curve, the LRAC curve is tangent to the minimum point on a short-run ATC curve. Figure IX-6 also depicts the long-run marginal cost curve. **Long-run marginal cost** (LRMC) is the change in *total cost* that results from a one unit increase in production. This reflects the fact that, in the long-run, all resources are variable. Thus, everything from labor to plant size is incorporated

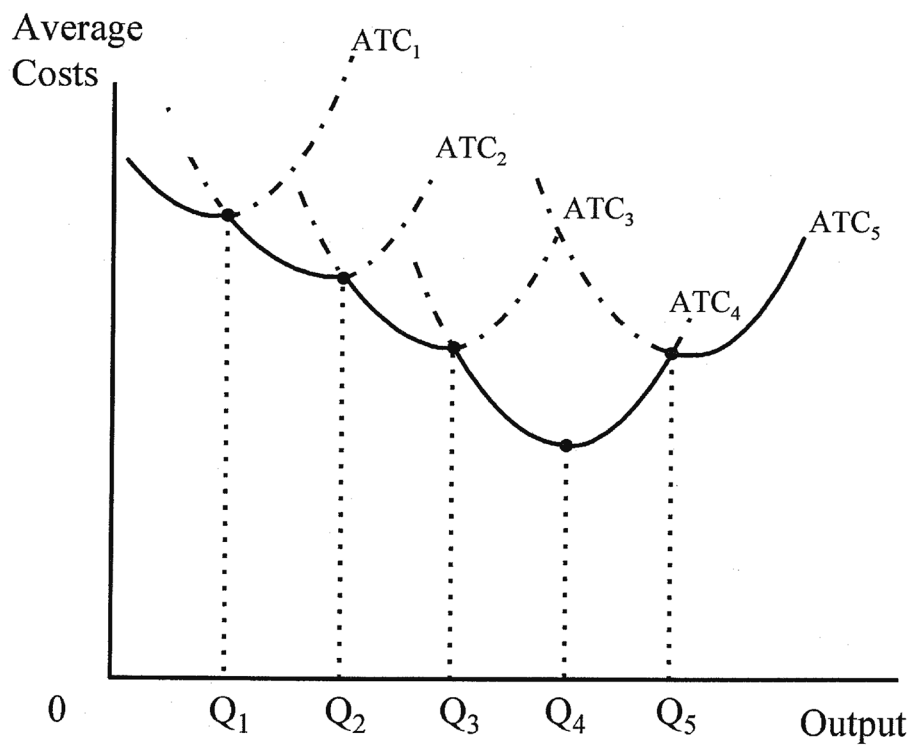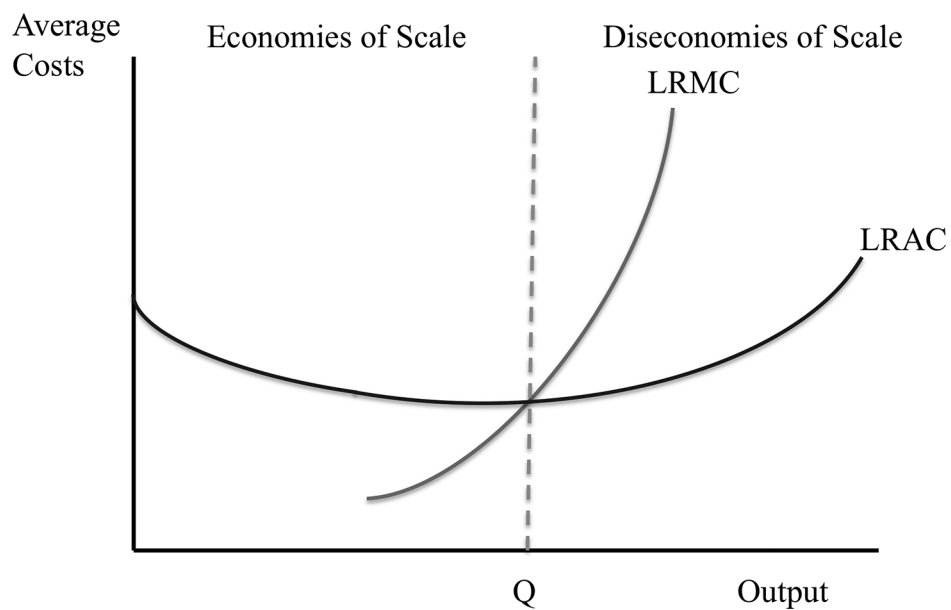**Figure IX-5. Adjusting Plant Size to Minimize Costs**

Figure IX-6. Adjusting Plant Size to Minimize Costs

into the LRMC number. As with short-run average costs curves, LRMC intersects long-run average curves at their minimum point. Thus, as long as LRMC is below LRAC, LRAC will decrease. When LRMC is greater than LRAC, LRAC will increase.

Like the short run ATC curve, the LRAC curve is also U-shaped over the full range of possible levels of output. First, the firm enjoys economies of scale as output increases because the opportunities for specialization and division of labor increase. Second, firms often learn how to reduce the costs of production after gaining some experience in actually producing the product. In other words, as the firm's output grows, managers and employees gain more and more experience with the production process and hence learn more efficient production techniques. Third, as output expands, firms can begin to take advantage of mass production technology. The combination of these factors helps to explain why LRAC decreases as output approaches Q from the left in Figure IX-6.

The Sam's Shack example elaborates on economies of scale. In the long run, when all factors of production are variable, Sam can vary his output by purchasing more heavy equipment and more tools. As he increases equipment, he is able to keep more workers busy and change the point where he experiences diminishing marginal returns. He is even able to buy enough heavy equipment and tools to have his company building many homes at the same time. Sam enjoys economies of scale because of all three potential sources listed in the previous paragraph. First, as he expands the company so that it builds many homes at once, his workers can very narrowly specialize and still keep busy by moving from one home to the next. For example, he can have workers that dig one foundation each day and then move to the next home, workers that only put shingles on, and workers that only pour concrete sidewalks around the home. Because these projects cannot be done simultaneously with other projects (workers cannot put shingles on before the foundation is poured and the home is built), workers could not have specialized on such narrow areas without Sam's Shacks dramatically increasing its size. Second, as the Sam's Shacks workers build many homes a year, they will learn new tricks that speed up the process. Third, once Sam's Shacks is large enough, it becomes worthwhile for Sam to purchase expensive machinery that prefabricates the walls to a home and then delivers them to the construction site—something that could not have been profitable without substantial expansion. These three things allow Sam's Shacks to benefit from economies of scale.

Diseconomies of scale begin to set in as the firm's output grows larger. Diseconomies of scale arise primarily as a result of the fact that, at some point, the firm grows too large for its managers to control. As firm output expands, more employees and managers are hired, which increases the firm's bureaucracy. As this bureaucracy grows, communication becomes difficult and decision making can grind to a halt. Eventually, the diseconomies outweigh the economies of scale. Thus, in Figure IX-6, LRAC increases as output expands to the right of Q.

The output level that minimizes LRAC is called the **minimum efficient scale**. Minimum efficient scale can have a large impact on the number of firms that compete in a particular industry. When minimum efficient scale can be achieved at low levels of output relative to industry demand, many firms will be able to enter and survive in the industry. On the other hand, if minimum efficient scale can be reached only at levels of output that constitute a large portion of the entire market demand, then only a few firms will be able to survive in the industry.

In Figure IX-6, LRAC comes to a unique minimum point at Q. This implies that there is one specific plant size—output level—that minimizes the costs of production in the

Figure IX-7. Alternative LRAC Curves
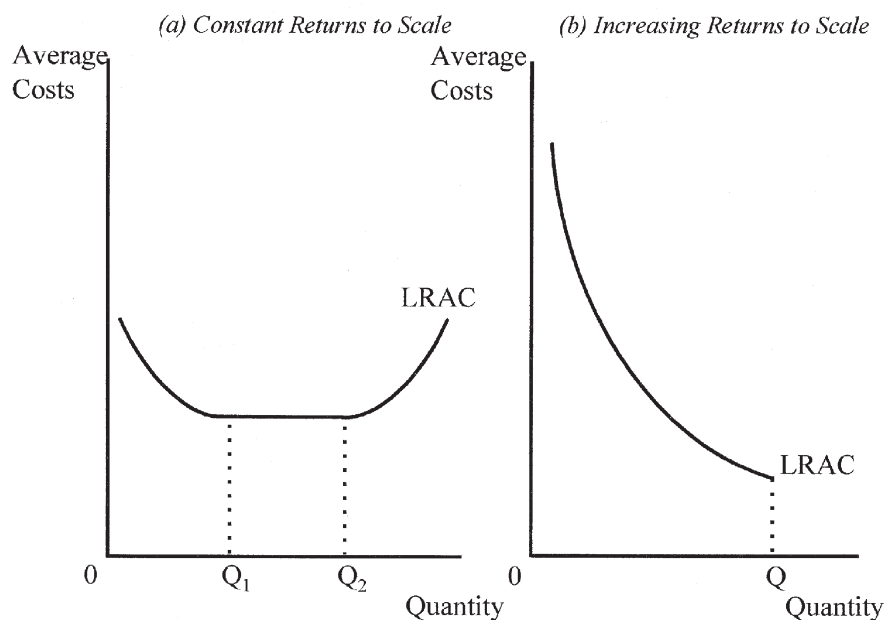
long-run. The LRAC curve does not have to decrease to a unique minimum point, however. Two other types of LRAC curves can be observed. First, Figure IX-7(a) depicts a LRAC curve that is exhibiting constant returns to scale. **Constant returns to scale** occur when long-run average cost does not change as output expands over some range. In Figure IX-7(a), economies of scale exist to the left of Q_1, constant returns to scale exist between Q_1 and Q_2, and diseconomies of scale exist to the right of Q_2. Constant returns to scale indicate that firms can produce at output levels between Q_1 and Q_2 without a difference in per unit costs.

Another LRAC cost curve of interest is one that exhibits increasing returns to scale, as depicted in Figure IX-7(b). **Increasing returns to scale** occur when LRAC decreases continually as output expands, indicating that economies of scale persist over all levels of output. This LRAC curve would indicate that the larger the firm, the better. In fact, such a curve suggests that a single firm would supply the entire industry output at the lowest average cost in the long-run. Public utilities such as electricity, natural gas, and water are often assumed to exhibit LRAC curves with increasing returns to scale.

B. Market Structures

1. Perfect Competition

Competition occupies a central role in the economic analysis of firm activity. In general, the effects of competition are seen as highly desirous and beneficial to overall social welfare. As the Supreme Court has said, "competition will produce not only lower prices, but also better goods and services. 'The heart of our national economic policy long has been faith

in the value of competition."... [C]ompetition is the best method of allocating resources in a free market ... [A]ll elements of a bargain—quality, service, safety, and durability—and not just the immediate cost, are favorably affected by the free opportunity to select among alternative offers."[1] This section examines the structural characteristics of competitive markets. These structural characteristics are enumerated in the model of perfect competition. A perfectly competitive market is a hypothetical model used to demonstrate the benefits of intense competition. In a perfectly competitive market both consumers and sellers complete all transactions that increase their welfare. This model is often used as a benchmark for analyzing real world markets.

To demonstrate this model, imagine a market where there are 1,000 sellers who are all selling exactly the same product. Each of these sellers is fully accessible to you. They are selling a product that you and everyone you know wants to buy. Despite this product's popularity, there are never any shortages or wait times to get your hands on it. The resources necessary to manufacture this product are readily available and cheap. And although some of the sellers may close down there is always another seller ready and able to take its place. Finally, you can readily find out everything you need to know about each seller, including its price, availability, location, and production capability. What would competition be like in this scenario?

First, no firm has an incentive to invest in improving technology because such innovations can be immediately imitated. Second, no firm has an incentive to advertise because all products and prices are alike. Third, no individual firm will have meaningful control over the price it can charge. Intuitively, this makes sense. The firm's output is so small relative to the market that its individual transactions do not affect the market price. Fourth, because all products are identical and market participants have complete information, no individual firm can raise its price without losing all of its sales. If firm A tries to raise its price, consumers already know about firm B that sells an identical product at the lower price. Because individual firms in a perfectly competitive market have no control over price, they are said to be **price takers**. Firms in a perfectly competitive market must offer their goods and services for sale at whatever price is established by the interaction of market demand and supply.

a. Short-Run Individual Firm Behavior: Profit Maximization

The perfectly competitive firm as a price taker is illustrated graphically in Figure IX-8. Market demand and supply appear in panel (b), and the demand curve for the output of a single firm is shown in panel (a). In panel (b), market demand and supply intersect at a price of $4.00 per unit. This market price is the price that the perfectly competitive firm must accept. Thus, the demand curve for a single competitive firm is a horizontal line drawn at the level of the market price—$4.00. Recall from the discussion of demand elasticity in Chapter II that a horizontal demand curve signifies totally elastic demand. If the firm were to try to raise its price above $4.00, its output would drop to zero. At the market price, the firm can sell as much as it can produce. A profit maximizing firm would never sell below the market price because it can already sell as much as it wants at the market price. Firms that face totally elastic demand have no power over price—no market power. In this situation, price is set by the market and individual firms have no ability to affect that price.

Firms in a perfectly competitive market attempt to maximize economic profit. Although a firm in a perfectly competitive industry has no control over price, it does control its level of output. Thus, the perfectly competitive firm must find the level of output that

---

1. Nat'l Soc. of Prof'l Engineers v. U. S., 435 U.S. 679, 695 (1978) (quoting Standard Oil Co. v. FTC, 340 U.S. 231, 248 (1951)).

**Figure IX-8. The Firm's Demand Curve and Market Equilibrium
in Perfect Competition**
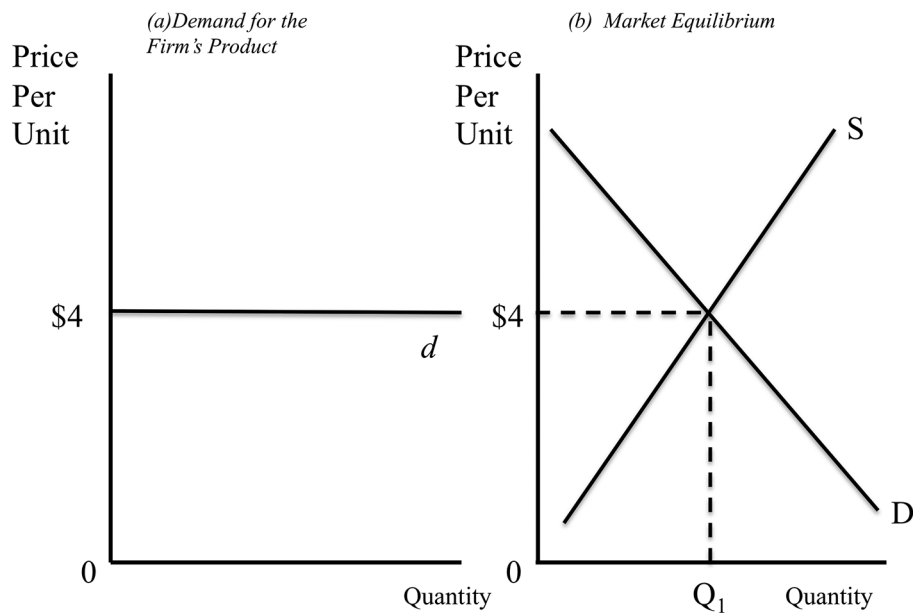
maximizes its profits. **Profit maximization** occurs at the level of output where the difference between total revenue (TR) and total costs (TC) is maximized. **Total revenue** is equal to the market price multiplied by the quantity of output sold.

A firm always maximizes its profits by producing that the point where MC equals **marginal revenue,** which is the additional revenue earned from a one unit increase in output. In a perfectly competitive market, the marginal revenue will always equal to the price because, as a price taker, the firm will always be able to sell at the market price (P). Therefore, the market price will always represent the marginal revenue gained by the firm when it sells an additional unit of output. Thus, for firms in a perfectly competitive industry, marginal revenue equals price — MR = P.

A firm should continue to produce additional units of output as long as the gain in TR exceeds the increase in TC, or stated differently, as long as MR is greater than MC. Every firm should ask itself if manufacturing and selling an additional unit of output will bring in more money in revenue than it will cost to produce. In other words, a firm will continue to produce additional units as long as the MR from producing those units exceeds the MC. This method for determining the profit maximizing level of output is illustrated in Figure IX-9(b), which shows the MR and MC curves for a firm in a perfectly competitive market. The MR curve is a horizontal line at $4.00 reflecting the fact that the firm is a price taker and will earn $4.00 in MR for each additional unit sold.

Consider the output decisions available to the profit maximizing firm faced with these MR and MC curves. At output levels from 0 to 10, MR > MC. Thus, the firm can increase profits by producing additional units. At any output level greater than 11, MC > MR and profits are increased by cutting back output. When the firm produces 11 units of output, MR = MC. At this level of output, the firm incurs an economic loss if it expands production and sacrifices profits if it reduces output — the firm can do no better than producing 11

Figure IX-9. Profit-Maximizing Output for a Purely Competitive Firm
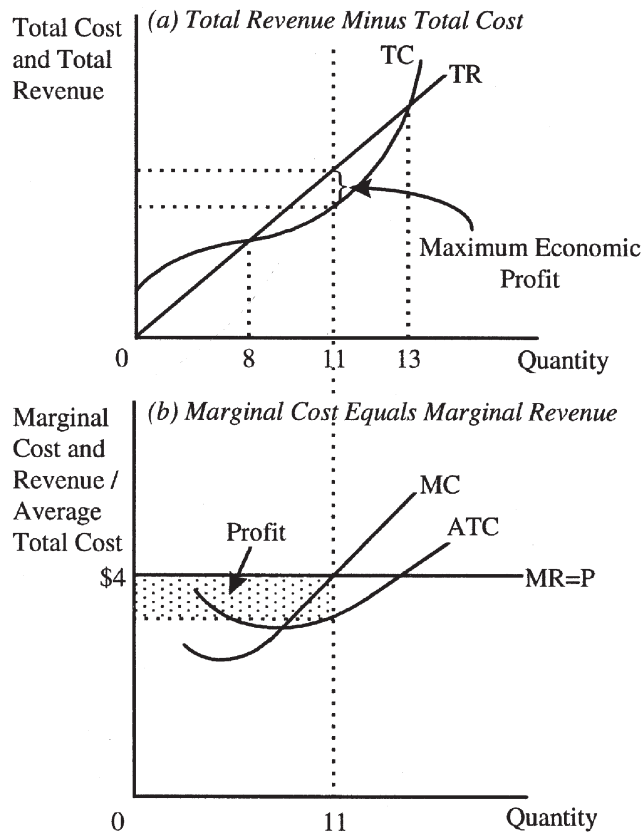
units. More generally, the profit maximizing level of output for all firms occurs where MR = MC. The amount of profit that the firm will earn at this level of production can be determined by multiplying the vertical difference between P and ATC by total output. Profit is shown by the shaded area in Figure IX-9(b).

Figure IX-9(a) can be used to illustrate the profit maximizing decision making process. As each additional unit is sold, the firm earns additional revenue equal to $4.00. Thus, the TR curve is a straight line which increases $4.00 for every additional unit of output sold. The TC curve first increases at a decreasing rate — a result of increasing marginal returns — and then increases at an increasing rate — caused by the inevitable diminishing marginal returns. Figure IX-9(a) shows that TC exceeds TR at output levels below 8 and above 13. Thus, if the firm were to produce less than 8 or more than 13 units of output, it would incur an economic loss. Figure IX-9(a) also shows that TR exceeds TC at output levels between 8 and 13. It follows then that an economic profit can be earned by the firm when it produces at any level between 8 and 13 units of output. However, economic profits are maximized when the vertical distance between the TR curve and the TC curve is maximized. This profit maximizing result occurs at an output of 11 units in Figure IX-9(a). Another way to think about this result is that firms want to maximize the difference between TR and TC because that is the point at which they earn the largest possible profit.

### b. *Loss Minimization: The Shut-Down Decision*

Whenever market price falls below a firm's ATC, the firm will incur an economic loss. This loss will be equal to the difference between ATC and price multiplied by output — (P – ATC) Q. In the short-run, the firm has two options: it can continue to operate and attempt to minimize losses; or, it can shut down, thereby reducing its variable costs to zero, and suffer losses equal to its fixed costs.

Because the firm has both fixed and variable costs in the short-run, it is possible to incur smaller losses by operating than by shutting down. When a firm shuts down, it will suffer an economic loss equal to its FC. The shut-down solution may not always be the best, however. The profit maximizing firm will make its decision on whether to operate in the face of economic losses based upon the relationship between market price and AVC. If price exceeds AVC, then the revenue generated by the firm will cover all VC and some revenue will be left over. This residual revenue is offsets some of the FC. By continuing to operate, the firm's economic loss is less than its FC.

It might be useful to consider an example. Imagine a factory that is operating at a point where its price exceeds AVC, but where its MC are greater than its MR. It is easy to think that the owner of the factory should shut it down because it is losing money. But the factory will, in the short-run, continue to incur the costs of owning the property (paying the mortgage), and it has already incurred the costs of buying machinery, training employees, and purchasing the inputs for its production. Even if the factory shuts down, those costs cannot be avoided in the short-run — the fixed costs are considered sunk. The important thing for the factory owner to consider is whether the factory is covering its variable costs, which only include the price of paying its employees and operating its machinery. The FC will be its losses if it shuts down. So long as the factory's TR is higher than its VC, then it is preferable to operate in the short-run because operation will reduce the firm's costs below its unavoidable FC.

Figure IX-10(a) demonstrates how to derive the loss minimizing result via marginal analysis. The figure shows the MC, ATC, AVC, and MR curves for a perfectly competitive

Figure IX-10. Loss Minimization: The Shutdown Decision
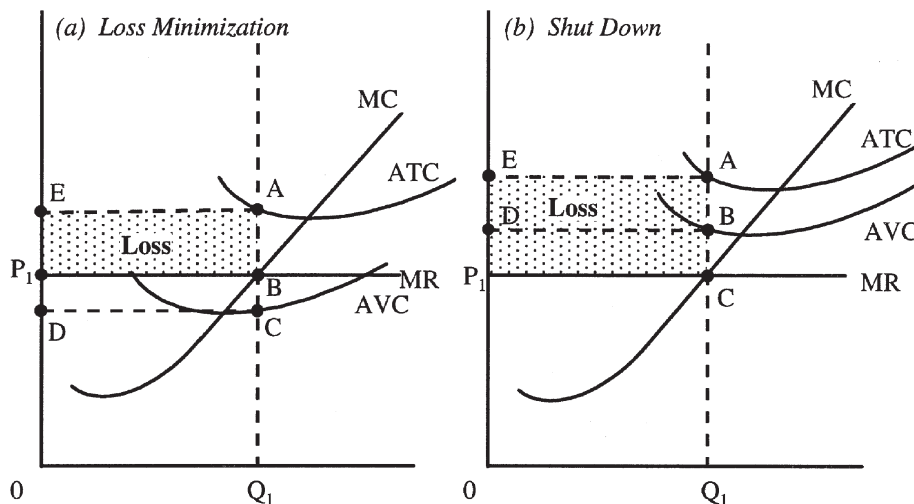
firm. The loss minimizing level of output is found by increasing production as long as MR > MC. This MR = MC rule is subject to one limitation. Namely, the MR and MC curves must intersect—be equal—at a point above the AVC curve. Recall that the vertical distance between the ATC and AVC curves is equal to AFC and that FC = AFC(Q). Thus, when price moves above AVC, revenue cuts into some portion of the AFC that lies between ATC and AVC. In Figure IX-10(a), the loss minimizing strategy is indeed met where MR = MC at Q_1 units of output. Total revenue generated at this level of output is equal to the area in the rectangle framed by (P_1BQ_10). Because the firm is able to cover all of its VC and a portion of its FC it is better off not shutting down. The economic loss is equal to (ATC – P) Q.

The case for shutting down is considered in Figure IX-10(b). Notice that at Q_1, (P_1 < AVC), (AFC > (ATC – AVC)), total revenue ($0P_1CQ_1$) earned by the firm does not cover variable costs ($0DBQ_1$), and losses (P_1EAC) are greater than fixed costs (ABDE). The rational firm will minimize losses by shutting down and incurring losses only equal to FC. Another way to consider this situation is to realize that different prices result in substantially different business decisions for a single firm. In other words, the prevailing exogenous market price can determine whether or not a factory will operate or shut down in the short-run, which is dependent on whether the factory is able to cover its VC.

c. The Short-Run Supply Curve

The perfectly competitive firm considers three factors in determining its profit maximizing level of output: (1) price (where P = MR); (2) marginal cost (MC); and (3) average variable cost (AVC). As P changes, the firm will produce the quantity where P = MC as long as P is above AVC. This relationship can be seen in Figure IX-11, which demonstrates that at different market prices, firms make different choices about their

Figure IX-11. Summary of Short-Run Decisions
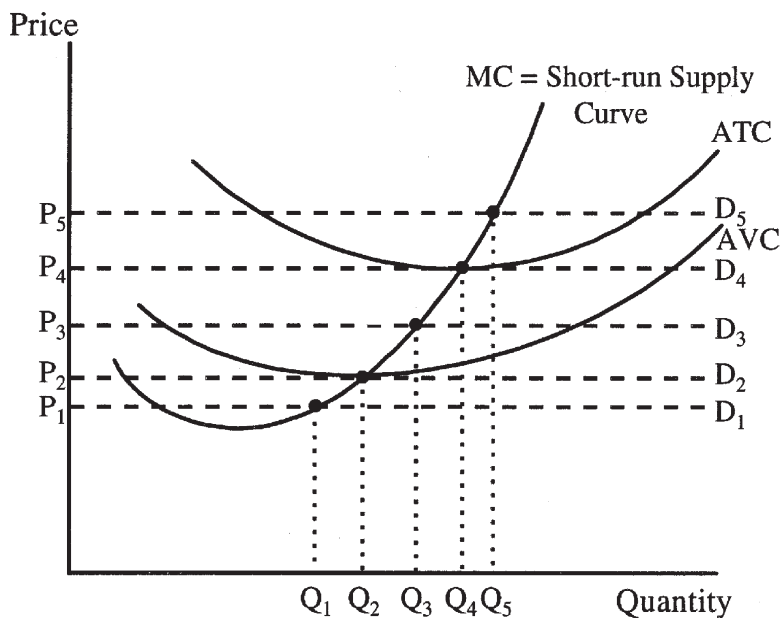
short-run operations. At P_5, P > ATC and the firm earns positive economic profits. At P_4, P = ATC of production and the firm earns a normal economic profit. At P_3, AVC < P < ATC, the firm seeks to minimize its economic loss by continuing to operate. At P_2, P = AVC, the firm is indifferent between shutting down with a loss equal to its FC or producing an output just sufficient to cover VC. Regardless of whether the firm produces or shuts down at this price level, the economic loss is equal to FC. Finally, P_1 < AVC and, thus, the firm will produce zero output in order to minimize economic losses. The above analysis leads us to an important conclusion regarding the perfectly competitive firm's MC curve in the short-run. For the perfectly competitive firm, the portion of the MC curve that lies above the AVC curve is the firm's short-run supply curve. Thus, in the short-run, firms minimize losses or maximize profits by adjusting their output subject to their MC curve.

d. Long-Run Firm Behavior: Industry Supply

Firms have substantially more options in the long-run. In the long-run, all resources can be varied—there are no fixed inputs or fixed costs—and firms are capable of making any number of structural changes to the way they do business. Firms can expand or contract their plant size as well as enter or exit an industry. In the short-run, firms are capable of earning economic profits because it takes time for new firms to construct plants and enter the market, and existing firms cannot immediately expand production beyond the capacity of current plants. In the long-run, economic profits attract new firms to the market and encourage existing firms to expand their plant and output. As firms move into industries earning positive economic profits, industry supply expands. The result of this increase in supply is to force the market price down. As the market price drops, so do economic profits. In the long-run, the entry of firms into industries earning positive economic profits eliminates those profits. This is the long-run process of industry competition.

A short-run economic loss will have a similar impact on the industry in the long-run. Firms suffering from an economic loss will observe other industries that are either making a normal return or a positive economic return. The rational response for these firms is to transfer their resources to a more profitable use. In the long-run, profit maximizing firms exit industries with short-run economic losses. As firms leave the industry, industry supply decreases. This decrease in supply causes the price for the good or service in that industry to increase. This process continues until supply contracts to the point that a normal profit is earned by the remaining firms.

The tendency for perfectly competitive firms to earn only a normal profit in the long-run is at the core of understanding long-run production decisions. As firms enter or exit a market, they impact industry supply and market price. The process of entry or exit will continue, in the long-run, until the market supply curve intersects the market demand curve at a price that is equal to the lowest point on each firm's LRAC curve. At any price that does not meet this condition, further adjustments to market supply will occur. A price above the minimum point on the LRAC curve will result in economic profits, attracting more firms or expanded output to the industry. On the other hand, a price below the minimum point on the LRAC curve would indicate that economic losses were being incurred. This would encourage additional firms to exit or cut back on production. This process makes a lot of intuitive sense. When one firm is making a lot of money producing a good or offering a service, other firms are going to want to enter that industry and try to gain some of those profits. When an industry is, instead, losing money, firms will drop out of that industry and seek another venture that will make them economic profit.

**Figure IX-12. Long-Run Equilibrium for the Firm and the Industry**
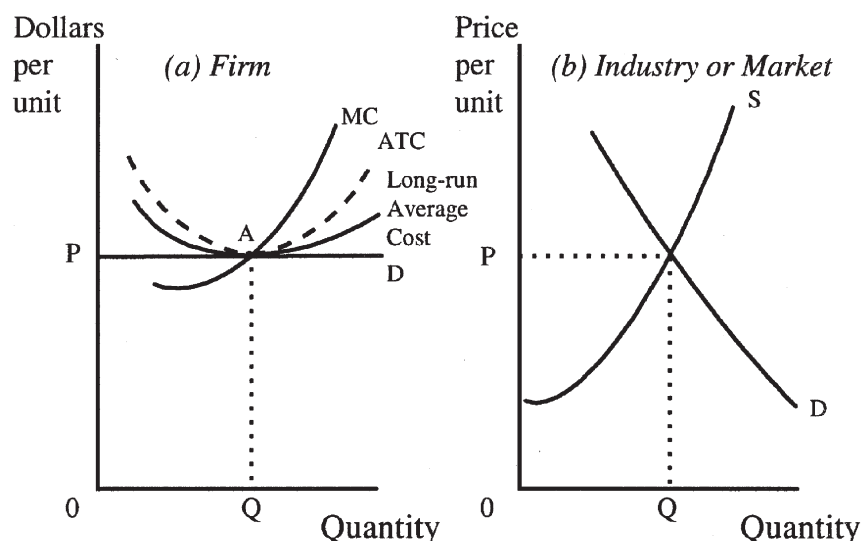
In this quest to maximize economic profits in the long-run, firms vary their scale of production until their average costs of production are minimized. Figure IX-12 demonstrates the equilibrium position at which the firm will earn a normal profit. At point A, the firm's MC, short-run average total cost (ATC), and LRAC are all equal. At this point, no firm in the industry has a reason to alter its output, and no firm outside of the industry has an incentive to enter the market. Thus, the market is in a stable equilibrium as shown in Figure IX-12(b).

Further insight into the long-run adjustment process can be gained by considering a specific example that shows short-run profits and a long-run shift in the market. For purposes of this example, assume that the costs of production do not change based on the number of firms in the industry. Figure IX-13(b) shows a perfectly competitive market in equilibrium at point A. Here, S_1 intersects D_1 and equilibrium price and quantity are P_1 and Q_1 respectively. Consider what occurs as a result of an increase in demand for the industry product in the short-run. In Figure IX-13(b) this is represented by movement to point B as demand shifts from D_1 to D_2. In the short-run, equilibrium price is P_2. Figure IX-13(a) demonstrates the reaction of individual firms to this increase in demand. At P_2, MR > MC at the original level of production, q_1. Thus, the profit maximizing firms will increase output along their short-run supply curve, up to the point where MR = MC. For this particular firm, MR = MC after the demand shift at an output of q_2. As all firms in the industry respond in this way to the price increase, industry supply will increase to Q_2 in panel (b). Note that in the short-run, each firm is earning an economic profit equal to the area P_2EFG. This indicates that an increase in demand can generate profits for firms because in the short-run, before more competitors enter the market, firms are able to sell more quantity at a higher price without substantial increases to their costs.

In the long-run, the firms' economic profits disappear because of the entry of new firms into the industry. Firms that are in industries earning a normal profit or an economic loss have an incentive to switch their productive resources to industries in which positive economic profits are being earned. The entry of these new firms causes the supply of the industry product to increase. New firms will continue to enter and industry supply will

### Figure IX-13. Long-Run Adjustment to an Increase in Demand
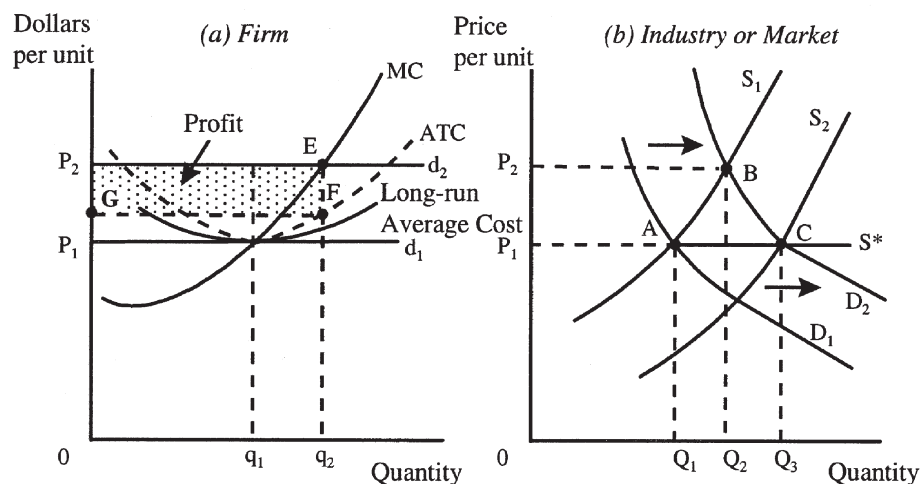
increase as long as individual firms in the industry are capable of earning economic profits. Thus, in the long-run, equilibrium moves from point B to point C as supply shifts from S_1 to S_2. At this point, equilibrium quantity has increased to Q_3 and equilibrium price is once again P_1. This example also demonstrates the difference between a move along the supply curve and a shift in the supply curve. In the short-run the firm is able to supply more and gain economic profits whereas in the long-run the entire industry will produce more and those profits will reduce to 0.

The effect on the individual firm of the decrease in price that results from new firms entering the industry can be seen in Figure IX-13(a). The fall in the market price causes the demand curve faced by the individual firm to decline from d_2 to d_1. As a result, the firm reduces output from q_2 to q_1. At this equilibrium price and quantity combination, the firm earns only a normal profit. In other words, the individual firm's marginal cost, price (marginal revenue), and long-run average cost are equal and the long-run average cost curve is at a minimum—MC = MR = P = LRAC. In the long-run, the price of a good or service is determined by the minimum point on the firm's long-run average cost curve. Despite the fact that industry output increases from Q_1 to Q_3, individual firm output remains at q_1 in the long-run. Thus, the increase in industry output is supplied by the entry of new firms.

A decrease in demand has an opposite effect on both the industry and the individual firm. Figure IX-14(b), shows the industry in equilibrium at point A, with demand curve, D_1, supply curve, S_1, market price, P_1, and quantity, Q_1. Point A represents the current long-run equilibrium, where MC = MR = ATC = LRAC. The equilibrium condition for the firm is shown in Figure IX-14(a).

Let us suppose that demand in this industry suddenly decreases. This decrease is seen as a shift in demand from D_1 to D_2. This decline in demand causes the market price to decrease to P_2—the market is now in equilibrium at point B in the short-run. Because of the decrease in market price, the demand curve facing the individual firm declines. This event can be seen in Figure IX-14(a) as movement from d_1 to d_2. Again, the firm will determine the level of output to produce by identifying the point at which MR = P = MC. This causes individual firms to cut production from q_1 to q_2. As a result of each

Figure IX-14. Long-Run Adjustments to a Decrease in Demand
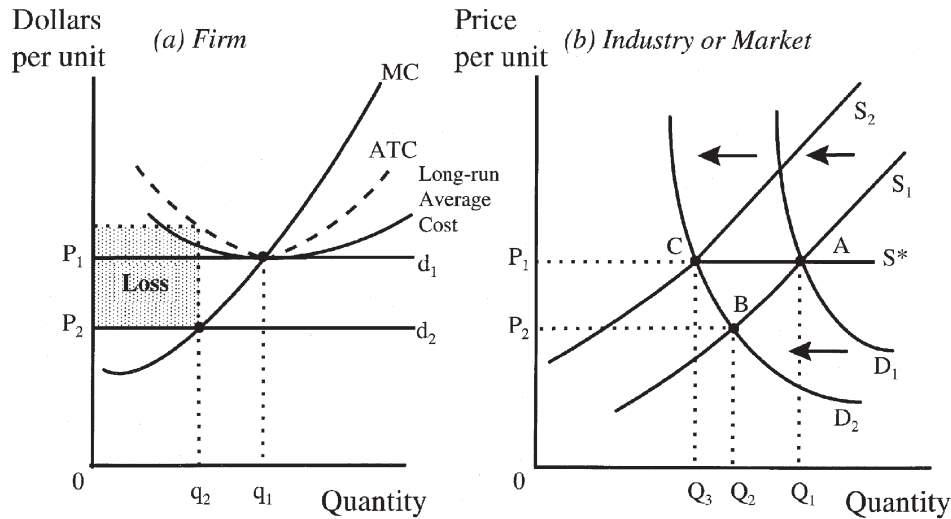
individual firm cutting back output in the short-run, industry-wide output declines from Q_1 to Q_2. This is a change in the quantity supplied which is indicated by movement along the supply curve from point A to point B. Thus, the market is now in equilibrium at point B, with an output of Q_2 and a price P_2.

At this lower price, the firm is selling q_2 units at a price below the average total cost of production. In other words, the firm operates at a loss equal to the vertical difference between the ATC curve and the demand curve at q_2 multiplied by total output, q_2. This is the shaded portion of Figure IX-14(a). In the long-run, these losses force some firms to exit the industry. As these firms leave, market supply decreases. Firms will continue to exit the market until the supply curve has shifted from S_1 to S_2. In the long-run, the market will be in equilibrium at point C, where S_2 and D_2 intersect. As a result of these changes, industry output has declined to Q_3 and market price has returned to P_1. At this new equilibrium position, marginal cost, price (marginal revenue), short-run average total cost, and long-run average cost are equal—P = MC = ATC = LRAC. In other words, individual firms in the industry are once again earning a normal profit. As was the case with an increase in demand, after long-run adjustments, individual firms again produce an output of q_1— the same output as before the change in demand. However, market output has declined from Q_1 to Q_3. The decline in industry output came as the result of firms exiting the market and not as a result of cuts in the remaining firms' production. As you might have expected this result was quite predictable. In the short-run, when there is a sudden decline in demand, industry output and price will drop and firms will lose money. In the long-run, some firms will drop out of the market reducing supply which will bring price back up to a stable equilibrium. Because demand is often out of control of the firms selling goods, markets constantly see this give and take of fluctuating prices and adjustments to supply.

The above analysis considered the response of an industry and its individual firms to a change in demand. In the short-run, firms responded by changing the quantity supplied to the market—movement along their marginal cost curve. In the long-run, however, firms entered or exited the industry until a new short-run supply curve was established that returned individual firms to a normal profit. The long-run supply curve for any particular industry can be identified by connecting the points of long-run equilibrium.

This can be seen in Figures IX-13(b) and IX-14(b) as the horizontal line marked S*. This long-run supply curve is horizontal and, thus, totally elastic. This implies that as output expands, the costs of production do not increase because in the long-run, firms do not need a price increase to supply the additional output. This type of industry is known as a **constant cost industry** because it can expand without affecting the prices of the resources it employs.

A constant cost industry is usually one in which the industry consumes only a small portion of the total amount of a particular productive resource available in the market. The producers of golf tees are an example of a constant cost industry. A large increase in the demand for golf tees is not likely to cause the market price of timber to increase. The production of golf tees utilizes only a small portion of all wood that is available as a production input.

Increasing cost industries are industries where increasing output in the long-run leads to an increase in the costs of production. Because the costs of some resources increase as output expands in these industries, individual firms' cost curves shift upward. Thus, long-run supply curves for increasing cost industries have a positive slope. The oil industry is an example of an increasing cost industry. As production of oil goes up, the obvious and convenient places to drill for oil run out and it becomes increasingly costly and difficult to find and reach new sources of oil. **Decreasing cost industries** experience lower production costs as output increases in the long-run. This decrease in the cost of production causes a downward shift in the individual firms' cost curves. Thus, such firms experience downward sloping supply curves in the long-run. An example of a decreasing cost industry would be many manufacturing industries because the marginal costs of manufacturing often decline as output expands because of economies of scale. Regardless of the effect of expanding output on the costs of production, the long-run conclusion is always the same. Individual firms in perfectly competitive markets can earn economic profits in the short-run, but the entry or exit of firms in the long-run results in a normal economic profit.

e. The Economic Benefits of Competition

When firms produce exactly the goods that consumers desire most at the lowest possible price, they are said to be achieving **allocative efficiency**. Perfect competition results in allocative efficiency. The market demand curve reflects the consumers' value for each additional unit of consumption, and the value attached to the last unit of output purchased by consumers is equal to the market price. In the perfectly competitive model, market price equals the marginal cost of the last unit producers are willing to supply. Furthermore, marginal cost is defined as a measure of the opportunity cost of the resources used in the production of each additional unit. By assumption the marginal benefit consumers attach to the last unit consumed is equal to the marginal opportunity cost of the resources used to produce that same unit, and thus the last unit produced is valued as much or more than any other good or service that could have been produced using those same resources. In other words, resources could not be reallocated, in any way, in order to increase the value of output to society. That is, allocative efficiency is achieved—and social welfare is maximized—in perfectly competitive industries. This is also a condition that economists refer to as Pareto efficiency. It represents a point where no one can be made any better off by reallocating resources.

The perfect competition model provides a **competitive benchmark** for the economic analysis of other industry structures. In general, industry production at the quantity where P = MC provides a benchmark for comparing an industry's performance with the allocative efficiency of the ideal perfectly competitive market.

In the perfect competition model, market price reflects the value of the resources that are used in the production of a product. This is important because prices are the most important signal in guiding the allocation of resources. If prices are distorted by a lack of competition, then the allocation of resources that results is less than efficient. Consider the following example.

Suppose, in the beginning, everyone in a small town bakes his or her own bread. One day a baker moves into town and decides that there is a profit to be made by baking bread for others. The baker adds all the expenses for a loaf of bread, including labor time, plus a premium for risk bearing, and determines that this cost should be the price to charge customers. Many people in town decide that it is cheaper to buy bread than to bake it themselves, and use their increased time to produce other products. Those people who do not produce other products or services with their time continue to bake their own bread. After a while, the baker determines that she could make more profit by exerting her market power and raising her price. Although some customers would return to baking their own bread, the baker would still make more revenue from the loaves sold than she would lose from the reduced volume.

Although it may appear that the increase in profits to the baker is simply a transfer of wealth from her customers and that there is no net effect on society, there is a hidden efficiency loss that makes society worse off. The customers who return to baking their own bread because of the artificially higher price must now reduce the time they spend producing other products in order to bake bread for themselves. Hence, society loses other valuable products because the price of bread has increased.

If baking bread were a relatively easy industry to enter, new competitors would enter the market and charge a lower price. The original baker, finding that she is losing her customers, would have to match her new competitor's price or perhaps charge less in order to get all of her customers back. The price war would continue until neither baker could lower prices any more without losing money. At this point, the price of bread accurately reflects the value of the resources used in manufacture. That is, competition has generated a market price that equals the marginal cost of production. Customers who returned to baking their own bread can buy their bread again and return to producing other products.

2. Monopoly

A monopoly exists when there is only one producer of a particular good or service in a market with no competition. Often these markets are characterized by a lack of close substitutes and the producer is protected from competition by impenetrable barriers to entry. Barriers to entry allow monopolies to earn higher than normal profits by reducing their output and raising prices above competitive levels. Monopolies are a form of market failure that results in a misallocation of resources. Yet, the very existence of monopolies poses a difficult problem for policy makers—in a market economy, monopoly profits encourage innovation and efficiency. Public policies designed to restrain monopolistic practices must be carefully designed to avoid dampening competitive incentives.

a. Profit Maximizing Price and Quantity

A monopolist is the only producer of a particular good or service and, therefore, by its actions alone, determines the market price. In contrast to the perfectly competitive model where the demand curve faced by price taking firms is horizontal and equal to a market price established by the interaction of industry supply and demand, the monopolist is the only firm in the industry and its demand curve is the same as the industry demand

curve. Thus, the monopolist faces a downward sloping demand curve—the monopolist must lower price in order to sell additional units. This has important implications for the monopolist's production decisions because, unlike the price taker, the monopolist's marginal revenue does not equal price.

For the monopolist, marginal revenue is less than price except for the very first unit sold. The divergence of marginal revenue and price occurs because the demand curve faced by the monopolist is downward sloping. When demand is downward sloping, and a uniform price is charged for all units sold, marginal revenue will always be less than market price. This result is illustrated in Figure IX-15. A price cut from P_1 to P_2 increases the quantity sold from Q_1 to Q_2. So, total revenue at point A on the demand curve is equal to P_1AQ_10; at point B, it is equal to P_2BQ_20. The change in total revenue in moving from A to B is the result of two offsetting influences. First, revenue is increased as a result of the sale of additional units of output from Q_1 to Q_2. This increase in revenue is the area CBQ_2Q_1. Second, revenue is lost due to a lower price on units 0 to Q_1 that could have been sold at the higher price P_1 instead of P_2. This reduction is represented by the area P_1ACP_2. The marginal revenue derived from the sale of an additional unit is the difference between these two offsetting effects. That is, marginal revenue is the difference between the total revenue at A and the total revenue at B. Because the marginal revenue for any output charged is less than the corresponding price for that level of output, the marginal revenue (MR) curve will always lie below the demand curve—with the one exception being the first unit sold. For the first unit sold, marginal revenue is equal to market price. Beyond that the general rule always holds—$P > MR$. When a monopolist lowers prices it must ask whether the new revenue from increased sales will offset the lost revenue from lower prices. The answer to that question depends heavily on understanding the effect of demand elasticity.

Figure IX-16, which shows a standard demand curve with MR below price at every level of output, demonstrates the relationship between total revenue, marginal revenue,

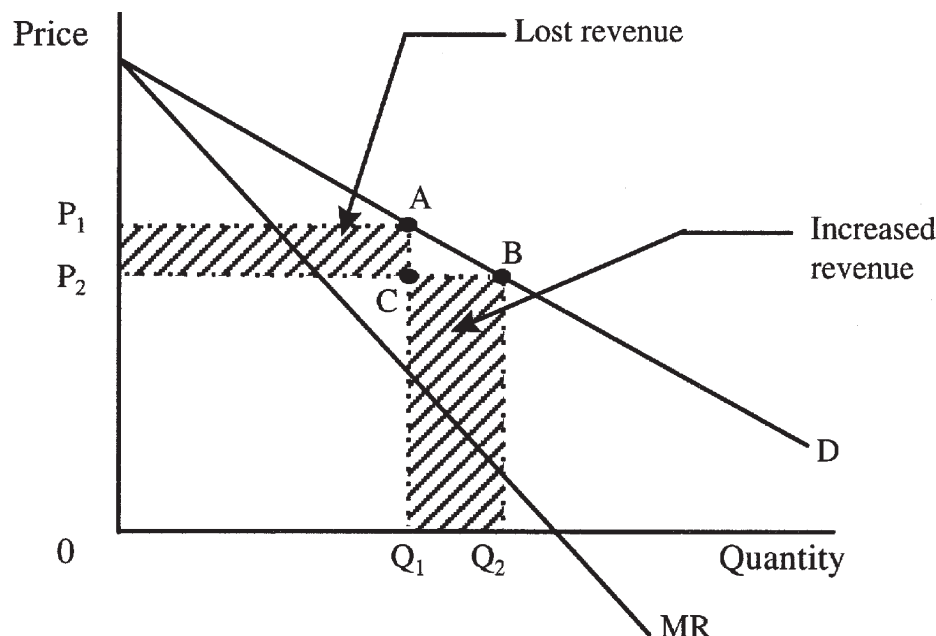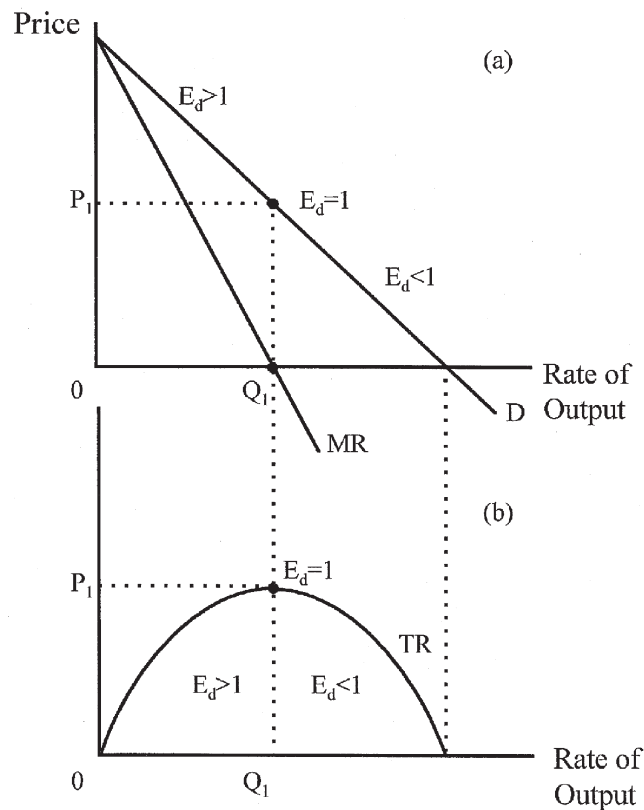**Figure IX-15. The Dual Effects of a Price Reduction on Total Revenues**

Figure IX-16. Changes in Elasticity of Demand and Total Revenue as Price Changes

demand, and elasticity. Recall from Chapter II that when price reductions cause total revenue to rise, demand is elastic. If MR > 0, then TR increases with increases in output. The only way to sell the increased output is to lower the price. If the lower price is more than offset by increased output, then TR increases and demand is elastic. Thus, if MR > 0, then demand is elastic. In Figure IX-16, $E_d > 1$ to the left of Q_1 and $E_d < 1$ to the right of Q_1. A minute change in price around P_1 will not change total revenue, so $E_d = 1$ at output Q_1.

The relationship between total revenue, marginal revenue, and demand elasticity is the starting point for analyzing the profit maximizing output and pricing decisions of a monopolist. Common sense tells us that no rational profit maximizer will produce output in the range where marginal revenue is negative. Thus, the profit maximizing monopolist will always operate along the elastic or unit elastic portion of the demand curve—at a quantity below Q_1 in Figure IX-16. If a firm were to find itself at an output level in the inelastic range, it could increase profits by simply reducing output (and costs) while increasing revenue.

Of course, in determining the profit maximizing level of output for the monopolist, the costs of production must be considered. Generally speaking, the monopolist has no special power in the resource or input market. In other words, when it comes to buying factors of production, the monopolist is just like any perfectly competitive firm. Thus, the cost curves for the monopolist have the same characteristics as those of the perfectly

competitive firm. Moreover, profit maximizing monopolists follow the same general rule as firms in a perfectly competitive industry — expand production until marginal revenue (MR) equals marginal cost (MC). The monopolist follows this rule for exactly the same reasons as the perfect competitor — it will only produce the next unit of production if it earns more money selling the unit than it costs to produce that unit. This is the only way to grow profits.

The profit maximizing decision of the monopolist is demonstrated in Figure IX-17. The profit maximizing output is Q*, where MR = MC. To the left of Q*, MR > MC and the firm can increase profits by expanding production. To the right of Q*, MC > MR and the firm can increase profits by reducing production. At Q*, the firm cannot increase profits by changing the level of production — profits are maximized at Q*. The demand curve shows that Q* units can be sold at P*. In Figure IX-17, this corresponds with point A on the demand curve. Thus, P* is the profit maximizing price.

The level of monopoly profits is calculated in the same manner as profits for the perfectly competitive firm. Specifically, profit is equal to Q*(P* − ATC), for ATC at the output level where MR = MC. In Figure IX-17, monopoly profits are equal to the area P*ABC.

When the monopolist earns a profit, it means that the firm has earned a return in excess of the opportunity cost of the resources used to produce its output. In a perfectly competitive industry, above normal profits attract the entry of new firms. However, entry is not possible for a pure monopoly industry because, by definition, new entry is precluded. Thus, unlike the competitive industry, which is restored to a normal return in the long-run, economic profits can continue to exist in the long-run for the monopolist.

An important note of caution is in order here. Not all monopolies are profitable. Consider the cost and revenue information for the monopolist graphed in Figure IX-18.
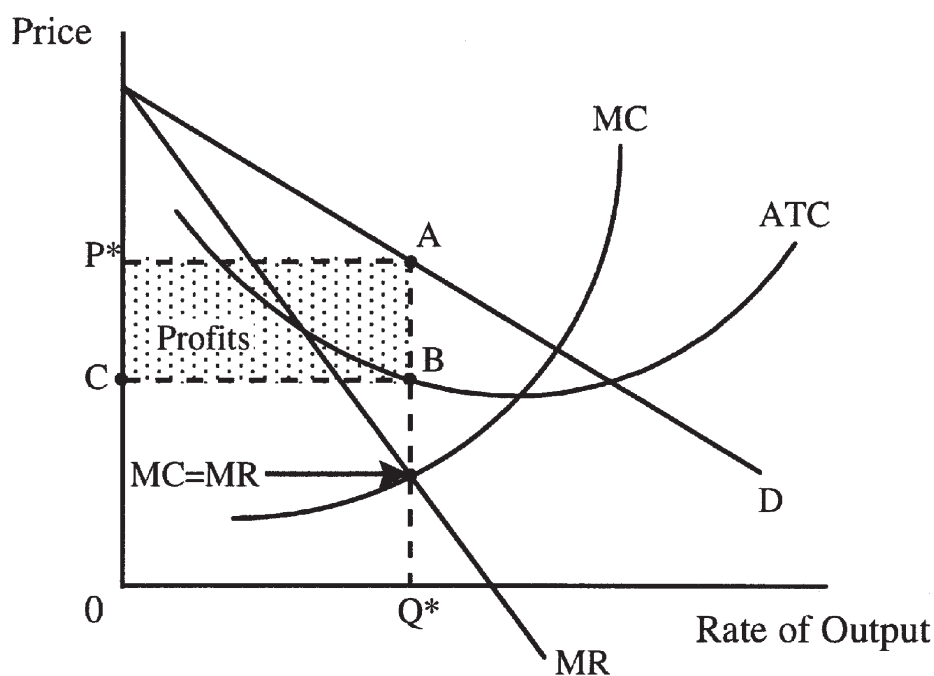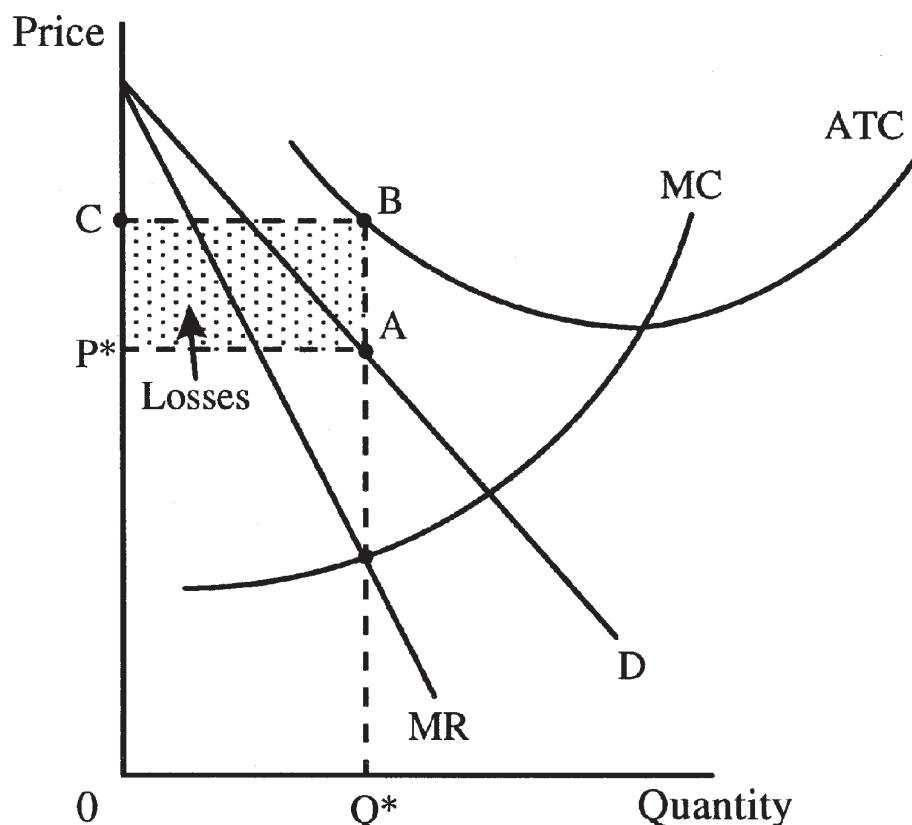
Figure IX-17. Monopoly Profits

Figure IX-18. An Unprofitable Monopoly

MR = MC at Q* which yields a profit maximizing price of P*. Q* corresponds to point A on the demand curve and point B on the ATC curve. Thus, at Q*, ATC > P*, and losses equal to Q*(P* – ATC) are shown by the rectangle ABCP*. When a monopolist suffers a loss, the firm is earning a return on the resources it uses that is less than their opportunity cost. As under perfect competition, in the face of a short-run economic loss, the monopolist has a choice between continuing production or shutting down. The same shutdown rule applies here as well. That is, as long as revenue exceeds variable cost, the monopolist firm will continue to produce in the short-run. If these loss conditions persist in the long-run, however, the firm will exit the industry and transfer its resources to an industry in which it can earn a normal return.

One final point must be made regarding the profit maximizing price and quantity decision of a monopolist. A supply curve is evidence of a unique relationship between market price and the quantity that the firm is willing and able to supply. In a perfectly competitive market, the price taking firm's marginal cost curve is the same as its supply curve. The monopolist, however, does not set price equal to marginal cost. In order to determine the profit maximizing output, the monopolist needs to know the marginal revenue and marginal cost curves. After determining the profit maximizing quantity, the monopolist finds the corresponding price on the demand curve. Thus, for the monopolist, there is no unique relationship between quantity supplied and price. In other words, because the monopolist is the only seller in the market, it does not have a supply curve.

b. Barriers to Entry

But for barriers to entry, long-run profits would not exist in a pure monopoly. Monopoly profits would dissipate as a result of competition from new entrants. Thus, the fundamental source of monopoly power is the fact that barriers to entry persist even in the long-run.

Economic barriers to entry can result from very large economies of scale or the total control of an essential resource. Previously, we introduced the notion of a long-run average cost curve that decreases continually as output expands. This suggests that one firm could supply the entire industry output at the lowest possible cost. When such conditions arise, the process of competition guarantees that only one firm will survive. Again, firms that gain monopoly power because of these pronounced economies of scale are called **natural monopolies**. The early development of railroads is a good example of a natural monopoly because building a railroad requires huge fixed costs and massive construction capabilities. These costs represent a substantial barrier to entry to competitors, and railroad companies benefit from economies of scale in production.

Another economic barrier to entry is the control of the entire supply of an essential resource. In this case, entry by new firms is blocked because they cannot gain access to the particular resource. A common example is De Beers Diamond Jewellers, which controls a large portion of the world's diamond mines. You might be surprised to discover that diamonds are not a particularly rare gem but that this company's control has artificially created scarcity and raised prices. Absolute control over a resource will be reflected in the value of the firm. Thus, individuals who purchase a firm that controls an essential resource will not earn a monopoly profit.

Legal barriers to entry are restrictions placed upon the entry of new firms into an industry by the government. A familiar example of a government granted monopoly is the right to be the sole provider of basic public utilities. Other examples include patents and licenses. **Patents** give the inventor of a product or process a monopoly over the use of that invention for 20 years from the filing date. **Licenses** are required to participate in many professions — e.g., medical and legal — and to operate many types of businesses — e.g., radio stations, cable, and airlines. The rent-seeking costs of monopoly discussed below (and more generally in Chapter III) often result from attempts by businesses to obtain legal barriers to entry. Firms sometimes attempt to erect entry barriers through the use of private contracts or even technological hurdles. Consider the case below.

United States v. Microsoft Corp.

United States Court of Appeals for the District of Columbia Circuit

253 F.3d 34 (2001)

Microsoft Corporation appeals from judgments of the District Court finding the company in violation of §§ 1 and 2 of the Sherman Act and ordering various remedies.

The action against Microsoft arose pursuant to a complaint filed by the United States and separate complaints filed by individual States. The District Court determined that Microsoft had maintained a monopoly in the market for Intel-compatible PC operating systems in violation of § 2; attempted to gain a monopoly in the market for internet browsers in violation of § 2; and illegally tied two purportedly separate products, Windows and Internet Explorer ("IE"), in violation of § 1. *United States v. Microsoft Corp.,* 87 F. Supp. 2d 30 (D.D.C. 2000).

* * *

II. MONOPOLIZATION

Section 2 of the Sherman Act makes it unlawful for a firm to "monopolize." The offense of monopolization has two elements: "(1) the possession of monopoly power in the relevant market and (2) the willful acquisition or maintenance of that power as distinguished from growth or development as a consequence of a superior product, business acumen, or historic accident." *United States v. Grinnell Corp.*, 384 U.S. 563, 570–71 (1966). The District Court applied this test and found that Microsoft possesses monopoly power in the market for Intel-compatible PC operating systems. Focusing primarily on Microsoft's efforts to suppress Netscape Navigator's threat to its operating system monopoly, the court also found that Microsoft maintained its power not through competition on the merits, but through unlawful means. Microsoft challenges both conclusions. We defer to the District Court's findings of fact, setting them aside only if clearly erroneous. We review legal questions *de novo*.

We begin by considering whether Microsoft possesses monopoly power … and finding that it does, we turn to the question whether it maintained this power through anticompetitive means. Agreeing with the District Court that the company behaved anticompetitively, and that these actions contributed to the maintenance of its monopoly power, we affirm the court's finding of liability for monopolization.

A. *Monopoly Power*

While merely possessing monopoly power is not itself an antitrust violation, *see* Northeastern Tel. Co. v. AT & T, 651 F.2d 76, 84–85 (2d Cir.1981), it is a necessary element of a monopolization charge, The Supreme Court defines monopoly power as "the power to control prices or exclude competition." United States v. E.I. du Pont de Nemours & Co., 351 U.S. 377, 391 (1956). More precisely, a firm is a monopolist if it can profitably raise prices substantially above the competitive level. Where evidence indicates that a firm has in fact profitably done so, the existence of monopoly power is clear. FTC v. Indiana Fed'n of Dentists, 476 U.S. 447, 460–61 (1986) (using direct proof to show market power in Sherman Act § 1 unreasonable restraint of trade action). Because such direct proof is only rarely available, courts more typically examine market structure in search of circumstantial evidence of monopoly power. 2A Areeda et al., Antitrust Law ¶ 531a, at 156. Under this structural approach, monopoly power may be inferred from a firm's possession of a dominant share of a relevant market that is protected by entry barriers. "Entry barriers" are factors (such as certain regulatory requirements) that prevent new rivals from timely responding to an increase in price above the competitive level.

The District Court considered these structural factors and concluded that Microsoft possesses monopoly power in a relevant market. Defining the market as Intel-compatible PC operating systems, the District Court found that Microsoft has a greater than 95% share. It also found the company's market position protected by a substantial entry barrier.

Microsoft argues that the District Court incorrectly defined the relevant market. It also claims that there is no barrier to entry in that market. Alternatively, Microsoft argues that because the software industry is uniquely dynamic, direct proof, rather than circumstantial evidence, more appropriately indicates whether it possesses monopoly power. Rejecting each argument, we uphold the District Court's finding of monopoly power in its entirety.

1. Market Structure
a. Market definition

"Because the ability of consumers to turn to other suppliers restrains a firm from raising prices above the competitive level," Rothery Storage & Van Co. v. Atlas Van Lines,

Inc., 792 F.2d 210, 218 (D.C.Cir.1986), the relevant market must include all products "reasonably interchangeable by consumers for the same purposes." *du Pont,* 351 U.S. at 395. In this case, the District Court defined the market as "the licensing of all Intel-compatible PC operating systems worldwide," finding that there are "currently no products—and … there are not likely to be any in the near future—that a significant percentage of computer users worldwide could substitute for [these operating systems] without incurring substantial costs." Calling this market definition "far too narrow," Microsoft argues that the District Court improperly excluded three types of products: non-Intel compatible operating systems (primarily Apple's Macintosh operating system, Mac OS), operating systems for non-PC devices (such as handheld computers and portal websites), and "middleware" products, which are not operating systems at all.

* * *

b. Market power

Having thus properly defined the relevant market, the District Court found that Windows accounts for a greater than 95% share. The court also found that even if Mac OS were included, Microsoft's share would exceed 80%. Microsoft challenges neither finding, nor does it argue that such a market share is not predominant. [The court cited three Supreme Court decisions that found market shares of 87%, 80%, and 75% predominant.]

Instead, Microsoft claims that even a predominant market share does not by itself indicate monopoly power. Although the "existence of [monopoly] power ordinarily may be inferred from the predominant share of the market," *Grinnell,* 384 U.S. at 571, we agree with Microsoft that because of the possibility of competition from new entrants, looking to current market share alone can be "misleading." Hunt-Wesson Foods, Inc. v. Ragu Foods, Inc., 627 F.2d 919, 924 (9th Cir.1980). In this case, however, the District Court was not misled. Considering the possibility of new rivals, the court focused not only on Microsoft's present market share, but also on the structural barrier that protects the company's future position. That barrier—the "applications barrier to entry"—stems from two characteristics of the software market: (1) most consumers prefer operating systems for which a large number of applications have already been written; and (2) most developers prefer to write for operating systems that already have a substantial consumer base. This "chicken-and-egg" situation ensures that applications will continue to be written for the already dominant Windows, which in turn ensures that consumers will continue to prefer it over other operating systems.

Challenging the existence of the applications barrier to entry, Microsoft observes that software developers do write applications for other operating systems, pointing out that at its peak IBM's OS/2 supported approximately 2,500 applications. This misses the point. That some developers write applications for other operating systems is not at all inconsistent with the finding that the applications barrier to entry discourages many from writing for these less popular platforms. Indeed, the District Court found that IBM's difficulty in attracting a larger number of software developers to write for its platform seriously impeded OS/2's success.

Microsoft does not dispute that Windows supports many more applications than any other operating system. It argues instead that "[i]t defies common sense" to suggest that an operating system must support as many applications as Windows does (more than 70,000, according to the District Court) to be competitive. Consumers, Microsoft points out, can only use a very small percentage of these applications. As the District Court explained, however, the applications barrier to entry gives consumers reason to prefer the dominant operating system even if they have no need to use all applications written for it:

> The consumer wants an operating system that runs not only types of applications that he knows he will want to use, but also those types in which he might develop an interest later. Also, the consumer knows that if he chooses an operating system with enough demand to support multiple applications in each product category, he will be less likely to find himself straitened later by having to use an application whose features disappoint him. Finally, the average user knows that, generally speaking, applications improve through successive versions. He thus wants an operating system for which successive generations of his favorite applications will be released—promptly at that. The fact that a vastly larger number of applications are written for Windows than for other PC operating systems attracts consumers to Windows, because it reassures them that their interests will be met as long as they use Microsoft's product.

Findings of Fact ¶ 37. Thus, despite the limited success of its rivals, Microsoft benefits from the applications barrier to entry.

Of course, were middleware to succeed, it would erode the applications barrier to entry. Because applications written for multiple operating systems could run on any operating system on which the middleware product was present with little, if any, porting, the operating system market would become competitive. But as the District Court found, middleware will not expose a sufficient number of APIs [Application Programming Interfaces] to erode the applications barrier to entry in the foreseeable future.

Microsoft next argues that the applications barrier to entry is not an entry barrier at all, but a reflection of Windows' popularity. It is certainly true that Windows may have gained its initial dominance in the operating system market competitively—through superior foresight or quality. But this case is not about Microsoft's initial acquisition of monopoly power. It is about Microsoft's efforts to maintain this position through means other than competition on the merits. Because the applications barrier to entry protects a dominant operating system irrespective of quality, it gives Microsoft power to stave off even superior new rivals. The barrier is thus a characteristic of the operating system market, not of Microsoft's popularity, or, as asserted by a Microsoft witness, the company's efficiency.

Finally, Microsoft argues that the District Court should not have considered the applications barrier to entry because it reflects not a cost borne disproportionately by new entrants, but one borne by all participants in the operating system market. According to Microsoft, it had to make major investments to convince software developers to write for its new operating system, and it continues to "evangelize" the Windows platform today. Whether costs borne by all market participants should be considered entry barriers is the subject of much debate. *Compare* 2A Areeda & Hovenkamp, Antitrust Law §420c, at 61 (arguing that these costs are entry barriers), *and* Joe S. Bain, Barriers to New Competition: Their Character and Consequences in Manufacturing Industries 6–7 (1956) (considering these costs entry barriers). We need not resolve this issue, however, for even under the more narrow definition it is clear that there are barriers. When Microsoft entered the operating system market with MS-DOS and the first version of Windows, it did not confront a dominant rival operating system with as massive an installed base and as vast an existing array of applications as the Windows operating systems have since enjoyed. Moreover, when Microsoft introduced Windows 95 and 98, it was able to bypass the applications barrier to entry that protected the incumbent Windows by including APIs from the earlier version in the new operating systems. This made porting existing Windows applications to the new version of Windows much less costly than porting them to the operating systems of other entrants who could not freely include APIs from the incumbent Windows with their own.

* * *

Notes and Questions

1. Market Shares and Monopolies: The court finds that Microsoft's 95 percent market share is persuasive evidence of monopoly power. Are market shares always a good proxy for monopoly power? What is the difference between market power and monopoly power? Is there a clear cut-off between a firm with monopoly power and one without it, or is it more of a spectrum of influence and control? Although a rough rule, Judge Learned Hand in the case United States v. Aluminum Co. of Am., 148 F.2d 416, 424 (2d Cir. 1945), articulated a rule of thumb for market shares and monopolists whereby "it is doubtful whether sixty or sixty-four percent would be enough; and certainly thirty-three per cent is not." However, he concluded that the Aluminum Company of America was a monopolist with around ninety percent of the market. Think about some of the biggest and most dominant brand names in today's American economy. Can you think of any monopolies?

2. Market Power and Emerging Technologies: The government's case for Microsoft's market power was much more complex than just market shares. The government pointed to and relied on multiple kinds of proof, including market shares, barriers to entry, network effects, pricing data, profit margins, and more. Considering the rise of competing technology giants such as Apple, Google, and Facebook, does the court's opinion seem wrong or overly cautious by today's standards? Doesn't the subsequent existence of competing technology giants disprove the anticompetitive nature of emerging technology markets? This was a concern that the D.C. Circuit certainly wrestled with in its opinion but ultimately rejected. The court considered that:

> Rapid technological change leads to markets in which "firms compete through innovation for temporary market dominance, from which they may be displaced by the next wave of product advancements." [A. Shelanski & J. Gregory Sidak, Antitrust Divestiture in Network Industries, 68 U. Chi. L. Rev. 1, 11–12 (2001)] (discussing Schumpeterian competition, which proceeds "sequentially over time rather than simultaneously across a market"). Microsoft argues that the operating system market is just such a market.
>
> Whether or not Microsoft's characterization of the operating system market is correct does not appreciably alter our mission in assessing the alleged antitrust violations in the present case.

United States v. Microsoft Corp., 253 F.3d 34, 49–50 (D.C. Cir. 2001).

3. How Can Monopolies Compete?: The allegation in the Microsoft case was that Microsoft had impeded the growth and entry of potential rivals Java and Netscape into competing markets. How much of a threat do those rivals have to be to raise antitrust claims? If Microsoft could prove that these companies were going to fail anyway, would it change the analysis? Is Microsoft unable to compete vigorously with potential competitors solely because it is a monopolist? If so, what limitations are there to Microsoft's ability to compete?

4. Drastic Measures: The district court sought a remedy to break up Microsoft into two smaller companies to prevent it from causing anticompetitive harm to the market. The decision sought to separate the hardware and operating system of Microsoft from its related products and software. This drastic remedy was related to a claim in the case that Microsoft was unlawfully using its monopoly in the operating system market to gain market share

in the web browser market—a cause of action referred to as tying in antitrust law. This severe remedy was never realized because it was vacated and remanded by the D.C. Circuit, and the case settled before that remand was issued. The D.C. Circuit explained that:

> [J]ust over six years have passed since Microsoft engaged in the first conduct plaintiffs allege to be anticompetitive. As the record in this case indicates, six years seems like an eternity in the computer industry. By the time a court can assess liability, firms, products, and the marketplace are likely to have changed dramatically. This, in turn, threatens enormous practical difficulties for courts considering the appropriate measure of relief in equitable enforcement actions, both in crafting injunctive remedies in the first instance and reviewing those remedies in the second. Conduct remedies may be unavailing in such cases, because innovation to a large degree has already rendered the anticompetitive conduct obsolete (although by no means harmless). And broader structural remedies present their own set of problems, including how a court goes about restoring competition to a dramatically changed, and constantly changing, marketplace. That is just one reason why we find the District Court's refusal in the present case to hold an evidentiary hearing on remedies—to update and flesh out the available information before seriously entertaining the possibility of dramatic structural relief—so problematic.

United States v. Microsoft Corp., 253 F.3d 34, 49 (D.C. Cir. 2001). This remedy, however, was used against AT&T in the government's earlier case in the 1980s. AT&T was segmented into a number of smaller "Baby Bells" that have subsequently re-merged into a similarly powerful telecommunications conglomerate. Is this remedy a feasible means of deterring anticompetitive conduct or preventing future harmful conduct? Should courts have this power to forcibly divide and split up companies? Can you think of an alternative, less strict remedy that would accomplish the same goal?

c. Price Discrimination

The practice of selling the same product to different customers at different prices is known as **price discrimination**. If three conditions are met, firms can increase their economic profits by engaging in price discrimination. First, the firm must face a downward sloping demand curve—the firm must be a price searcher. Recall that perfectly competitive firms are price takers. These firms have no control over price and therefore, cannot charge different customers different prices. On the other hand, a pure monopolist has a great deal of control over market price. Second, the firm must have some way of identifying separate groups of customers with differing demands for its product. In general, the price discriminating seller will charge a higher price to those customers with a relatively inelastic demand. Third, the firm must be able to prevent **arbitrage**—resale by low-price purchasers to the higher price purchasers.

Commonly observed forms of price discrimination satisfy these criteria. Take the example of "senior citizen discounts" at restaurants. Senior citizens typically have more elastic demand curves, they are easily identified, and the goods and services (like meals) sold to them at a discount are difficult to resell. Judge Frank Easterbrook recently described an innovative manner by which software developers can price discriminate between consumer and commercial users of software:

> ProCD, the plaintiff, has compiled information from more than 3,000 telephone directories into a computer database.... ProCD sells a version of the database, called SelectPhone (trademark), on CD-ROM discs.... The database in

SelectPhone (trademark) cost more than $10 million to compile and is expensive to keep current. It is much more valuable to some users than to others. The combination of names, addresses, and SIC codes enables manufacturers to compile lists of potential customers. Manufacturers and retailers pay high prices to specialized information intermediaries for such mailing lists; ProCD offers a potentially cheaper alternative. People with nothing to sell could use the database as a substitute for calling long distance information, or as a way to look up old friends who have moved to unknown towns, or just as an electronic substitute for the local phone book. ProCD decided to engage in price discrimination, selling its database to the general public for personal use at a low price (approximately $150 for the set of five discs) while selling information to the trade for a higher price....

If ProCD had to recover all of its costs and make a profit by charging a single price—that is, if it could not charge more to commercial users than to the general public—it would have to raise the price substantially over $150. The ensuing reduction in sales would harm consumers who value the information at, say, $200. They get consumer surplus of $50 under the current arrangement but would cease to buy if the price rose substantially. If because of high elasticity of demand in the consumer segment of the market the only way to make a profit turned out to be a price attractive to commercial users alone, then all consumers would lose out—and so would the commercial clients, who would have to pay more for the listings because ProCD could not obtain any contribution toward costs from the consumer market.

To make price discrimination work, however, the seller must be able to control arbitrage. An air carrier sells tickets for less to vacationers than to business travelers, using advance purchase and Saturday-night-stay requirements to distinguish the categories. A producer of movies segments the market by time, releasing first to theaters, then to pay-per-view services, next to the videotape and laserdisc market, and finally to cable and commercial TV. Vendors of computer software have a harder task. Anyone can walk into a retail store and buy a box. Customers do not wear tags saying "commercial user" or "consumer user." Anyway, even a commercial-user-detector at the door would not work, because a consumer could buy the software and resell to a commercial user. That arbitrage would break down the price discrimination and drive up the minimum price at which ProCD would sell to anyone.

Instead of tinkering with the product and letting users sort themselves—for example, furnishing current data at a high price that would be attractive only to commercial customers, and two-year-old data at a low price—ProCD turned to the institution of contract. Every box containing its consumer product declares that the software comes with restrictions stated in an enclosed license. This license, which is encoded on the CD-ROM disks as well as printed in the manual, and which appears on a user's screen every time the software runs, limits use of the application program and listings to non-commercial purposes.

ProCD, Inc. v. Zeidenberg, 86 F.3d 1447, 1449–1450 (7th Cir. 1996).

Figure IX-19 illustrates the advantages of price discrimination to a monopolist. The monopolist is able to divide the market into two segments of buyers. The demand for each of these segments is depicted in the two panels of Figure IX-19. The demand for the monopolist's product is relatively more elastic in panel (a) than in panel (b). Because one

Figure IX-19. Price Discrimination

(a) Market A: Buyers With More Elastic Demand

(b) Market B: Buyers With Less Elastic Demand
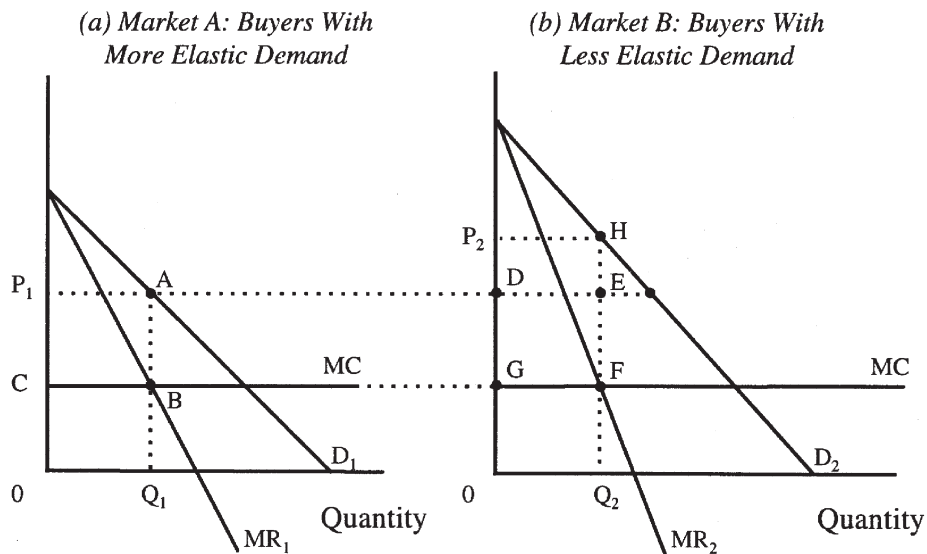
firm supplies both of these "markets," the marginal cost of supplying the product is the same in both markets.

The monopolist desires to maximize profit in each of these "markets." To maximize profits, the monopolist will set $MR_1 = MC$ in panel (a) and $MR_2 = MC$ in panel (b). In panel (a), the market with the more elastic demand, the monopolist charges P_1. However, in panel (b), the market with the relatively inelastic demand, the monopolist charges a higher price of P_2. In this case, individuals with relatively inelastic demand are charged a high price and those with a relative elastic demand are charged a low price.

Price discrimination can increase total welfare by expanding sales to consumers who would not purchase at the single profit-maximizing price. Further price discrimination can lead to more intense competition if firms lure consumers, who prefer rival brands, with special offers as compensation for purchasing a less-preferred brand. See James C. Cooper et al, *Does Price Discrimination Intensify Competition? Implications for Antitrust*, 72 Antitrust Law Journal 327 (2005).

d. Predatory Pricing

Predatory Pricing is a strategy by which a firm reduces its price below its costs of production in order to drive its rivals out of the market in and then to raise prices to capture monopoly profits once it has the market to itself. In practice it is difficult to distinguish predatory behavior from normal rivalrous competition. Because it is rational for a profit maximizing firm to continue to operate in the short-run even if P < ATC, as long as P > AVC, setting price below ATC is not considered evidence of predatory pricing—price must be less than AVC in order to establish something other than normal competitive behavior. By setting price below its average variable cost, the predator firm operates in a range where most firms shut down in the short-run. Thus, the theory of predatory pricing states that the predator attempts to earn monopoly profits in the long run by incurring economic losses in the short-run. It is alleged that such behavior can

increase the value of the firm. For this to occur, the present value of the long-run monopoly profits must be greater than the short-run losses.

A major difficulty for predatory pricing theory is explaining why new firms do not attempt to enter the market once the successful predator begins charging monopoly prices. Remember that the key to reaping monopoly profits in the long-run is barriers to entry. It is not at all clear where the predator gains such a barrier. A common response to this criticism is to trust that the predator's past price reductions discourage potential entrants from entering the market.

Predatory pricing may be attacked under Section 2 of the Sherman Act as an attempt to monopolize. However, the following case raises doubts about the future of predatory pricing as an antitrust violation.

Matsushita Elec. Indus. Co. v. Zenith Radio Corp.

Supreme Court of the United States
475 U.S. 574 (1986)
and accompanying Notes and Questions
Read *supra,* Chapter I

1. Predatory Pricing: The Probability of Success: The Court observes that the success of any predatory pricing scheme depends on the ability of the predator to charge monopoly profits after its rivals have exited the market. Moreover, to be economically rational, the present value of the monopoly profits must exceed the losses incurred in driving rivals out of the market. Zenith and NUE allege that the Japanese firms have been pricing below cost for over twenty years — thereby incurring substantial losses in a not yet successful effort to drive the American firms out of the CEP industry. How long will it take to recoup these losses? Does it seem likely that Japanese firms will be able to charge monopoly prices for a period long enough to break even? How about making an economic profit?

2. Predatory Pricing by a Cartel: The Probability of Success?: For any single firm, predatory pricing is a risky strategy. The risk arises because of the need to maintain barriers to entry after the extinction of rivals for a period long enough to recapture losses and make an economic profit. Zenith and NUE do not accuse a single firm of predatory pricing. Rather, they claim that over twenty-one Japanese firms engaged in this predatory pricing scheme. In other words, Zenith and NUE have accused a cartel of predatory pricing. What about the inherent difficulties of organizing and maintaining a cartel? If the probability of a single firm engaging in predatory pricing is remote, what would happen to that probability if the accusation were against a cartel? The payoff for a successful predatory pricing scheme is the monopoly profits earned after all rivals have been eliminated. Thus, post elimination, a monopoly price and output must be set. For a cartel to engage in predatory pricing, all firms in the cartel must agree on price and output. Not only would coming to such an agreement be difficult, but there is a huge incentive to cheat. After sustaining losses in order to drive rivals out of the market, each individual firm in the cartel will be eager to earn monopoly profits. Each individual firm will believe that by increasing its output by only a small amount over the agreed upon level, no other firm in the cartel will notice and the cheating firm will be able to earn even greater monopoly profits. The greater the number of firms in the cartel, the greater the incentive to cheat. Moreover, the longer the time period involved, the less likely the cartel will succeed. What do these facts say about the alleged Japanese cartel's chance for success? If all of these factors are combined, what can be said about the probability that a predatory pricing scheme was in fact under way in the CEP market?

3. An Alternative Story: Price Discrimination: Recall the prerequisites for a firm to engage in price discrimination. First, the seller must have some way of identifying separate groups of customers with differing demands for the seller's product. Second, the seller must be able to prevent the resale of its good or service from the low-price purchasers to the high-price purchasers. For the Japanese firms, there are clearly two markets for CEPs—one in the United States and one in Japan. Japanese consumers are faced with a well-organized cartel that is protected by government restrictions on foreign competition. Therefore, Japanese consumers face monopoly prices and output. On the other hand, the American CEP market is much more competitive and prices are lower. Thus, Japanese firms face two markets with differing elasticities of demand for the seller's product, and resale between these two markets is prohibited by the Japanese government. What is the effect on American consumers of Japanese price discrimination?

e. The Social Costs of Monopoly

Recall that perfect competition provides a benchmark, P = MC, for evaluating the allocative efficiency of a market. A monopolist operates at a quantity where P > MC and, thus, the monopolist falls short of the competitive benchmark for allocative efficiency. Monopoly is a form of market failure. This section explores the social costs of this market failure.
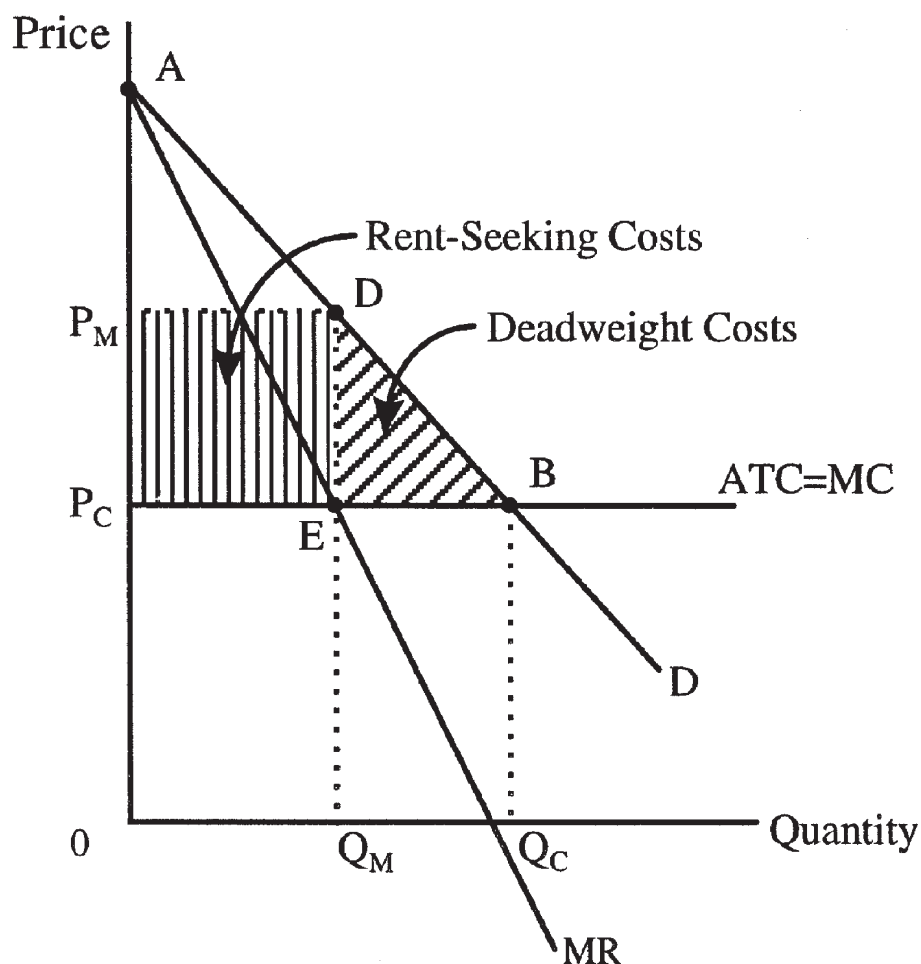
i. Deadweight Costs

The most widely accepted criticism of monopolies is that monopolists distort the allocation of resources by restricting output when buyers are willing to pay a price greater than the marginal costs of producing additional output. Consider the cost and demand curves for a hypothetical industry presented in Figure IX-20. For simplicity, assume that the average total cost of production is the same at all levels of production—that is, this is a constant cost industry and MC = ATC at all levels of output.

If this were a perfectly competitive industry, the firms would produce Q_C where PC = MC. At Q_C, the sum of producer and consumer surplus (ABP_C) is maximized. All potential gains from trade are realized, and the last unit produced is valued by customers at an amount equal to the opportunity cost of the resources used in its production.

The monopolist's profit maximizing quantity occurs at Q_M where MR = MC. Q_M is less than the competitive level, Q_C. As a consequence of the restriction of output, consumer surplus is reduced to the area ADP_M, which is considerably less than the consumer surplus (ABP_C) in the competitive industry. Some of the lost consumer surplus goes to the monopolist in the form of profits, P_MDEP_C. By reducing output and increasing price above a competitive level, monopolists are able to capture a portion of the consumer surplus that would exist in a competitive market. There can be a reasonable debate about whether society benefits more from those profits going to consumers or monopolists. Such distributional issues (i.e., how the economic pie is divided) are usually not addressed by economists because it is difficult to develop and test hypotheses that deal with interpersonal comparisons of utility from wealth.

Economists are concerned, however, with another portion of consumer surplus that disappears when monopolists restrict output—the triangle DBE in Figure IX-20. The monopolist's decision to not produce beyond QM means that buyers located between points D and B on the demand curve are unable to purchase the product even though they are willing to pay a price greater than the marginal costs of producing those additional units. The triangle DBE represents unrealized potential gains from trade. Those

Figure IX-20. Social Costs of Monopoly

opportunities are lost. The triangle DBE represents a reduction in social welfare, which is referred to as the **deadweight costs of monopoly.**

The impact of this deadweight cost on society can be illustrated by the example of a baker. Suppose that the total cost of making a loaf of bread (assuming constant marginal cost) is $1.00, including a normal return to the baker for labor time and entrepreneurship. Suppose, further, that it takes a particular individual, who paints houses competitively for $1.50 an hour, a total of one hour to bake a loaf of bread. When the baker opens the bakery selling bread for $1.00, the painter can buy his bread instead of baking it himself because the value of his time is better spent painting a house for $1.50 than baking bread. Instead of using an hour to bake bread, the painter spends the time painting houses. Not only is the painter better off by $0.50, but society is also better off because it has gained $1.50 of painted houses at a cost of only $1.00. Now imagine that the baker has no competition (that is, she is a monopolist), and she discovers that she can increase profits by increasing the price of her bread to $2.00. As a result, the painter would return to baking his own bread. Society loses $1.50 of output and gains only $1.00 (not $2.00) because the true value of baking bread is only $1.00 — this is the point where P > MC. The deadweight cost equals $0.50. This lost production can never be regained — bygones are bygones.

The size of the deadweight costs has been estimated in several empirical studies by economists.[2] Although the results of the studies are controversial because of their methodology and assumptions, it is worth noting that the studies indicate that the deadweight costs of monopoly are less than one percent of gross national product (GNP). These results sent industrial organization economists scurrying to come up with explanations for why the estimates were so low. After all, it is difficult to justify any government antitrust policy if the potential gains from an effective policy are less than one percent of GNP—and the repeal of the antitrust laws would eliminate a major source of consulting income for economists. Of course, one explanation for the low estimates of deadweight costs is that they prove that antitrust policy was effectively preventing monopolistic practices.

One general criticism of the deadweight cost estimates is that they use above-normal profits as one of the proxies for the presence of monopoly power. Thus, the explanations for the relatively low estimates are, at least in part, really concerned with explaining why monopoly profits are not very high. In that regard, one of the most persuasive explanations for the low deadweight cost estimates is that political competition for the right to earn monopoly profits (economists refer to monopoly profits as **rents**) may have increased the monopolist's costs to the point that the monopoly return was simply a normal profit. This explanation is investigated in the following subsection.

ii. Rent-Seeking Costs

The potential of earning monopoly profits (rents) leads firms to use resources to capture those rents. Such **rent-seeking** expenditures represent an additional social cost of monopoly. Rent-seeking expenditures attempt to use the power of the state to transfer wealth from one group to another. See Chapter III. This use of scarce resources represents a social loss, because those resources could have been used for more productive purposes. That is, rent-seeking expenditures are concerned with merely dividing up the economic pie, rather than increasing the size of the pie.

Firms invest resources in the pursuit of monopoly profits up to the point where the marginal opportunity cost of the last dollar invested is equal to its expected return in increased profits. For example, if the sole domestic producer of a certain product expects a proposed import restriction or tariff (designed to solidify its monopoly) to increase the value of the firm by $1 million (the present value of future monopoly profits), then the domestic producer would be willing to spend up to $1 million on lobbying and campaign contributions in order to secure the favorable regulation. Moreover, those who expect to be injured by the change in the law would be willing to make similar investments to prevent the adoption of the import restriction or tariff. Consequently, rent-seeking expenditures can equal, or even exceed, the total amount of potential monopoly profits.[3] That is, the rectangle representing monopoly profits, $P_M DEP_C$, will be consumed by rent-seeking expenditures. Thus, in Figure IX-19, the social costs of monopoly include the deadweight costs, DBE, plus the rent-seeking costs, $P_M DEP_C$. In an ironic twist of fate,

---

2. The most famous of these studies is Arnold C. Harberger, *Monopoly and Resource Allocation*, 54 American Economic Review 77 (1954).

3. It also can be the case that those who expect to be injured might not have the resources to fight the rent-seeking, they may not even be aware that the rent-seeking is occurring, or there can be a collective action problem—that is, the cumulative injury is spread among many individuals and the transaction costs of cooperating are higher than the resulting benefit of fighting the regulation. In these cases, the monopolist may not need to expend substantial resources on rent-seeking.

American antitrust laws have an exception for private firms that seek to influence the political process and pass laws that might have anticompetitive effects.[4]

iii. Other Criticisms of Monopoly

Some economists have argued that monopolies should be attacked because the lack of product market competition makes the managers lazy and allows them to deviate from technically efficient decisions concerning cost minimization and innovation. As Sir John Hicks, a famous English economist, once put it: "The best of all monopoly profits is a quiet life."[5] There are, however, theoretical reasons for doubting the validity of this hypothesis.

Competition in the product market is not the only constraint on managerial behavior. Managers are forced to attempt to maximize profits by a number of competitive mechanisms other than product market competition. For example, from the typical shareholder's perspective, profits that are sacrificed due to managerial laziness are an opportunity cost that is reflected in a lower stock price. Competition in the stock market will lead to the identification of the poorly performing firm as an attractive takeover target, and managers may lose their jobs in the event of a takeover. Thus, even in monopoly firms, managers have the incentive to maximize profits through cost minimization. Further, the quest to become a monopolist can drive innovation through dynamic competition. This type of competition is present in high-tech markets where disruptive technologies can take over the field only to be replaced by the next big thing. Such jockeying for market dominance is the force behind much of innovation.

Verizon Communications Inc. v. Law Offices of Curtis V. Trinko, LLP

Supreme Court of the United States
540 U.S. 398 (2004)

✳ ✳ ✳

III

The complaint alleges that Verizon denied interconnection services to rivals in order to limit entry. If that allegation states an antitrust claim at all, it does so under § 2 of the Sherman Act, which declares that a firm shall not "monopolize" or "attempt to monopolize." It is settled law that this offense requires, in addition to the possession of monopoly power in the relevant market, "the willful acquisition or maintenance of that power as distinguished from growth or development as a consequence of a superior product, business acumen, or historic accident." *United States v. Grinnell Corp.,* 384 U.S. 563 (1966). The mere possession of monopoly power, and the concomitant charging of monopoly prices, is not only not unlawful; it is an important element of the free-market system. The opportunity to charge monopoly prices—at least for a short period—is what attracts "business acumen" in the first place; it induces risk taking that produces innovation and economic growth. To safeguard the incentive to innovate, the possession of monopoly power will not be found unlawful unless it is accompanied by an element of anticompetitive *conduct.*

---

4. If you are interested in learning how this doctrine works in practice, take a look at *Enforcement Perspectives on the* Noerr-Pennington *Doctrine,* Fed. Trade Comm'n (2006), *available at* http://www.ftc.gov/reports/P013518enfperspectNoerr-Penningtondoctrine.pdf.

5. *See* J.R. Hicks, *Annual Survey of Economic Theory: The Theory of Monopoly*, 3 Econometrica 1, 8 (1935).

Firms may acquire monopoly power by establishing an infrastructure that renders them uniquely suited to serve their customers. Compelling such firms to share the source of their advantage is in some tension with the underlying purpose of antitrust law, since it may lessen the incentive for the monopolist, the rival, or both to invest in those economically beneficial facilities. Enforced sharing also requires antitrust courts to act as central planners, identifying the proper price, quantity, and other terms of dealing—a role for which they are ill suited. Moreover, compelling negotiation between competitors may facilitate the supreme evil of antitrust: collusion. Thus, as a general matter, the Sherman Act "does not restrict the long recognized right of [a] trader or manufacturer engaged in an entirely private business, freely to exercise his own independent discretion as to parties with whom he will deal." *United States v. Colgate & Co.,* 250 U.S. 300, 307 (1919).

* * *

Against the slight benefits of antitrust intervention here, we must weigh a realistic assessment of its costs. Under the best of circumstances, applying the requirements of §2 "can be difficult" because "the means of illicit exclusion, like the means of legitimate competition, are myriad." United States v. Microsoft Corp., 253 F.3d 34, 58 (C.A.D.C.2001) (en banc) *(per curiam).* Mistaken inferences and the resulting false condemnations "are especially costly, because they chill the very conduct the antitrust laws are designed to protect." Matsushita Elec. Industrial Co. v. Zenith Radio Corp., 475 U.S. 574, 594 (1986). The cost of false positives counsels against an undue expansion of §2 liability.

* * *

The 1996 Act is, in an important respect, much more ambitious than the antitrust laws. It attempts "*to eliminate the monopolies* enjoyed by the inheritors of AT&T's local franchises." *Verizon Communications Inc. v. FCC,* 535 U.S. [467,] 476 (2002) (emphasis added). Section 2 of the Sherman Act, by contrast, seeks merely to prevent *unlawful monopolization.* It would be a serious mistake to conflate the two goals. The Sherman Act is indeed the "Magna Carta of free enterprise," *United States v. Topco Associates, Inc.,* 405 U.S. 596, 610 (1972), but it does not give judges *carte blanche* to insist that a monopolist alter its way of doing business whenever some other approach might yield greater competition. We conclude that respondent's complaint fails to state a claim under the Sherman Act.

Notes and Questions

1. The Benefits of Monopolies: As noted in the case, monopolies are not illegal in the United States. Section 2 of the Sherman Antitrust Act makes illegal the acquisition or maintenance of monopoly power through anticompetitive means. Often it is difficult to distinguish between aggressive market competition and anticompetitive behavior. Monopolies are allowed in the United States because it is important to provide a carrot to encourage businesses to innovate by finding more efficient ways of operating and developing better products. When these better methods or processes are developed, companies should benefit from the ability to charge monopoly prices and earn higher profits. Accordingly, the quest for monopoly profits drives innovation and technological development.

2. Error Costs from Aggressive Antitrust Enforcement: One major concern noted by the court in *Trinko* is that in heavily regulated industries, the goals of Congressional statutory schemes may diverge from the antitrust laws. The Telecommunications Act, in that case, sought to end an artificially created government monopoly in favor of more robust competition. The Act forced Verizon to allow its competitors to use its infrastructure in order

to facilitate new competitors. The court concluded that although Congress imposed this burden on Verizon, the antirust laws do not impose such an obligation to deal with rivals. The court was concerned that the costs of imposing a bad antitrust rule would outweigh the benefits. This approach has been summarized by Judge Frank Easterbrook in his seminal article on error costs. He writes that:

> A fundamental difficulty facing the court is the incommensurability of the stakes. If the court errs by condemning a beneficial practice, the benefits may be lost for good. Any other firm that uses the condemned practice faces sanctions in the name of stare decisis, no matter the benefits. If the court errs by permitting a deleterious practice, though, the welfare loss decreases over time. Monopoly is self-destructive. Monopoly prices eventually attract entry. True, this long run may be a long time coming, with loss to society in the interim. The central purpose of antitrust is to speed up the arrival of the long run. But this should not obscure the point: judicial errors that tolerate baleful practices are self-correcting while erroneous condemnations are not.

Frank H. Easterbrook, *The Limits of Antitrust*, 63 Tex. L. Rev. 1, 2–3 (1984)

3. Sociopolitical Arguments Against Monopoly: In addition to the economic criticisms of monopoly, there are several sociopolitical arguments against monopoly. The following excerpt from Professor (now Judge) Richard Posner's *Antitrust Law: An Economic Perspective* summarizes many of these criticisms and their validity.

> Having considered the economic objections to monopoly, I want to discuss now three broadly political arguments against it. The first is that monopoly transfers wealth from consumers to the stockholders of monopolistic firms, a redistribution that goes from the less to the more wealthy. This appealing argument is undermined by the point made earlier that competition to become a monopolist will tend to transform the expected gains from monopoly into social costs. To the extent that this occurs, consumers' wealth will not be transferred to the shareholders of monopoly firms but will instead be dissipated in the purchase of inputs into the activity of becoming a monopolist.
>
> A second argument is that monopoly, or more broadly any condition (such as concentration) that fosters cooperation among competing firms, will facilitate an industry's manipulation of the political process to obtain protective legislation aimed at increasing the industry's profits. Often such protection takes the form of controls over entry and price competition, coupled with exemption from the antitrust laws, that result in cartelizing the industry much more effectively than could be done by private agreement. This is not the place to pursue the intricacies of the nascent economic analysis of the determinants of political power. It is enough to note that, while concentration may reduce the costs of organizing effectively to manipulate the political process, it may also reduce the demand for public assistance in suppressing competition, since … a concentrated industry, other things being equal, is in a better position to suppress competition through private agreement, express or tacit, than an unconcentrated industry. It is therefore unclear whether on balance concentrated, or monopolistic, industries will obtain greater help from the political process than unconcentrated, or competitive, industries. This theoretical indeterminacy is mirrored in the empirical world, where we observe many unconcentrated industries—agriculture, trucking, local broadcasting, banking, medicine, to name a few—enjoying governmental protection against competition.

In any event, however, this political objection to monopoly and concentration is not sharply different from the economic objection. The legislation sought by an industry—a tariff, a tax on substitute products, control of entry—will usually have economic effects similar or even identical to those of a private cartel agreement. The political argument—which is simply that concentration facilitates monopoly pricing indirectly through the legislative process, as well as directly through cartelization—thus implies no change in the character of an antitrust policy deduced from economic considerations. (This is also true, incidentally, of the wealth-redistribution argument: the implications for public policy are not sharply different whether one objects to monopoly pricing because it wastes resources or because it brings about undesirable changes in the redistribution of wealth.)

The last political argument that I shall discuss has, in contrast, implications for antitrust policy that diverge sharply from those of economic analysis. The popular (or Populist) alternative to an antitrust policy designed to promote economic efficiency by limiting monopoly is a policy of restricting the freedom of action of large business firms in order to promote small business. (It may be possible to conceive of a different alternative to an efficiency-based antitrust policy, but this is the only one suggested with any frequency.) The idea that there is some special virtue in small business compared to large is a persistent one. I am not prepared to argue that it has no merit whatever. I am, however, confident that antitrust enforcement is an inappropriate method of trying to promote the interests of small business as a whole. The best overall antitrust policy from the standpoint of small business is *no* antitrust policy, since monopoly, by driving a wedge between the prices and the costs of the larger firms in the market (it is presumably they who take the lead in forming cartels), enable the smaller firms in the market to survive even if their costs are higher than those of the large firms. The only kind of antitrust policy that would benefit small business would be one whose principal objective was to limit the attempts of large firms to underprice less efficient small firms by sharing their lower costs with consumers in the form of lower prices. Apart from raising in acute form the question of whether it is socially desirable to promote small business at the expense of the consumer, such a policy would be unworkable because it would require comprehensive and continuing supervision of the prices of large firms.... The tools of antitrust enforcement are poorly designed for effective discrimination in favor of small firms, compared, for example, to the effectiveness of taxing larger firms at higher rates. We shall have frequent occasion in this book to remark how difficult it is to press the antitrust laws into the service of small business. The realistic choice is between shaping antitrust policy in accordance with the economic (and congruent political) objections to monopoly and—if we think that limiting big business and promoting small is more important than efficiency—abandoning it.

Richard A. Posner, Antitrust Law: An Economic Perspective 18–20 (1976).

3. More Realistic Models of Competition: Differentiated Products Markets & Oligopoly

The perfect competition model is an interesting and important tool in economic analysis. Competition in the real world, however, is characterized by a wide variety of **rivalrous**

behavior. Competitors engage in advertising, price-cutting, and quality improvement programs. Consider the following observation by Professor McNulty regarding the dangers of confusing the model of perfect competition with the meaning of competition:

> [T]he most general tendency concerning the meaning of competition in economic theory is to regard it as the opposite of monopoly. An unfortunate result of this way of thinking has been no little confusion concerning the relationship between economic efficiency and business behavior. There is a striking contrast in economic literature between the analytical rigor and precision of competition when it is described as a market structure, and the ambiguity surrounding the ideas of competition whenever it is discussed in behavioral terms. Since, as Hayek has rightly noted, "the law cannot effectively prohibit states of affairs but only kinds of action," a concept of economic competition, if it is to be significant for economic policy, ought to relate to patterns of business behavior such as might reasonably be associated with the verb "to compete." That was the case with the competition which Adam Smith made the central organizing principle of economic society in the *Wealth of Nations....* Whether it was seen as price undercutting by sellers, the bidding up of prices by buyers, or the entry of new firms into profitable industries, the fact is that competition entered economics as a concept which had empirical relevance and operational meaning in terms of contemporary business behavior. Yet on the question of whether such common current practices as advertising, product variation, price undercutting, or other forms of business activity do or do not constitute competition, modern economic theory offers the clarification that they are "monopolistically" competitive. While this is a useful way of illustrating the truth that most markets are in some degree both controlled and controlling, it is less useful as a guide in implementing a policy, such as our antitrust policy, which seeks at once to restrain monopoly and promote competition.

McNulty, *Economic Theory and the Meaning of Competition*, 82 Quarterly Journal of Economics 639, 639–40 (1968). Thus, in reality, perfect competition rarely, if ever, exists. Instead, most firms participate in markets that lie on the spectrum somewhere between perfect competition and monopoly. The following sections discuss two such markets: differentiated products markets and oligopoly.

a. Competition with Differentiated Products

The perfect competition model assumes homogenous consumers and firms. In reality, however, consumers have different tastes and preferences and some firms enjoy lower costs structures than others or are more innovative. **Differentiated products** have distinguishing features relative to the goods or services they are in competition with, like McDonalds and the other fast food restaurants. Differentiation can occur across different dimensions. Product quality can be an important source of product differentiation. Quality differences can be measured in objective dimensions — for example, first class vs. coach air travel — or in subjective dimensions — for example, Coca-Cola vs. Pepsi. Another common source of product differentiation is customer service.

Another important distinction between the real world and the perfectly competitive market involves information. The perfect competition model assumes perfect information is available at zero cost. But, in reality, information is scarce and costly to obtain. The market process approach stresses that economic decision making in the real world is a "discovery process." The discovery process takes place under conditions of ignorance far removed from the perfect knowledge assumption in the model of perfect competition.

In the model of perfect competition, it was assumed that all firms produced an identical product. Thus, perfectly competitive firms face a totally-elastic, or horizontal demand curve, making them price takers. Relaxing the assumptions of perfect information and homogeneity, means that firms face downward sloping demand curve. That is, the firm with a differentiated product has some degree of price control. In general, the more (less) elastic the demand curve, the more (less) similar are competitor's products, and the less (more) control competitors have over price. As always, a firm maximizes it profits by setting price where its marginal cost equals marginal revenue. Because a firm's demand curve slopes downward, however, the price charged will be greater than marginal cost. The gap between marginal cost and price is largely dependent upon the availability of substitutes.

Just because price is greater than marginal cost, however, does not mean that firms earn economic profits in the long run. As in the model of perfect competition, entry and exit are relatively easy in differentiated products market, which means that new firms will be attracted to profitable product spaces. Consider the following example. Suppose Henry lives in a community that had always been exclusively zoned residential. Then, the town council changes the zoning to commercial and Henry decides to open a new fast food restaurant, Henry's Hamburgers. The restaurant is immediately successful and Henry is able to charge high enough prices to make significantly more money than he would using his resources in other ways; in economic parlance, Henry earns economic profits. National chains observe Henry's success and open six new fast food restaurants—KFC, McDonalds, Five Guys, Chipotle, Taco Bell, and Panda Express. These restaurants provide slightly distinguished services and different food, but are certainly close substitutes to Henry's Hamburgers. As each restaurant opens, Henry realizes that demand decreases for his hamburger and he must decrease his prices to keep selling hamburgers. Soon enough, Henry's restaurant makes the same amount as he could if he allocated his resources in other ways—he and the others experience zero economic profits.

Three additional entrepreneurs mistakenly think that Henry's Hamburgers and the other restaurants are still earning positive economic profits. Consequently they each build new fast food restaurants in Henry's community. This mistake is possible and unfortunately frequent in the dynamic process because firms do not have perfect information as assumed in perfectly competitive markets. As these firms enter the market, demand continues to fall for Henry's Hamburgers and the others and they begin to operate at a loss. This means that these firms could increase their profits by using their resources in other ways. Consequently, firms begin to exit the fast food industry in Henry's community. As each firm exits, demand for the remaining firms increases because there are fewer substitutes available. Firms will exit until economic profits are back to zero again.

One characteristic of firms that produce differentiated products is the use of **advertising**. Advertising is a means of communicating both the existence of alternate sellers as well as the distinguishing features of a product. Advertising is a critical and necessary condition for dynamic competitive markets. For example, new firms need to make consumers aware of their products' existence, location, and characteristics. Without advertising, it would be very difficult for new entrants to challenge an established firm. The greater consumer awareness of product alternatives, availability, prices, and other characteristics, the less the consumer is bound to any one seller. Economists have argued that the provision of product information contained in advertising lowers the total search costs to consumers. Moreover, empirical studies have shown that continued advertising improves product

quality, maximizes firm profitability, and decreases price. Thus, it seems clear that consumers are the prime beneficiaries of advertising.

Production and sale of goods and services is a complicated and competitive process. Specialization creates economic efficiencies derived from dividing the manufacturing, distribution, and retail sales functions among separate firms. **Vertical relationships** are arrangements down the chain of distribution — from input suppliers to manufacturers to wholesale distributors to retailers. A typical vertical relationship is organized in the following way. A bicycle manufacturer recognizes that certain efficiencies can be gained by selling its product to wholesale distributors (who, in turn, sell the bikes to local retailers) rather than performing all of these functions itself.

For the manufacturer, however, there is a fundamental danger in this type of vertical relationship. Specifically, the bicycle manufacturer wants to sell the profit maximizing quantity of bikes; however, it does not control the retail sales function. The manufacturer wants its retailers to engage in aggressive competitive behavior geared (no pun intended) toward selling its bikes. Retailers can take several actions that affect the sales of the manufacturer's bikes. Obviously, the price set by the retailer will affect sales. Further, the retailer can promote the manufacturer's bikes on its web site, it can place the bikes prominently in the store, and train employees to provide information about the bikes or to promote the bikes over other brands.

Because retailers and producers are different entities, however, they have different objectives; retailers do not necessarily have the same incentive — a desire to maximize the sale of the manufacturer's product. As a result, manufacturers have resorted to numerous contractual devices to encourage their retailers to promote their products, for example, limiting retail pricing discretion or requiring certain types of point-of-sale service. The case below involves an antitrust challenge to once such contractual limitation — **resale price maintenance**.

Leegin Creative Leather Products, Inc. v. PSKS, Inc.

Supreme Court of the United States
551 U.S. 877 (2007)

JUSTICE KENNEDY delivered the opinion of the Court

In Dr. Miles Medical Co. v. John D. Park & Sons Co., 220 U.S. 373 (1911), the Court established the rule that it is *per se* illegal under §1 of the Sherman Act for a manufacturer to agree with its distributor to set the minimum price the distributor can charge for the manufacturer's goods. The question presented by the instant case is whether the Court should overrule the *per se* rule and allow resale price maintenance agreements to be judged by the rule of reason, the usual standard applied to determine if there is a violation of §1. The Court has abandoned the rule of *per se* illegality for other vertical restraints a manufacturer imposes on its distributors. Respected economic analysts, furthermore, conclude that vertical price restraints can have procompetitive effects. We now hold that *Dr. Miles* should be overruled and that vertical price restraints are to be judged by the rule of reason.

I

Petitioner, Leegin Creative Leather Products, Inc. (Leegin), designs, manufactures, and distributes leather goods and accessories. In 1991, Leegin began to sell belts under the brand name "Brighton." The Brighton brand has now expanded into a variety of women's fashion accessories. It is sold across the United States in over 5,000 retail es-

tablishments, for the most part independent, small boutiques and specialty stores. Leegin's president, Jerry Kohl, also has an interest in about 70 stores that sell Brighton products. Leegin asserts that, at least for its products, small retailers treat customers better, provide customers more services, and make their shopping experience more satisfactory than do larger, often impersonal retailers. Kohl explained: "[W]e want the consumers to get a different experience than they get in Sam's Club or in Wal-Mart. And you can't get that kind of experience or support or customer service from a store like Wal-Mart."

Respondent, PSKS, Inc. (PSKS), operates Kay's Kloset, a women's apparel store in Lewisville, Texas. Kay's Kloset buys from about 75 different manufacturers and at one time sold the Brighton brand. It first started purchasing Brighton goods from Leegin in 1995. Once it began selling the brand, the store promoted Brighton. For example, it ran Brighton advertisements and had Brighton days in the store. Kay's Kloset became the destination retailer in the area to buy Brighton products. Brighton was the store's most important brand and once accounted for 40 to 50 percent of its profits.

In 1997, Leegin instituted the "Brighton Retail Pricing and Promotion Policy." Following the policy, Leegin refused to sell to retailers that discounted Brighton goods below suggested prices.

* * *

Leegin adopted the policy to give its retailers sufficient margins to provide customers the service central to its distribution strategy. It also expressed concern that discounting harmed Brighton's brand image and reputation.

* * *

In December 2002, Leegin discovered Kay's Kloset had been marking down Brighton's entire line by 20 percent. Kay's Kloset contended it placed Brighton products on sale to compete with nearby retailers who also were undercutting Leegin's suggested prices. Leegin, nonetheless, requested that Kay's Kloset cease discounting. Its request refused, Leegin stopped selling to the store. The loss of the Brighton brand had a considerable negative impact on the store's revenue from sales.

PSKS sued Leegin in the United States District Court for the Eastern District of Texas. It alleged, among other claims, that Leegin had violated the antitrust laws by "enter[ing] into agreements with retailers to charge only those prices fixed by Leegin." Leegin planned to introduce expert testimony describing the procompetitive effects of its pricing policy. The District Court excluded the testimony, relying on the *per se* rule established by *Dr. Miles*.

* * *

The jury agreed with PSKS and awarded it $1.2 million. Pursuant to 15 U.S.C. §15(a), the District Court trebled the damages and reimbursed PSKS for its attorney's fees and costs. It entered judgment against Leegin in the amount of $3,975,000.80.

The Court of Appeals for the Fifth Circuit affirmed. On appeal Leegin did not dispute that it had entered into vertical price-fixing agreements with its retailers. Rather, it contended that the rule of reason should have applied to those agreements.

* * *

We granted certiorari to determine whether vertical minimum resale price maintenance agreements should continue to be treated as *per se* unlawful.

* * *

A

Though each side of the debate can find sources to support its position, it suffices to say here that economics literature is replete with procompetitive justifications for a manufacturer's use of resale price maintenance. See, *e.g.,* Brief for Economists as *Amici Curiae* 16 ("In the theoretical literature, it is essentially undisputed that minimum [resale price maintenance] can have procompetitive effects and that under a variety of market conditions it is unlikely to have anticompetitive effects"); Brief for United States as *Amicus Curiae* 9 ("[T]here is a widespread consensus that permitting a manufacturer to control the price at which its goods are sold may promote *inter*brand competition and consumer welfare in a variety of ways"); ABA Section of Antitrust Law, Antitrust Law and Economics of Product Distribution 76 (2006) ("[T]he bulk of the economic literature on [resale price maintenance] suggests that [it] is more likely to be used to enhance efficiency than for anticompetitive purposes"); *see also* H. Hovenkamp, The Antitrust Enterprise: Principle and Execution 184–191 (2005) (hereinafter Hovenkamp); R. Bork, The Antitrust Paradox 288–291 (1978) (hereinafter Bork). Even those more skeptical of resale price maintenance acknowledge it can have procompetitive effects. See, *e.g.,* Brief for William S. Comanor et al. as *Amici Curiae* 3 ("[G]iven [the] diversity of effects [of resale price maintenance], one could reasonably take the position that a *rule of reason* rather than a *per se* approach is warranted"); F. Scherer & D. Ross, Industrial Market Structure and Economic Performance 558 (3d ed.1990) (hereinafter Scherer & Ross) ("The overall balance between benefits and costs [of resale price maintenance] is probably close").

* * *

The justifications for vertical price restraints are similar to those for other vertical restraints. Minimum resale price maintenance can stimulate interbrand competition—the competition among manufacturers selling different brands of the same type of product—by reducing intrabrand competition—the competition among retailers selling the same brand. The promotion of interbrand competition is important because "the primary purpose of the antitrust laws is to protect [this type of] competition." [*State Oil Co. v.*] *Khan,* 522 U.S. [3,] 15 [(1997)]. A single manufacturer's use of vertical price restraints tends to eliminate intrabrand price competition; this in turn encourages retailers to invest in tangible or intangible services or promotional efforts that aid the manufacturer's position as against rival manufacturers. Resale price maintenance also has the potential to give consumers more options so that they can choose among low-price, low-service brands; high-price, high-service brands; and brands that fall in between.

Absent vertical price restraints, the retail services that enhance interbrand competition might be underprovided. This is because discounting retailers can free ride on retailers who furnish services and then capture some of the increased demand those services generate. Consumers might learn, for example, about the benefits of a manufacturer's product from a retailer that invests in fine showrooms, offers product demonstrations, or hires and trains knowledgeable employees. R. Posner, Antitrust Law 172–173 (2d ed.2001) (hereinafter Posner). Or consumers might decide to buy the product because they see it in a retail establishment that has a reputation for selling high-quality merchandise. Marvel & McCafferty, *Resale Price Maintenance and Quality Certification*, 15 Rand J. Econ. 346, 347–349 (1984). If the consumer can then buy the product from a retailer that discounts because it has not spent capital providing services or developing a quality reputation, the high-service retailer will lose sales to the discounter, forcing it to cut back its services to a level lower than consumers would otherwise prefer. Minimum resale price maintenance alleviates the problem because it prevents the discounter from undercutting

the service provider. With price competition decreased, the manufacturer's retailers compete among themselves over services.

* * *

Resale price maintenance can also increase interbrand competition by encouraging retailer services that would not be provided even absent free riding. It may be difficult and inefficient for a manufacturer to make and enforce a contract with a retailer specifying the different services the retailer must perform. Offering the retailer a guaranteed margin and threatening termination if it does not live up to expectations may be the most efficient way to expand the manufacturer's market share by inducing the retailer's performance and allowing it to use its own initiative and experience in providing valuable services. See Mathewson & Winter, *The Law and Economics of Resale Price Maintenance*, 13 Rev. Indus. Org. 57, 74–75 (1998); Klein & Murphy, *Vertical Restraints as Contract Enforcement Mechanisms*, 31 J. Law & Econ. 265, 295 (1988); see also Deneckere, Marvel, & Peck, *Demand Uncertainty, Inventories, and Resale Price Maintenance*, 111 Q.J. Econ. 885, 911 (1996) (noting that resale price maintenance may be beneficial to motivate retailers to stock adequate inventories of a manufacturer's goods in the face of uncertain consumer demand).

* * *

While vertical agreements setting minimum resale prices can have procompetitive justifications, they may have anticompetitive effects in other cases; and unlawful price fixing, designed solely to obtain monopoly profits, is an ever–present temptation. Resale price maintenance may, for example, facilitate a manufacturer cartel. An unlawful cartel will seek to discover if some manufacturers are undercutting the cartel's fixed prices. Resale price maintenance could assist the cartel in identifying price-cutting manufacturers who benefit from the lower prices they offer. Resale price maintenance, furthermore, could discourage a manufacturer from cutting prices to retailers with the concomitant benefit of cheaper prices to consumers.

Vertical price restraints also "might be used to organize cartels at the retailer level." *Business Electronics* [*Corp. v. Sharp Electronics Corp*., 485 U.S. 717,]*,* 725–726 [(1988)]..A group of retailers might collude to fix prices to consumers and then compel a manufacturer to aid the unlawful arrangement with resale price maintenance. In that instance the manufacturer does not establish the practice to stimulate services or to promote its brand but to give inefficient retailers higher profits. Retailers with better distribution systems and lower cost structures would be prevented from charging lower prices by the agreement. Historical examples suggest this possibility is a legitimate concern. See, *e.g.,* Marvel & McCafferty, *The Welfare Effects of Resale Price Maintenance*, 28 J. Law & Econ. 363, 373 (1985) (providing an example of the power of the National Association of Retail Druggists to compel manufacturers to use resale price maintenance); Hovenkamp 186 (suggesting that the retail druggists in *Dr. Miles* formed a cartel and used manufacturers to enforce it).

* * *

Resale price maintenance, it is true, does have economic dangers. If the rule of reason were to apply to vertical price restraints, courts would have to be diligent in eliminating their anticompetitive uses from the market. This is a realistic objective, and certain factors are relevant to the inquiry. For example, the number of manufacturers that make use of the practice in a given industry can provide important instruction. When only a few manufacturers lacking market power adopt the practice, there is little likelihood it is facilitating a manufacturer cartel, for a cartel then can be undercut by rival manufacturers. Likewise, a retailer cartel is unlikely when only a single manufacturer in a competitive market uses

resale price maintenance. Interbrand competition would divert consumers to lower priced substitutes and eliminate any gains to retailers from their price-fixing agreement over a single brand. See Posner 172; Bork 292. Resale price maintenance should be subject to more careful scrutiny, by contrast, if many competing manufacturers adopt the practice. Cf. Scherer & Ross 558 (noting that "except when [resale price maintenance] spreads to cover the bulk of an industry's output, depriving consumers of a meaningful choice between high-service and low-price outlets, most [resale price maintenance arrangements] are probably innocuous"); Easterbrook 162 (suggesting that "every one of the potentially-anticompetitive outcomes of vertical arrangements depends on the uniformity of the practice").

The source of the restraint may also be an important consideration. If there is evidence retailers were the impetus for a vertical price restraint, there is a greater likelihood that the restraint facilitates a retailer cartel or supports a dominant, inefficient retailer. See Brief for William S. Comanor et al. as *Amici Curiae* 7–8. If, by contrast, a manufacturer adopted the policy independent of retailer pressure, the restraint is less likely to promote anticompetitive conduct. Cf. Posner 177 ("It makes all the difference whether minimum retail prices are imposed by the manufacturer in order to evoke point-of-sale services or by the dealers in order to obtain monopoly profits"). A manufacturer also has an incentive to protest inefficient retailer-induced price restraints because they can harm its competitive position.

As a final matter, that a dominant manufacturer or retailer can abuse resale price maintenance for anticompetitive purposes may not be a serious concern unless the relevant entity has market power. If a retailer lacks market power, manufacturers likely can sell their goods through rival retailers. See also *Business Electronics, supra*, at 727, n. 2 (noting "[r]etail market power is rare, because of the usual presence of interbrand competition and other dealers"). And if a manufacturer lacks market power, there is less likelihood it can use the practice to keep competitors away from distribution outlets.

* * *

The manufacturer has a number of legitimate options to achieve benefits similar to those provided by vertical price restraints. A manufacturer can exercise its *Colgate* right to refuse to deal with retailers that do not follow its suggested prices. The economic effects of unilateral and concerted price setting are in general the same. The problem for the manufacturer is that a jury might conclude its unilateral policy was really a vertical agreement, subjecting it to treble damages and potential criminal liability. Even with the stringent standards in *Monsanto* and *Business Electronics,* this danger can lead, and has led, rational manufacturers to take wasteful measures. A manufacturer might refuse to discuss its pricing policy with its distributors except through counsel knowledgeable of the subtle intricacies of the law. Or it might terminate longstanding distributors for minor violations without seeking an explanation. The increased costs these burdensome measures generate flow to consumers in the form of higher prices.

Furthermore, depending on the type of product it sells, a manufacturer might be able to achieve the procompetitive benefits of resale price maintenance by integrating downstream and selling its products directly to consumers. *Dr. Miles* tilts the relative costs of vertical integration and vertical agreement by making the former more attractive based on the *per se* rule, not on real market conditions. [S]ee generally Coase, The Nature of the Firm, 4 Economica, New Series 386 (1937). This distortion might lead to inefficient integration that would not otherwise take place, so that consumers must again suffer the consequences of the suboptimal distribution strategy. And integration, unlike vertical price restraints, eliminates all intrabrand competition.

There is yet another consideration. A manufacturer can impose territorial restrictions on distributors and allow only one distributor to sell its goods in a given region. Our cases have recognized, and the economics literature confirms, that these vertical nonprice restraints have impacts similar to those of vertical price restraints; both reduce intrabrand competition and can stimulate retailer services.... The same legal standard (*per se* unlawfulness) applies to horizontal market division and horizontal price fixing because both have similar economic effect. There is likewise little economic justification for the current differential treatment of vertical price and nonprice restraints. Furthermore, vertical nonprice restraints may prove less efficient for inducing desired services, and they reduce intrabrand competition more than vertical price restraints by eliminating both price and service competition. See Brief for Economists as *Amici Curiae* 17–18.

* * *

For these reasons the Court's decision in Dr. Miles Medical Co. v. John D. Park & Sons Co., 220 U.S. 373 (1911), is now overruled. Vertical price restraints are to be judged according to the rule of reason.

V

* * *

The judgment of the Court of Appeals is reversed, and the case is remanded for proceedings consistent with this opinion.

Notes and Questions

1. Resale Price Maintenance and the Leegin Factors: The business conduct at issue in this case was minimum resale price maintenance (RPM). RPM occurs where a manufacturer demands that retailers only sell the manufacturer's product above an established price floor. As evidenced in the opinion, the Court gave substantial weight to academic literature claiming that this practice carries with it both a possibility for anticompetitive harm to markets and procompetitive benefits. The Court rejected the previously standing 100 year rule of *per se* condemnation of RPM and replaced it with the rule of reason. The Court placed the onus of further development of the rule of reason on the district courts and advised them to be "diligent in eliminating their anticompetitive uses from the market." Justice Kennedy points to three factors as guidance for when to be concerned about the anticompetitive harms of minimum RPM. Those three factors include: (1) when RPM covers the bulk of an industry's goods, (2) when a single dominant retailer, rather than manufacturers, implements or encourages RPM, and (3) when RPM is instituted by firms with discernible market power. Do these three factors, purportedly increasing the dangers of RPM, make economic sense? How should a court go about weighing the magnitude of economic harm and procompetitive benefit? Is evidence of what the court calls free-riding by low-cost retailers on the customer service of higher cost retailers, such as showrooms, product demonstrations, and knowledgeable employees, dispositive evidence of economic harm?

2. Intrabrand v. Interbrand Rivalry: The Court recognizes a distinction between intrabrand and interbrand competition. Should these two different types of competition be analyzed in the same manner for purposes of antitrust regulation? The Court observes that some vertical nonprice restraints on intrabrand competition may actually improve interbrand competition. How does this work? Does this also mean that the losses in efficiency resulting from restraints in intrabrand competition are made up for by improved

interbrand competition? At least one prominent commentator is quite skeptical of the balancing the court seems to support. Judge Frank Easterbrook has written that:

> No one can sensibly weigh inter- and intrabrand competition against one another; they are not commensurable. The reduction in "intrabrand competition" is the *source* of the competitive benefit that helps one product compete against another. Intrabrand competition as such is worthless; one might as well complain when a corporation does not have internal competition to make the product most cheaply. Vertical integration eliminates this form of "competition," but in so doing it may enable the manufacturer to reduce its delivered price. No manufacturer wants to have less competition among its dealers for the sake of less competition. The reduction in dealers' rivalry in the price dimension is just the tool the manufacturer uses to induce greater competition in the service dimension. As I spelled out above, restricted dealing alters the product's attributes. There is no "less" in one column to "balance against a gain" in the other, any more than the manufacturer's sole prerogative to decide what physical product to make creates a "reduction in intrabrand competition."

Frank H. Easterbrook, *Vertical Arrangements and the Rule of Reason*, 53 Antitrust L.J. 135 (1984).

3. Development of the Rule of Reason in Vertical Restraint: Justice Kennedy's opinion recognized that the Court had slowly but surely been moving in the direction of handling most business conduct under the rule of reason rather than condemning the same conduct under a rule of *per se* illegality. *Leegin* was foretold by a major shift in the court's approach to nearly all vertical restraints. The Court cited *GTE Sylvania* for the proposition that non-price vertical restraints should be evaluated under a rule of reason. In that case, a manufacturer of television sets used vertical restraints in the form of location-specific exclusive contracts. The agreements only allowed the retailer to sell the product at specific stores specified in the agreement. One retailer, despite agreeing to this term and having notable success with the sale of Sylvania TVs, sued under the antitrust laws because it wanted to sell this product in other geographical regions disallowed under the contract. The Supreme Court, in an argument similar to the one adopted in *Leegin*, highlighted the possibility for vertical restrictions on goods to increase interbrand competition and allow smaller manufacturers to compete. It reasoned that:

> Vertical restrictions reduce intrabrand competition by limiting the number of sellers of a particular product competing for the business of a given group of buyers. Location restrictions have this effect because of practical constraints on the effective marketing area of retail outlets. Although intrabrand competition may be reduced, the ability of retailers to exploit the resulting market may be limited both by the ability of consumers to travel to other franchised locations and, perhaps more importantly, to purchase the competing products of other manufacturers. None of these key variables, however, is affected by the form of the transaction by which a manufacturer conveys his products to the retailers.

> Vertical restrictions promote interbrand competition by allowing the manufacturer to achieve certain efficiencies in the distribution of his products. These "redeeming virtues" are implicit in every decision sustaining vertical restrictions under the rule of reason. Economists have identified a number of ways in which manufacturers can use such restrictions to compete more effectively against other manufacturers. See, e. g., Preston, *Restrictive Distribution Arrangements: Economic Analysis and Public Policy Standards*, 30 Law & Contemp. Prob. 506, 511 (1965).

For example, new manufacturers and manufacturers entering new markets can use the restrictions in order to induce competent and aggressive retailers to make the kind of investment of capital and labor that is often required in the distribution of products unknown to the consumer. Established manufacturers can use them to induce retailers to engage in promotional activities or to provide service and repair facilities necessary to the efficient marketing of their products. Service and repair are vital for many products, such as automobiles and major household appliances. The availability and quality of such services affect a manufacturer's goodwill and the competitiveness of his product. Because of market imperfections such as the so-called "free rider" effect, these services might not be provided by retailers in a purely competitive situation, despite the fact that each retailer's benefit would be greater if all provided the services than if none did.

Economists also have argued that manufacturers have an economic interest in maintaining as much intrabrand competition as is consistent with the efficient distribution of their products. Bork, *The Rule of Reason and the Per Se Concept: Price Fixing and the Market Division (II)*, 75 Yale L.J. 373, 403 (1966). Although the view that the manufacturer's interest necessarily corresponds with that of the public is not universally shared, even the leading critic of vertical restrictions concedes that *Schwinn*'s distinction between sale and nonsale transactions is essentially unrelated to any relevant economic impact. Comanor, *Vertical Territorial and Customer Restrictions:* White Motor *and Its Aftermath*, 81 Harv. L. Rev. 1419, 1422 (1968).25 Indeed, to the extent that the form of the transaction is related to interbrand benefits, the Court's distinction is inconsistent with its articulated concern for the ability of smaller firms to compete effectively with larger ones. Capital requirements and administrative expenses may prevent smaller firms from using the exception for nonsale transactions. See, e. g., Phillips, Schwinn *Rules and the "New Economics" of Vertical Relation*, 44 Antitrust L.J. 573, 576 (1975).

* * *

In sum, we conclude that the appropriate decision is to return to the rule of reason that governed vertical restrictions prior to Schwinn. When anticompetitive effects are shown to result from particular vertical restrictions they can be adequately policed under the rule of reason, the standard traditionally applied for the majority of anticompetitive practices challenged under § 1 of the Act. Accordingly, the decision of the Court of Appeals is Affirmed.

Cont'l T. V., Inc. v. GTE Sylvania Inc., 433 U.S. 36, 54–59 (1977).

Another shift in the court's approach to vertical restraints came in *State Oil v. Khan*. There, a gasoline supplier terminated an agreement with an owner of a gas station who promptly sued alleging that State Oil's policies mandated a maximum resale price that should be condemned as a per se violation of the antitrust laws. The agreement provided the gas station owner with a 3.25 cent margin on the purchase price of every gallon of gasoline. If the gas station charged any price higher than the supplier's suggested retail price the extra money had to be rebated back to State Oil. In dismissing the application of the *per se* rule, a unanimous Court affirmed the reasoning, although not the decision, of Judge Posner in the Seventh Circuit. There, Judge Posner applied the *per se* rule in accordance with *stare decisis* but not without analyzing the answer to the question:

Why might competitors fix a maximum resale price? The difference between what a supplier charges his dealer and what the dealer charges the ultimate

customer is, functionally, compensation to the dealer for performing the resale service; so by agreeing on the resale prices of their goods competing sellers can reduce their dealers' margin below the competitive price for the dealers' service. This is a form of monopsony pricing, which is analytically the same as monopoly or cartel pricing and so treated by the law. E.g., Mandeville Island Farms, Inc. v. American Crystal Sugar Co., 334 U.S. 219 (1948).

The questionable next step (logically, not chronologically, next) in the evolution of antitrust law was to affix the per se label to contracts in which a single supplier, not acting in concert with any of its competitors, fixed its dealers' retail prices. Dr. Miles Medical Co. v. John D. Park & Sons Co., 220 U.S. 373 (1911). Here the economic difference between fixing a minimum resale price and fixing a maximum resale price becomes more pronounced, although most economists believe that neither form of price fixing is pernicious when the supplier is neither the cat's paw of colluding distributors nor acting in concert with his competitors. A supplier acting unilaterally might fix a minimum resale price in order to induce his dealers to furnish valuable point-of-sale services (trained salesmen, clean restrooms—whatever) to customers, which they could not afford to do without a guaranteed margin to cover the costs of the services, because the customers would use the services provided by the full-service dealers but then purchase the product from a competing dealer who could sell the product at a discount because he had not borne the expense of providing the services. Lester G. Telser, *Why Should Manufacturers Want Fair Trade?*, 3 J. Law & Econ. 86 (1962).

As for maximum resale price fixing, unless the supplier is a monopsonist he cannot squeeze his dealers' margins below a competitive level; the attempt to do so would just drive the dealers into the arms of a competing supplier. A supplier might, however, fix a maximum resale price in order to prevent his dealers from exploiting a monopoly position. We do not know anything about the competitive environment in which Khan and State Oil operate—which is why the district judge was right to conclude that if the rule of reason is applicable, Khan loses. But suppose that State Oil, perhaps to encourage the dealer services that we mentioned, has spaced its dealers sufficiently far apart to limit competition among them (or even given each of them an exclusive territory); and suppose further that Union 76 is a sufficiently distinctive and popular brand to give the dealers in it at least a modicum of monopoly power. Then State Oil might want to place a ceiling on the dealers' resale prices in order to prevent them from exploiting that monopoly power fully. It would do this not out of disinterested malice, but in its commercial self-interest. The higher the price at which gasoline is resold, the smaller the volume sold, and so the lower the profit to the supplier if the higher profit per gallon at the higher price is being snared by the dealer.

Khan v. State Oil Co., 93 F.3d 1358, 1361–62 (7th Cir. 1996) *vacated*, 522 U.S. 3 (1997).

4. A Spirited Dissent: There was a four-Justice dissent in the *Leegin* opinion written by Justice Breyer that stressed the value of *stare decisis* and the limited authority of the court to overrule well-established decisions. Arguing that antitrust policy required stable and predictable rules he contended that "the fact that a rule of law has become 'embedded' in our 'national culture' argues strongly against overruling. The *per se* rule forbidding minimum resale price maintenance agreements has long been 'embedded' in the law of antitrust. It involves price, the economy's 'central nervous system.' It reflects a basic

antitrust assumption (that consumers often prefer lower prices to more service). It embodies a basic antitrust objective (providing consumers with a free choice about such matters). And it creates an easily administered and enforceable bright line, 'Do not agree about price,' that businesses as well as lawyers have long understood." Leegin Creative Leather Products, Inc. v. PSKS, Inc., 551 U.S. 877, 926–27 (2007) (citations omitted).

Is predictability a good reason to keep old precedent? What do you think happens in the business world when the Supreme Court changes a long-established rule that substantially affects the way that corporations compete? The Supreme Court frequently overrules itself. Why should this opinion be any different?

5. Manufacturers' Interest in Retailer Service and Prices: For minimum resale price maintenance, a manufacturer conditions a retailer's right to sell the manufacturer's products on the retailer's agreement not to sell the product below a certain price. This seems counterintuitive for the manufacturer because a retailer's margin over cost does not impact the manufacturer's profit. Put another way, a manufacturer sets its price regardless of what the retailer does so why does he care about the retailer's price? Some insight may be provided by a statement made by Leegin's president: "[W]e want the consumers to get a different experience than they get in Sam's Club or in Wal-Mart. And you can't get that kind of experience or support or customer service from a store like Wal-Mart." 551 U.S. at 882. Why does the president care how much service retailers provide? Are you convinced that high quality service benefits Leegin, as the manufacturer? If so, what is the logic?

6. Rule of Reason's Structure: The majority recognizes that there is a valid concern that minimum resale price maintenance is nothing but a ruse among retailers to cartelize the market. This valid concern, as elaborated by the dissent, is why most scholars promulgate a rule or reason for minimum resale price maintenance instead of a standard of per se legality. *But see* Richard A. Posner, *The Next Step in the Antitrust Treatment of Restricted Distribution: Per Se Legality*, 48 U. Chi. L. Rev. 6 (1981). *Leegin's* actual impact will not be fully established until courts determine how the rule of reason should apply. If the factfinder finds that the minimum resale price maintenance was imposed by the retailers, what does this say about the likely anticompetitive effect of the agreement? Is the analysis different if the factfinder finds that the manufacturer initiated the minimum resale price agreement? How should the answers to this question impact how courts should apply the rule of reason to these agreements? For more information on the potential structure for the rule of reason, see Thomas A. Lambert, Dr. Miles *Is Dead. Now What?: Structuring A Rule of Reason for Evaluating Minimum Resale Price Maintenance*, 50 Wm. & Mary L. Rev. 1937 (2009).

7. The Intersection of State and Federal Antitrust Law: Leegin's impact on firms' actual pricing practices is very difficult to predict given the complicated overlap of state and federal law in this area. Some state antitrust laws are identical to the Sherman Act, and precedent or statutes require the state court to follow federal precedent interpreting the Sherman Act. Those states presumably will follow *Leegin*. Other states, however, have codified the per se illegal standard for minimum resale price maintenance. In these states, a firm implementing resale price maintenance may not be liable under federal law but still will be liable under state law. See Michael A. Lindsay, *An Update on State RPM Laws Since* Leegin, Antitrust Source, Dec. 2010 (providing a table of all the states and how each treats minimum resale price maintenance). With this structure of laws, is it likely that national firms will even bother with resale price maintenance because they will essentially have a complicated patchwork of pricing schemes depending on state boundaries and statutes? If a national firm proceeds and implements a patchwork of pricing strategy even

though it is very complicated and costly to implement, does this lend support to the majority's view that minimum resale price maintenance is procompetitive?

8. Who Is the Loser: The majority argues that minimum resale price maintenance can alleviate the market of the free-rider problem:

> If the consumer can … buy the product from a retailer that discounts because it has not spent capital providing services or developing a quality reputation, the high-service retailer will lose sales to the discounter, forcing it to cut back its services to a level lower than consumers would otherwise prefer. Minimum resale price maintenance alleviates the problem because it prevents the discounter from undercutting the service provider. With price competition decreased, the manufacturer's retailers compete among themselves over services.

551 U.S. at 891. This argument is criticized because preventing the free-rider problem converts consumer welfare into producer welfare—certainly not the goal of antitrust. Critics argue it is better to know that consumers are getting lower prices than to hope that pricing schemes that increase prices somehow result in higher demand and increased total welfare. Does the knowledge that rational consumers will choose more consumer welfare over less resolve this concern since rational profit maximizing firms will choose a strategy that benefits consumers?

The most likely loser under the rule of reason is an individual that does not want point of service sales and is satisfied with little or no service and lower prices. Is the Court right to ignore this consumer in its analysis? Is this resolved by the difference in a per se illegal standard, the rule of reason standard, and a per se legal standard?

9. RPM and Increased Prices: Justice Breyer stated that "[m]ost economists today agree that, in the words of a prominent antitrust treatise, 'resale price maintenance tends to produce higher consumer prices than would otherwise be the case.'" 551 U.S. at 912. At least one empirical study attempted to measure *Leegin*'s impact on prices and found some evidence that prices did increase after *Leegin* was decided. Nathaniel J. Harris, Leegin's *Effect on Prices: An Empirical Analysis*, 9 J.L. Econ. & Pol'y 251 (2013). Is this price increase evidence dispositive and mean *Leegin* was incorrectly decided? If price increases, does consumer welfare—the ultimate standard for antitrust—necessarily have to decrease? If minimum resale price maintenance shifted the demand curve outward, would price increase? Would this shift necessarily increase consumer welfare?

10. Vertical Restraints and Evidence: Economists have continued to seek ways to use data, empirical methods, and quantitative approaches to test and confirm economic theories. Reviews of empirical literature seem to suggest that vertical restraints are associated with higher levels of output. For example, economists Francine Lafontaine and Margaret Slade concluded from their review that:

> Specifically, it appears that when manufacturers choose to impose such restraints, not only do they make themselves better off, but they also typically allow consumers to benefit from higher quality products and better service provision.... The evidence thus supports the conclusion that in these markets, manufacturer and consumer interests are apt to be aligned, while interference in the market is accomplished at the expense of consumers (and of course manufacturers).

For the details of their quantitative results, read Francine Lafontaine & Margaret Slade, *Exclusive Contracts and Vertical Restraints: Empirical Evidence and Public Policy*, 45 J. Econ

Lit. 629 (2005). A group of economists from the FTC reached a similar conclusion from their review of the empirical literature. *See* James C. Cooper, Luke M. Froeb, Dan O'Brien, & Michael G. Vita, *Vertical Antitrust Policy as a Problem of Inference*, 23 Int'l J. Indus. Org. 639 (2005).

b. Oligopoly

A market that consists of a few firms is called an oligopoly. Common oligopoly markets include wireless service providers (AT&T, Verizon, T-Mobile, and Sprint), commercial airline manufacturers (Boeing and Airbus), tennis ball manufacturers (Wilson, Penn, Dunlop, and Spalding), and accounting service providers (PWC, Deloitte, KPMG, and Ernst & Young). There are only a few firms in these markets because the optimal size of firm in these industry is very large (i.e., minimum ATC is only achieved when the firms are large) and the demand is not sufficient to support many large firms. The costs required for a firm to get to the point where minimum ATC is achieved and the fact that oligopoly markets can only sustain a limited number of firms serve as natural barriers to entry. Entry can also be limited by the strategic behavior of incumbent firms that do not want to be replaced by another firm.

Firms in an oligopolistic market can influence price and have the potential to earn monopoly profits. Each firm in an oligopoly controls such a large portion of the market that if it suddenly increases or decreases its quantity, the price will change drastically. This is unlike perfect competition and monopolistic competition, because individual firms in those markets are unable to substantially impact market price. An oligopoly firm's decision making process also varies from that of a monopolist because an oligopolist's decisions must be conditioned on what the other firms in the market—which also have the ability to impact the price—will do. Below, we examine how interdependence can lead firms to engage in parallel conduct. We also explore the extent to which oligopolies can coordinate their actions through explicit collusion.

i. Interdependence

One of the distinguishing characteristics of oligopolistic markets is interdependence; each firm's profitability depends on the decisions of the other firms in the market. As discussed earlier in this chapter, the profit-maximizing level of output is where its marginal cost equals marginal revenue. Marginal cost does not depend on the action of other firms, but marginal revenue does. Revenue is price multiplied by quantity sold, and price in the market will depend on all the firms' decisions because each of the firms in an oligopoly market can impact price depending on the quantity they sell. Thus, a firm's profit maximizing decision depends on pricing and output decisions of other firms.

Not surprisingly, in oligopoly industries, firms can arrive at similar prices and output levels due solely to rational independent responses to other firms' decisions. The U.S. antitrust laws recognize the interdependent nature of oligopoly markets, which is why similar pricing patterns or other tandem behavior alone do not violate the antitrust laws.

Bell Atlantic Corporation v. Twombly

Supreme Court of the United States
550 U.S. 544 (2007)

SOUTER, J.

I

The upshot of the 1984 divestiture of the American Telephone & Telegraph Company's (AT&T) local telephone business was a system of regional service monopolies (variously called "Regional Bell Operating Companies," "Baby Bells," or "Incumbent Local Exchange Carriers" (ILECs)), and a separate, competitive market for long-distance service from which the ILECs were excluded. More than a decade later, Congress withdrew approval of the ILECs' monopolies by enacting the Telecommunications Act of 1996 (1996 Act), which "fundamentally restructure[d] local telephone markets" and "subject[ed] [ILECs] to a host of duties intended to facilitate market entry."

* * *

"Central to the [new] scheme [was each ILEC's] obligation ... to share its network with competitors," which came to be known as "competitive local exchange carriers" (CLECs). A CLEC could make use of an ILEC's network in any of three ways: by (1) "purchas[ing] local telephone services at wholesale rates for resale to end users," (2) "leas[ing] elements of the [ILEC's] network 'on an unbundled basis,' " or (3) "interconnect[ing] its own facilities with the [ILEC's] network."

* * *

Respondents William Twombly and Lawrence Marcus (hereinafter plaintiffs) represent a putative class consisting of all "subscribers of local telephone and/or high speed internet services ... from February 8, 1996 to present."

* * *

The complaint alleges that the ILECs conspired to restrain trade in two ways, each supposedly inflating charges for local telephone and high-speed Internet services. Plaintiffs say, first, that the ILECs "engaged in parallel conduct" in their respective service areas to inhibit the growth of upstart CLECs. Their actions allegedly included making unfair agreements with the CLECs for access to ILEC networks, providing inferior connections to the networks, overcharging, and billing in ways designed to sabotage the CLECs' relations with their own customers. According to the complaint, the ILECs' "compelling common motivatio[n]" to thwart the CLECs' competitive efforts naturally led them to form a conspiracy; "[h]ad any one [ILEC] not sought to prevent CLECs ... from competing effectively ..., the resulting greater competitive inroads into that [ILEC's] territory would have revealed the degree to which competitive entry by CLECs would have been successful in the other territories in the absence of such conduct."

Second, the complaint charges agreements by the ILECs to refrain from competing against one another. These are to be inferred from the ILECs' common failure "meaningfully [to] pursu[e]" "attractive business opportunit[ies]" in contiguous markets where they possessed "substantial competitive advantages," and from a statement of Richard Notebaert, chief executive officer (CEO) of the ILEC Qwest, that competing in the territory of another ILEC " 'might be a good way to turn a quick dollar but that doesn't make it right[.]' "

* * *

II

A

Because § 1 of the Sherman Act "does not prohibit [all] unreasonable restraints of trade … but only restraints effected by a contract, combination, or conspiracy," *Copperweld Corp. v. Independence Tube Corp.,* 467 U.S. 752, 775 (1984), "[t]he crucial question" is whether the challenged anticompetitive conduct "stem[s] from independent decision or from an agreement, tacit or express," *Theatre Enterprises,*[*Inc. v. Paramount Film Distr. Corp.,*] 346 U.S. [537,] 540 [(1954)]. While a showing of parallel "business behavior is admissible circumstantial evidence from which the fact finder may infer agreement," it falls short of "conclusively establish[ing] agreement or … itself constitut[ing] a Sherman Act offense." *Id.,* at 540–541. Even "conscious parallelism," a common reaction of "firms in a concentrated market [that] recogniz[e] their shared economic interests and their interdependence with respect to price and output decisions" is "not in itself unlawful." *Brooke Group Ltd. v. Brown & Williamson Tobacco Corp.,* 509 U.S. 209, 227 (1993); see 6 P. Areeda & H. Hovenkamp, Antitrust Law ¶ 1433a, p. 236 (2d ed.2003) ("The courts are nearly unanimous in saying that mere interdependent parallelism does not establish the contract, combination, or conspiracy required by Sherman Act § 1")[.]

The inadequacy of showing parallel conduct or interdependence, without more, mirrors the ambiguity of the behavior: consistent with conspiracy, but just as much in line with a wide swath of rational and competitive business strategy unilaterally prompted by common perceptions of the market. Accordingly, we have previously hedged against false inferences from identical behavior at a number of points in the trial sequence. An antitrust conspiracy plaintiff with evidence showing nothing beyond parallel conduct is not entitled to a directed verdict; proof of a § 1 conspiracy must include evidence tending to exclude the possibility of independent action; and at the summary judgment stage a § 1 plaintiff's offer of conspiracy evidence must tend to rule out the possibility that the defendants were acting independently.

B

This case presents the antecedent question of what a plaintiff must plead in order to state a claim under § 1 of the Sherman Act. Federal Rule of Civil Procedure 8(a)(2) requires only "a short and plain statement of the claim showing that the pleader is entitled to relief," in order to "give the defendant fair notice of what the … claim is and the grounds upon which it rests," *Conley v. Gibson,* 355 U.S. 41, 47 (1957). While a complaint attacked by a Rule 12(b)(6) motion to dismiss does not need detailed factual allegations, a plaintiff's obligation to provide the "grounds" of his "entitle[ment] to relief" requires more than labels and conclusions, and a formulaic recitation of the elements of a cause of action will not do.

* * *

In applying these general standards to a § 1 claim, we hold that stating such a claim requires a complaint with enough factual matter (taken as true) to suggest that an agreement was made. Asking for plausible grounds to infer an agreement does not impose a probability requirement at the pleading stage; it simply calls for enough fact to raise a reasonable expectation that discovery will reveal evidence of illegal agreement. And, of course, a well-pleaded complaint may proceed even if it strikes a savvy judge that actual proof of those facts is improbable, and "that a recovery is very remote and unlikely." In identifying facts that are suggestive enough to render a § 1 conspiracy plausible, we have the benefit of the prior rulings and considered views of leading commentators, already quoted, that lawful parallel conduct fails to bespeak unlawful agreement. It makes sense to say, therefore,

that an allegation of parallel conduct and a bare assertion of conspiracy will not suffice. Without more, parallel conduct does not suggest conspiracy, and a conclusory allegation of agreement at some unidentified point does not supply facts adequate to show illegality. Hence, when allegations of parallel conduct are set out in order to make a § 1 claim, they must be placed in a context that raises a suggestion of a preceding agreement, not merely parallel conduct that could just as well be independent action.

The need at the pleading stage for allegations plausibly suggesting (not merely consistent with) agreement reflects the threshold requirement of Rule 8(a)(2) that the "plain statement" possess enough heft to "sho[w] that the pleader is entitled to relief." A statement of parallel conduct, even conduct consciously undertaken, needs some setting suggesting the agreement necessary to make out a § 1 claim; without that further circumstance pointing toward a meeting of the minds, an account of a defendant's commercial efforts stays in neutral territory. An allegation of parallel conduct is thus much like a naked assertion of conspiracy in a § 1 complaint: it gets the complaint close to stating a claim, but without some further factual enhancement it stops short of the line between possibility and plausibility of "entitle[ment] to relief." Cf. *DM Research, Inc. v. College of Am. Pathologists,* 170 F.3d 53, 56 (C.A.1 1999) ("[T]erms like 'conspiracy,' or even 'agreement,' are border-line: they might well be sufficient in conjunction with a more specific allegation—for example, identifying a written agreement or even a basis for inferring a tacit agreement, … but a court is not required to accept such terms as a sufficient basis for a complaint").

—————

ii. Collusion

A **cartel** is a combination of independent producers attempting to limit competition among themselves by acting together to fix prices, divide markets, or restrict entry into a market. A cartel exists when individual firms collude to act as one firm—as a monopoly—and make price and output decisions that maximize the profits to the industry. Such **collusive behavior** is a major focus of antitrust law because, if successful, it can result in a misallocation of resources similar to that generated by single-firm monopoly behavior.

Palmer v. BRG of Georgia

Supreme Court of the United States
498 U.S. 46 (1990)

PER CURIAM.

In preparation for the 1985 Georgia Bar Examination, petitioners contracted to take a bar review course offered by respondent BRG of Georgia, Inc. (BRG). In this litigation they contend that the price of BRG's course was enhanced by reason of an unlawful agreement between BRG and respondent Harcourt Brace Jovanovich Legal and Professional Publications (HBJ), the Nation's largest provider of bar review materials and lecture services. The central issue is whether the 1980 agreement between respondents violated § 1 of the Sherman Act.[6]

—————————

6. Section 1 of the Sherman Act, 26 Stat. 209, as amended and set forth in 15 U.S.C. § 1, provides in relevant part:

"Every contract, combination in the form of trust or otherwise, or conspiracy, in restraint of trade or commerce among the several States, or with foreign nations, is declared to be illegal."

HBJ began offering a Georgia bar review course on a limited basis in 1976, and was in direct, and often intense, competition with BRG during the period from 1977 to 1979. BRG and HBJ were the two main providers of bar review courses in Georgia during this time period. In early 1980, they entered into an agreement that gave BRG an exclusive license to market HBJ's material in Georgia and to use its trade name "Bar/Bri." The parties agreed that HBJ would not compete with BRG in Georgia and that BRG would not compete with HBJ outside of Georgia.[7] Under the agreement, HBJ received $100 per student enrolled by BRG and 40% of all revenues over $350. Immediately after the 1980 agreement, the price of BRG's course was increased from $150 to over $400.

On petitioners' motion for partial summary judgment as to the § 1 counts in the complaint and respondents' motion for summary judgment, the District Court held that the agreement was lawful. The United States Court of Appeals for the Eleventh Circuit, with one judge dissenting, agreed with the District Court that *per se* unlawful horizontal price fixing required an explicit agreement on prices to be charged or that one party have the right to be consulted about the other's prices. The Court of Appeals also agreed with the District Court that to prove a *per se* violation under a geographic market allocation theory, petitioners had to show that respondents had subdivided some relevant market in which they had previously competed. The Court of Appeals denied a petition for rehearing en banc that had been supported by the United States.

In *United States v. Socony-Vacuum Oil Co.,* 310 U.S. 150 (1940), we held that an agreement among competitors to engage in a program of buying surplus gasoline on the spot market in order to prevent prices from falling sharply was unlawful, even though there was no direct agreement on the actual prices to be maintained. We explained that "[u]nder the Sherman Act a combination formed for the purpose and with the effect of raising, depressing, fixing, pegging, or stabilizing the price of a commodity in interstate or foreign commerce is illegal *per se.*" *Id.* at 223.

The revenue-sharing formula in the 1980 agreement between BRG and HBJ, coupled with the price increase that took place immediately after the parties agreed to cease competing with each other in 1980, indicates that this agreement was "formed for the purpose and with the effect of raising" the price of the bar review course. It was, therefore, plainly incorrect for the District Court to enter summary judgment in respondents' favor. Moreover, it is equally clear that the District Court and the Court of Appeals erred when they assumed that an allocation of markets or submarkets by competitors is not unlawful unless the market in which the two previously competed is divided between them.

In *United States v. Topco Associates, Inc.,* 405 U.S. 596 (1972), we held that agreements between competitors to allocate territories to minimize competition are illegal:

> One of the classic examples of a *per se* violation of § 1 is an agreement between competitors at the same level of the market structure to allocate territories in order to minimize competition.... This Court has reiterated time and time again

---

We do not reach the other claims alleged in petitioners' nine-count complaint, including violations of § 2 of the Sherman Act, 15 U.S.C. § 2.

7. The 1980 agreement contained two provisions, one called a "Covenant Not to Compete" and the other called "Other Ventures." The former required HBJ not to "directly or indirectly own, manage, operate, join, invest, control, or participate in or be connected as an officer, employee, partner, director, independent contractor or otherwise with any business which is operating or participating in the preparation of candidates for the Georgia State Bar Examination." Plaintiffs' Motion for Partial Summary Judgment, Attachment E, p. 10. The latter required BRG not to compete against HBJ in States in which HBJ currently operated outside the State of Georgia. *Id.* at 15.

that "[h]orizontal territorial limitations … are naked restraints of trade with no purpose except stifling of competition." Such limitations are *per se* violations of the Sherman Act.

Id. at 608 (citations omitted).

The defendants in *Topco* had never competed in the same market, but had simply agreed to allocate markets. Here, HBJ and BRG had previously competed in the Georgia market; under their allocation agreement, BRG received that market, while HBJ received the remainder of the United States. Each agreed not to compete in the other's territories. Such agreements are anticompetitive regardless of whether the parties split a market within which both do business or whether they merely reserve one market for one and another for the other.[6] Thus, the 1980 agreement between HBJ and BRG was unlawful on its face.

The petition for a writ of certiorari is granted, the judgment of the Court of Appeals is reversed, and the case is remanded for further proceedings consistent with this opinion.

Notes and Questions

1. Agreements on Price: In *Palmer*, BRG argues that their arrangement with HBJ never involved an actual agreement on price. Antitrust law requires violations of the Sherman Act to result from an unlawful agreement. Do you see why their arrangement was still price fixing? Direct competition in the Georgia market for bar review services should bring down prices for law students, something the readers of this textbook would certainly be sympathetic to. By agreeing to clearly divide the markets and avoid competition, these two companies unlawfully created a small monopoly where BRG can raise the price of bar review services. The economic effect is as simple as an artificial reduction in supply that will reduce output and raise prices.

2. Application of the Per Se Rule: In *BRG*, the Supreme Court remands for further proceedings consistent with its opinion. This case will be dealt with on remand under a rule of *per se* illegality. That is, the conduct is inherently harmful and the mere existence of the agreement on price will be enough to find a violation of the antitrust laws. This can be contrasted with the rule of reason, where the plaintiff will have to fulfill a much heavier burden of proving actual harm to consumers through evidence of higher prices or reduced output. Why draw a distinction between the burdens of proof in antitrust cases? The rule of reason is applied to other circumstances in which competitors work together and cooperate through arrangements such as joint ventures, partnerships, and other forms of corporate integration.

———————

The typical cartel is faced with three major problems: (1) agreeing on a common price, (2) dividing the market share among members, (3) discouraging entry, and (4) monitoring the agreement. Looking at the first problem, consider that when a pure monopolist sets price, it only needs to compare its own cost and revenue data. When a cartel attempts to set price, it must aggregate all of the members' cost curves in order to select the profit-maximizing price and output combination. In order for the cartel to be profitable, it requires a restriction in output relative to the quantity the firms were producing when they were competing with one another. The restricted output will lead to a higher price, and the price will be greater than the marginal cost of producing the last unit sold by each firm.

The second problem facing the cartel is the division of the market, because each firm desires to increase its output at the cartel price and thus maximize its own profits. For

example, suppose the cartel has agreed to restrict total output by 20%. The most straight-forward solution to this would be for every firm to agree to restrict its output by exactly 20%. However, the firms with the lowest costs of production may argue that they shouldn't have to reduce their output as much because they are more profitable. In other words, it would cost them more to reduce output by 20%. Alternatively, the firms with the highest costs of production might argue that the other firms should absorb most of the restrictions in output because the higher-cost firms need to sell more in order to earn as much from the cartel as do the other firms. There is no way to predict how this conflict will be resolved, but it seems clear that haggling over the division of the market can threaten the formation of a cartel.

Third, cartels must prevent entry by firms that will erode the cartel's monopoly profits. Suppose that firm A wants to enter the automobile market. Firm A's cars are higher quality and lower priced than two of the three firms in an existing cartel of automobile manufacturers and firm A wants to steal business from the incumbent firms and enjoy monopoly profits. The incumbent firms do not want firm A to enter the market so they agree to flood the market with cars or engage in a price war as soon as firm A enters in order to prevent firm A from obtaining enough revenue to offset its costs. Other potential entrants observe how the cartel reacted to firm A's attempted entry and may be deterred from even attempting to enter the market. Of course, there is a limit to how many times the cartel can decrease its prices to ward off entrants before it is unable to recoup enough profits to offset the losses from the price war.

The fourth problem faced by the cartel involves the enforcement of the cartel agreement because there is much to be gained from a member of the cartel cheating on the agreement. Assume that the cartel is formed and that its members have agreed upon a price and the division of the restricted output. The cartel must now deal with the incentive that each member has to cheat on the cartel. In Figure IX-21, the cartel and each firm has agreed to restrict output by 20%—from Q_C to Q_M for the industry and from q^* to q_M for each firm. Restricting output causes price to increase from P_C to P_M. If the cartel agreement works as planned, then each firm earns monopoly profits of abP_CP_M. However, each cartel

Figure IX-21. Individual Firm's Incentive to Cheat Under a Cartel
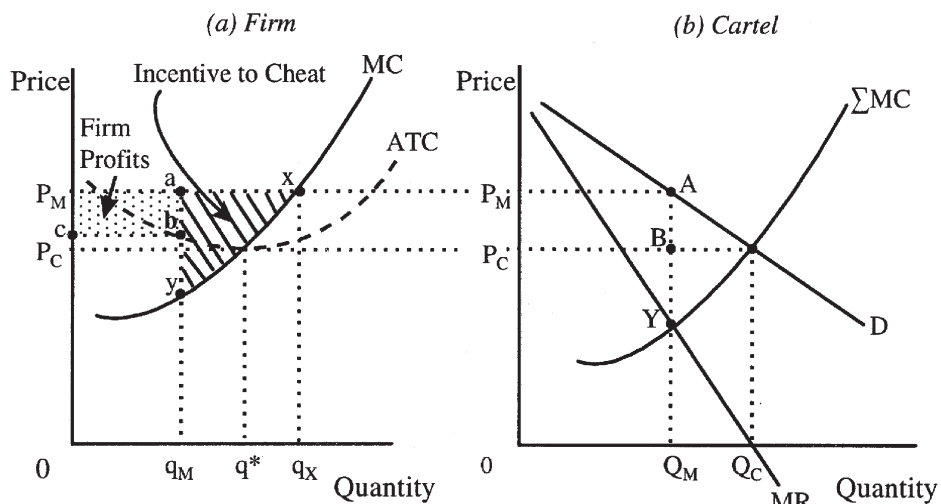
member is aware that the sale of a few additional units at the cartel price, which is well above the marginal cost of producing those additional units, will yield a substantial profit in addition to its cartel (monopoly) profits. That is, on the margin, each individual firm observes that P_M is greater than its marginal cost of producing each unit of output from q_M to q_X. Each firm has the incentive to increase its output to q_X and earn additional profits equal to the area axy. Because the cartelized industry has numerous competitors, each cartel member thinks that its increased output is so small that it will not have an adverse impact on the cartel price and thus will not be detected. For example, suppose that 100 firms form a cartel; a firm with 1 percent of the cartel's market share would not expect to have a significant impact on the total output of the cartel, so the firm has the impression that it can cheat—and earn even higher profits—by increasing its output and still have little chance of detection. The problem for the cartel and its members is that each firm feels that it can get away with the cheating, so everyone cheats and the cartel collapses.

Because of these incentives, economists argue that all cartels are unstable in the long-run. This, however, is not a justification for an antitrust policy that ignores the existence of cartels, because the cartels may impose considerable costs on society before they collapse. Cartels will also attempt to lengthen the life of the cartel through different enforcement mechanisms. In some scenarios where law enforcement is underdeveloped, cartels may be enforced through violence. However, this is not common. Cartels can also be enforced by internally imposed penalties—the agreement could be that if a firm cheats on the cartel, the next year that firm's output is restricted as punishment. The question arises though, what will keep C from cheating again and ignoring the penalty? Firms recognize the benefits of a cartel in the long run and a cheating firm would prefer monopoly profits on a smaller percentage for a short time period as a penalty with a potential of more monopoly profits after the fact instead of operating in a crumbled cartel where each firm is likely earning zero economic profits. This analysis is better understood through the lens of game theory, as discussed in Chapter I.

An understanding of the incentives to cheat on a cartel can also provide some guidance to antitrust authorities about conditions conducive to cartelization and, thus, where to look for collusion. For example, firms in highly concentrated industries are more prone to cartelization because agreements can be more easily formed and policed between a small number of firms. Conversely, as the number of firms in an industry increase, successful cartelization is less likely because it is more difficult to reach agreement, it is more difficult to detect cheating, and there is more incentive to cheat because each firm feels that it will have an inconsequential impact on total output.

Another factor conducive to cartelization is the presence of highly standardized products across firms. Standardized products permit easier monitoring of the cartel agreement. Price differences between firms are not due to differences in quality because of the standardized product, but instead most likely reflect cheating on the cartel arrangement. Firms are not able to raise prices above the cartel price because consumers will switch to other producers. However, if one firm lowers its price, consumers will switch to that producer. The other producers will observe shifts by consumers and, in all likelihood, learn of the lower price and the cheating. Because all producers know they will be detected if they cheat, cheating is less likely. It is interesting to note that the most famous attempt to form a cartel in recent decades, the Organization of Petroleum Exporting Countries (OPEC), has not been successful in setting a uniform price. Saudi Arabia, in particular, consistently undercuts the cartel price.

For more information on what makes a market conducive to coordination, see George J. Stigler, *The Theory of Oligopoly*, 72 J. Pol. Econ. 44 (1964); and Hospital Corporation of America v. FTC, 807 F.2d 1381 (1986) (Posner, J.).

Finally, the existence of a uniform price in an industry is not definitive evidence of the presence of a cartel. The perfectly competitive model of economic theory—the model that is often used as the example of the optimum allocation of resources—also predicts that all firms in an industry will sell at the same price. A major challenge in the application of the antitrust laws is determining whether a certain observed practice is indicative of competition or collusion.

Global Cartels Redux: The Amino Acid Lysine Antitrust Litigation

John M. Connor (1996)

In the evening of June 27, 1995, more than 70 FBI agents simultaneously raided the world headquarters of Archer-Daniels-Midland Company (ADM) in Decatur, Illinois and interviewed a number of ADM officers in their homes. Serving subpoenas authorized by a federal grand jury sitting in Chicago, the agents collected documents related to ADM's lysine, citric acid, and corn sweeteners businesses. Within a day or two, investigators had also raided the offices of four other companies that manufactured or imported lysine. These subpoenaed documents, together with hundreds of secret tape recordings of the conspirators' meetings and conversations, built a strong case that five companies had been illegally colluding on lysine prices around the world for three years.

The FBI raids were widely reported in the mass media and unleashed a torrent of legal actions, some of which were still unresolved seven years later. The three major federal antitrust actions were the result of an undercover investigation by the U.S. Department of Justice (DOJ) that had begun in November 1992 with the cooperation of the ADM lysine-division president. The first suit was a treble-damages class action settled in the summer of 1996. A few months later, the DOJ sought and obtained convictions for criminal price fixing by the five corporate lysine sellers. Although all the corporate members of the cartel pleaded guilty and paid historic fines, not all of the executives who managed the conspiracy were willing to plead guilty. Therefore, the DOJ prosecuted four lysine executives in a highly publicized jury trial held in Chicago in the summer of 1998; three of the four were found guilty and heavily sentenced. The five corporate conspirators were later investigated and fined by the antitrust authorities of Canada, Mexico, Brazil, and the European Union.

Within a year of the FBI raids, more than 40 civil antitrust suits were filed in federal district courts by direct buyers of lysine, each suit incorporating multiple plaintiffs. In early 1996, approximately 400 plaintiffs were certified as a single federal class, and the case called *Amino Acid Lysine Antitrust Litigation* was assigned to a judge of the U.S. District Court of Northern Illinois. In April 1996, the three largest defendants offered the class $45 million to settle the damages allegedly caused by their price fixing. Final approval of the settlement occurred in July 1996. Additional follow-up suits include about 15 actions filed by farmers, consumers and other indirect buyers of lysine in the courts of six states and two Canadian provinces. ADM was further distracted by derivative shareholders' suits charging mismanagement by the company's managers and board of directors.

The three federal lysine cases were important for at least four reasons. First, it was the U.S. Government's first completely successful conviction of a global cartel in more than four decades.

* * *

Second, the conviction of the lysine cartel was the first public manifestation of a sea change in enforcement priorities by U.S. and overseas antitrust officials. Prior to 1995, less than 1 percent of the price-fixing indictments by the DOJ involved at least one non-U.S.-based corporation or non-U.S. resident. By contrast, beginning during 1998–2000 more than half of all criminal price-fixing indictments were brought against international conspirators. The investigation of the lysine cartel led directly to the discovery and successful prosecution of 30 multinational corporations that participated in global price fixing in the markets for lysine, citric acid, sodium gluconate, and ten bulk vitamins. Since 1996, more than a score of global cartels have been uncovered and prosecuted by the DOJ, the Competition Policy Directorate of the European Commission (DG-IV), and other antitrust agencies around the world. Cartel enforcement remains a high priority for the Antitrust Division of the DOJ, which is devoting 30 percent of its resources to criminal price-fixing prosecution.

* * *

Finally, the lysine cases and those that followed soon thereafter showed that the sanctions for criminal price-fixing had escalated enormously in the 1990s. Not only has Congress steadily raised the statutory fine for Sherman Act violations (up to $10 million for corporations), it also in 1994 made criminal antitrust violations felonies instead of misdemeanors. Combined with the U.S. Sentencing Guidelines first promulgated in 1987, corporate price fixers are now liable for criminal penalties as high as "double the harm" caused by a cartel. That is, corporations can be fined by the government up to twice the monopoly overcharge generated by a cartel, an amount that can easily exceed the $10-million statutory cap when market sales are large. ADM, the leader of the lysine cartel, was fined $100 million for its role in two criminal price-fixing schemes—a record amount that was twice eclipsed in the late 1990s by leaders of highly injurious global cartels. In fiscal years 1998–2001, the Antitrust Division collected more than $2 billion in fines for criminal price fixing, of which more than three-fourths was from members of international cartels. The EU's DG-IV, which operates on a somewhat delayed schedule, imposed record fines of 1.84 billion euros on hard-core cartels in 2001 alone; these fines are loosely based on the cartels' overcharges to customers in the European Economic Space. Both U.S. and EU authorities are empowered to base their fines on worldwide overcharges rather than their jurisdictional injuries, and the U.S. DOJ has done so at least twice.

U.S. Government fines are mere paper cuts compared to the financial wounds that may be inflicted by plaintiffs in civil actions. Direct buyers suing in federal courts, the principal focus of this chapter, are entitled to seek treble damages. In some cases, direct buyers abroad are permitted to seek treble damage in U.S. courts. However, antitrust liability does not stop there. Nearly 20 states allow their residents who are *indirect* purchasers to sue in state courts, most of which permit treble damages. In addition, state attorneys general increasingly have banded together to pursue antitrust claims in federal courts (*parens patriae* suits) to recover treble damages for their state governments and for corporate and individual indirect buyers residing in their states. For example, in October 2000, the attorneys general of more than 40 states announced a settlement totaling $340 million to be paid by the six largest members of the vitamins cartels. Not counting the losses associated with derivative shareholders' suits, legal fees, and reputational effects, corporations accused of criminal price fixing now face maximum antitrust liabilities that range from *eight to twelve times* the cartel's U.S. overcharges. The fines and prison terms meted out to cartel managers have also risen.

The major role played by economic analysis in horizontal price-fixing cases is the calculation of the *overcharge* on buyers in markets affected by a cartel. The overcharge is

the value of purchases of a cartelized product actually made minus what the sales would have been for the same volume of product absent the cartel. Accurate estimates of conspiracy-induced overcharges are important not only because of recovery of civil damages, but also because overcharges are the basis for the calculation of government fines.

Notes and Questions

1. The Informant: The investigation of the lysine cartel was successful much to the credit of one Mark Whitacre. Whitacre was the president of ADM's bioproducts division and came forward as an informant whose aid to the government was invaluable. The story is more complicated than that and Whitacre's actions are well-documented in Kurt Eichenwald, The Informant: A True Story (2001). These actions have also been chronicled in a popular NPR radio show (see This American Life: The Fix is In, Chicago Public Radio (Sept. 15, 2000)) and a Hollywood big screen adaptation starring Matt Damon. To give you just a taste of the twists and turns of the case read the following description from the Seventh Circuit Court of Appeals:

The Investigation:

Mark E. Whitacre joined ADM in 1989 as president of its bioproducts division. That year, ADM announced that it would enter the lysine market dominated by Asian producers. Whitacre, who held a Ph.D. in biochemistry from Cornell University and degrees in agricultural science, answered directly to Mick Andreas. Just 32 years old when he joined the company, Whitacre's star clearly was rising fast at ADM, and some industry analysts thought he could be the next president of ADM.

In 1992, Whitacre began working with Wilson, and the two attended the first meetings of the lysine producers in Mexico City. Also in 1992, Whitacre began embezzling large sums of money from ADM and eventually stole at least $9 million from the company by submitting to ADM phony invoices for work done by outside companies, who would then funnel the money to Whitacre's personal offshore and Swiss bank accounts. To cover up the embezzlement, Whitacre hatched a scheme in the summer of 1992 to accuse Ajinomoto of planting a saboteur in ADM's Decatur plant. Whitacre would accuse the saboteur of contaminating the delicate bacterial environment needed for the production of lysine, a story made believable because of the many early difficulties the ADM lysine plant encountered.

In accordance with the plot, Whitacre told Mick Andreas that an engineer at Ajinomoto named Fujiwara had contacted him at his home and offered to sell ADM the name of the saboteur in exchange for $10 million. The story was a lie. However, Dwayne Andreas believed it and feared it could jeopardize relations between the United States and Japan. He called the CIA, but the CIA, considering the matter one of federal law enforcement rather than national security, directed the call to the FBI, which sent agents out to ADM to interview Whitacre and other officials about the extortion. Whitacre apparently had not expected this and realized quickly that his lie would be discovered by the FBI, particularly after Special Agent Brian Shepard asked Whitacre if he could tap Whitacre's home telephone to record the next extortion demand. Whitacre knew that when the extortionist failed to call, Shepard would know Whitacre had invented the story. Whitacre confessed the scheme to Shepard, but to save himself, he agreed to

become an undercover informant to help the FBI investigate price fixing at ADM. He did not come totally clean with the FBI, however; he failed to mention the millions he embezzled and in fact continued to embezzle after he began working for the government. For the next two-and-a-half years, Whitacre acted as an undercover cooperating witness — legally a government agent — and secretly taped hundreds of hours of conversations and meetings with Wilson, Mick Andreas and the other conspirators. In addition, the FBI secretly videotaped meetings of the lysine producers.

Whitacre made between 120 and 130 tapes for the FBI during the investigation, beginning with a November 9, 1992, conversation with Yamamoto, by using recording equipment, tapes and instruction provided by the government. FBI agents met with Whitacre more than 150 times during the investigation. The tapes were collected and reviewed usually within a day or two of the FBI receiving them, and Department of Justice (DOJ) attorneys regularly participated in reviewing the tapes and monitoring the supervision of Whitacre. However, the FBI's supervision of Whitacre was not flawless. Whitacre was, to say the least, a difficult cooperating witness to handle. Whitacre lied to the FBI during the probe, failed polygraph tests, bragged to his gardener about his role as an FBI mole, all while continuing to embezzle millions of dollars from the company. He even envisioned himself ascending to the ADM presidency as a hero once Andreas, Wilson and Randall were taken down in the FBI sting. In short, he was out of control, and the FBI struggled to keep him on track. Nonetheless, the FBI and the DOJ considered him the best opportunity to stop a massive price-fixing scheme.

Whitacre exercised much discretion in deciding which conversations to record. He was given a tape recorder that could be hidden in his coat breast pocket and another that could be stowed in his briefcase. Agent Shepard showed him how to use the devices and sometimes affixed a recording device to Whitacre's body. Another recording device was used to tap one of Whitacre's home telephones, but not his cellular telephone. All recordings were done with Whitacre's express, signed consent, and all but one were done after Whitacre confessed that his story about a saboteur was a hoax and he began cooperating.

Whitacre was told to record conversations relevant to the conspiracy, but not to record anything about ADM's legitimate business. In direct contravention of the FBI's recording policy, Whitacre did not record many conversations he had with the alleged conspirators. The record shows Whitacre telephoned Ajinomoto and Kyowa 114 times, but 80 were never recorded or documented by the FBI as required. In addition, many conversations with co-defendants Wilson and Andreas were never recorded or documented.

Whitacre once claimed that Shepard ordered him to destroy tapes bearing exculpatory conversations, but Shepard denied this charge and Whitacre later recanted it in a sworn affidavit. Both Whitacre and a friend he entrusted with some of the tapes testified that no tapes were destroyed at Shepard's command. A tape expert testified for the government that none of the tapes exhibited evidence of splicing or alteration and that only a few showed evidence of "bulk erasure" or over-recording. Although that meant that some recordings may have been taped over, the expert expressed an opinion that none of the final recordings had been altered.

Andreas and Wilson moved to suppress the inculpatory tapes before trial and to allow them to introduce evidence that exculpatory tapes had been destroyed.

> For reasons explained below, the motion was denied although the trial court found that the FBI's supervision of Whitacre and its blatant inability to follow its own internal policies "border on gross negligence."

United States v. Andreas, 216 F.3d 645, 654–56 (7th Cir. 2000).

2. Other Global Cartels, Vitamins and DRAM: The discovery of such a massive price fixing scheme highlighted the general failure of the government's ability to adequately enforce the antitrust laws. Even the lysine cartel was stumbled onto basically by accident. In an effort to increase the likelihood of discovery the government changed its leniency program to be more effective by making both the individual and the corporation immune from criminal and civil prosecution if an investigation had not already been opened. As the Justice Department describes it, the program "is its most important investigative tool for detecting cartel activity. Corporations and individuals who report their cartel activity and cooperate in the Division's investigation of the cartel reported can avoid criminal conviction, fines, and prison sentences if they meet the requirements of the program." *See* The Antitrust Division's Leniency Program, U.S. Dep't of Justice, available online at http://www.justice.gov/atr/public/criminal/leniency.html.

This change has been credited with helping to increase successful detection of cartels. A small list of cartels that have been uncovered with cooperation through the program include the Vitamins Cartel (garnering over $700 million in U.S. fines), the Dynamic Random Access Memory Cartel (over $600 million in U.S. fines from microchip manufacturing companies), the Graphite Electrodes Cartel (over $350 million in U.S. fines), and the Fine Arts Auctions Cartel (with $45 million in U.S. fines). This strategy has been defined as a carrot and a stick approach, offering harsh fines and jail times for participants and leniency for those who come forward and cooperate.

3. The Rise of Criminal Antitrust: Adam Smith wrote in the Wealth of Nations, his groundbreaking treatise on capitalism, that "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." The antitrust laws have strived to deter this behavior and have increasingly sought out more strict punishments as a means of enforcement. Recent results have been substantial. Corporate fines increased from $1.6 billion in the 1990s to $4.2 billion in the 2000s. Criminal fines have seen a similar increase. Jail time, representing individual criminal accountability for cartel behavior, has also seen a nearly five-fold increase with the DOJ seeking jail time for more participants and higher profile corporate executives. There has also been a strong expansion of antitrust laws worldwide and greater cooperation between governments to help investigate and bring cartels to justice. For more information, see Scott D. Hammond, The Evolution of Criminal Antitrust Enforcement Over the Last Two Decades, U.S. Dep't of Justice (Feb. 25, 2010), available at http://www.justice.gov/atr/public/speeches/255515.pdf.

4. Apple, eBooks, and Price-Fixing. Consider the recent case of United States v. Apple Inc., 12 CIV. 2826 DLC, 2013 WL 3454986 (S.D.N.Y. July 10, 2013). In this case, Apple Inc. (Apple), a massive U.S. technology company renowned for its personal computers, smart phones (iPhones), personal music players (iPods), electronic tablets (iPads), and other computerized electronic devices, organized a new bookstore for ebooks, which could be read on Apple's new iPad device—an electronic tablet capable of serving as a multi-functional handheld computer and ebook reader. The bookstore offered books from five of the six largest ebook publishers in the U.S. The marker had previously been dominated by Amazon.com Inc. (Amazon).

Amazon offered the Kindle, a handheld ebook reader, and had controlled around 90 percent of ebook sales in the U.S. Amazon priced its books based on the wholesale model, where Amazon purchased the right to sell and determine the final sales price of ebooks. Amazon offered most of its ebooks for the very low price of $9.99 and substantial anecdotal evidence indicated that the ebook publishers believed this price was shockingly low and jeopardized their ability to continue to profitably operate in the market.

The Department of Justice alleged that Apple had organized a price-fixing conspiracy between the five ebook publishers that led to a change from the wholesale model to an agency model where the publishers retained control and conspired to raise the final sales price of ebooks. Having already settled this case with each of the five publishers, the DOJ proceeded to trial against Apple alone and alleged that Apple directly organized a horizontal conspiracy between the publishers.

The DOJ also alleged that Apple negotiated maximum pricing tiers on all ebooks and then imposed a most-favored-nations clause (MFN) that promised that Apple would charge the lowest price that any ebook retailer offered on the market. The effect of the MFN was that if the publishers wanted to raise prices, they had to institute their new and higher pricing scheme on all the other ebook retailers including Amazon, something that happened briefly after the formation of the Apple bookstore. Despite Apple's defense that it did not force the publishers to raise prices and was unaware of the effect the agreement would have on the market, comments from its then-CEO Steve Jobs seemed to indicate that Apple knew exactly what it was doing. The court recounted Jobs' interaction with a reporter after announcing the iBookstore when he was asked:

> [W]hy people would pay $14.99 in the iBookstore to purchase an e-book that was selling at Amazon for $9.99, Jobs told a reporter, "Well, that won't be the case." When the reporter sought to clarify, "You mean you won't be 14.99 or they won't be 9.99?" Jobs paused, and with a knowing nod responded, "The price will be the same," and explained that "Publishers are actually withholding their books from Amazon because they are not happy."

The court then concluded that Apple violated Section 1 of the Sherman Antitrust Act. Consider Apple's arguments that it 1) was vertically related to the publishers, 2) was opening a bookstore that required the participation of multiple competitors, 3) lacked any market power or even a presence in this market prior to the opening of the bookstore, and 4) did not directly set or determine the final prices of ebooks. Is it fair to find Apple liable for an antitrust violation? Many of the publishers argued that Amazon sold ebooks at a loss and that the $9.99 price was too low to operate their business. Accordingly, they contended that any higher prices in the market were reasonable, still very low, and a return to a market clearing equilibrium. If this fact is true does it change the approach to a price-fixing conspiracy? Should it? The DOJ's case drew criticism from commentators who believed that the shift from the wholesale model to the agency model was not one motivated by price but by an industry's desire to retain control over its own final products. Is this kind of pricing model any different from the resale price maintenance in the *Leegin* case? Are there any procompetitive business justifications for using the agency model over the wholesale model? Currently, Apple intends to appeal the decision, if you were the appellate court judge how would you rule? For a more nuanced analysis of this case, see David Bosco, et al., *e-Books and the Boundaries of Antitrust*, 3-2012 Concurrences (Sept. 3, 2012), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id= 2140778.

5. Using Government to Enforce a Cartel. Price fixing is an agreement between individual firms that are competitors for the purpose of setting a specific price, a range of prices, or

otherwise raising, depressing or stabilizing the market price. One of the greatest difficulties that cartels face in earning monopoly profits is enforcing the cartel price. Given the difficulty of enforcing the cartel agreement, it is not surprising that many cartels seek government assistance in policing the agreement. In lobbying for the government to help enforce the cartel agreement, cartels often attempt to invoke a public interest rationale.

Consider the case Goldfarb v. Virginia State Bar, 421 U.S. 773 (2004). The Fairfax County State Bar published a minimum fee schedule for all attorney services. Although the Virginia State Bar had never formally punished anyone for violations of the fee schedule they enforced it by publicizing support for the fee schedule and had twice issued ethical opinions making clear that attorneys are not free to ignore the suggested prices. One such opinion stated that "an attorney [who] habitually charges less than the suggested minimum fee schedule adopted by his local bar Association, raises a presumption that such lawyer is guilty of misconduct."

The plaintiffs in the case were a husband and wife who were contracting to buy a house in Fairfax, Virginia. In order to secure financing they were required to secure title insurance and have a title examination done on the property. Only a registered Virginia attorney could legally perform the service. The attorney in question always set his fees in line with the minimum fee schedule. The couple sent letters to 36 other Fairfax County lawyers and found that none that responded offered price lower than the suggested minimum prices of the fee schedule. The couple then brought a class action against the State and County Bar Associations alleging violations of § 1 of the Sherman Act, claiming that the minimum fee schedule was a price-fixing scheme and an agreement between competitors in restraint of trade. The Supreme Court agreed, finding that:

> The County Bar argues that because the fee schedule is merely advisory, the schedule and its enforcement mechanism do not constitute price fixing. Its purpose, the argument continues, is only to provide legitimate information to aid member lawyers in complying with Virginia professional regulations. Moreover, the County Bar contends that in practice the schedule has not had the effect of producing fixed fees. The facts found by the trier belie these contentions, and nothing in the record suggests these findings lack support.

> A purely advisory fee schedule issued to provide guidelines, or an exchange of price information without a showing of an actual restraint on trade, would present us with a different question. The record here, however, reveals a situation quite different from what would occur under a purely advisory fee schedule. Here a fixed, rigid price floor arose from respondents' activities: every lawyer who responded to petitioners' inquiries adhered to the fee schedule, and no lawyer asked for additional information in order to set an individualized fee. The price information disseminated did not concern past standards, but rather minimum fees to be charged in future transactions, and those minimum rates were increased over time. The fee schedule was enforced through the prospective professional discipline from the State Bar, and the desire of attorneys to comply with announced professional norms; the motivation to conform was reinforced by the assurance that other lawyers would not compete by underbidding. This is not merely a case of an agreement that may be inferred from an exchange of price information, for here a naked agreement was clearly shown, and the effect on prices is plain.

Moreover, in terms of restraining competition and harming consumers like petitioners the price-fixing activities found here are unusually damaging. A title examination is in-

dispensable in the process of financing a real estate purchase, and since only an attorney licensed to practice in Virginia may legally examine a title, consumers could not turn to alternative sources for the necessary service. All attorneys of course, were practicing under the constraint of the fee schedule. The County Bar makes much of the fact that it is a voluntary organization; however, the ethical opinions issued by the State Bar provide that any lawyer, whether or not a member of his county bar association, may be disciplined for "habitually charg[ing] less than the suggested minimum fee schedule adopted by his local bar Association...." These factors coalesced to create a pricing system that consumers could not realistically escape. On this record respondents' activities constitute a classic illustration of price fixing.

Why does the State Bar Association publish a schedule of suggested minimum fees? In light of the costs imposed upon society by these fees, the answer to this question is very important. The State Bar's Committee on Legal Ethics seems to play a large role in the justification for the minimum fee schedule. What is the ethical justification for the minimum fee schedule? Consider what the Court observes in footnote 16: "The reason for adopting the fee schedule does not appear to have been wholly altruistic. The first sentence in respondent State Bar's 1962 Minimum Fee Schedule Report states: 'The lawyers have slowly, but surely, been committing economic suicide as a profession.'"

How would you characterize the industry structure in *Goldfarb*? The market for legal services is limited to those who have gone to law school and passed a state bar examination. Thus, the number of sellers in the market is limited. Moreover, the State Bar Association could be used as a tool by which attorneys could engage in collusion. For example, the State Bar could help the legal profession act in concert for restriction of output and increased prices. This is one reason that the U.S. Government has kept a close eye on centrally organized professional organizations. The regulation of these organizations is both necessary and beneficial for consumers by maintaining standards and quality yet they also represent a danger of collusion, cartel-behavior, and price-fixing. If you are interested in learning more about these kinds of cases take a look at National Society of Professional Engineers v. United States, 435 U.S. 679 (1978); Federal Trade Commission v. Superior Court Trial Lawyers Association, 493 U.S. 411 (1990); and California Dental Association v. Federal Trade Commission, 526 U.S. 756 (1999).

6. Policing the Cartel: A cartel for legal services would face the same fundamental problem that all cartels face. Each individual attorney has an incentive to cheat by setting his price just below the minimum imposed by the cartel. In this way he can increase the quantity of his services provided and earn a greater portion of the monopoly profits. If enough cheating were to occur, the cartel would fall apart and the attorneys would earn only a normal profit for their services. The challenge to the Bar Association is how to stop individual attorneys from cheating on the cartel. Footnote 4 of the Court's opinion contains a quote from the State Bar Committee on Legal Ethics suggesting that any attorney who lowers his price below the "customary charges" risks violation of ethical canons forbidding solicitation and encroachment. In addition, professional sanctions were available to enforce these "ethical" considerations against those who would cut prices. Moreover, the State Bar's second fee schedule stated that "lawyers should feel free to charge more than the recommended fees; and to avoid condemnation of higher fees charged by some lawyers." What if, hypothetically, the fees on the schedule actually represented the competitive market price for lawyers' services? If the Bar Association could prove that fact would it still be price-fixing? In other words, is price fixing only illegal under the Sherman Act when prices are unreasonable or is the coordination of rivals sufficient to constitute a violation?

iii. Other Horizontal Arrangements and the Rule of Reason

Many agreements among competitors have nothing to do with price fixing or market allocation, but instead are attempts to increase efficiency. For example, the vast majority of mergers—which are agreements among competitors to merge their businesses—are not challenged because they are unlikely to pose a competitive threat and often generate efficiencies that the parties could not achieve absent the merger.

By law, parties to mergers involving assets above a certain threshold must inform the antitrust agencies—the FTC and DOJ—prior to the consummation of the merger. This filing, under the Hart-Scott-Rodino Act,[8] allows the agencies to review these transactions before they occur—thereby providing them an opportunity to prevent any competitive harm that is likely to result. Broadly, the agencies apply a cost benefit test to mergers, weighing likely competitive harms from combining rivals' economic decision-making against efficiencies generated from the transaction. U.S. Dep't of Justice & Fed. Trade Comm'n, Horizontal Merger Guidelines (2010). In 2012, 1,429 transactions were filed with the federal government. This is slightly less than in 2011. Of those mergers, 44 were challenged by either the DOJ or the FTC. The FTC brought 25 actions; 15 led to consent orders for public comment, seven mergers were abandoned or restructured, and 3 resulted in FTC-initiated administrative litigation. The DOJ Antitrust Division challenged 19 mergers; in eight the Antitrust Division filed a complaint in federal district court, of which seven mergers settled and one was abandoned. In the other eleven cases, six mergers were abandoned, two mergers were restructured, and in three the companies altered their conduct to avoid anticompetitive concerns. A full report reviewing these actions is filed annually by the agencies and can be found at U.S. Dep't Justice & Fed. Trade Comm'n, Hart-Scott-Rodino Annual Report: Section 7A of the Clayton Act, Antitrust Improvements Act of 1976 (2012), *available at* http://www.ftc.gov/os/2013/04/130430hsrreport.pdf.

Other types of horizontal agreements fall short of mergers, but are also necessary to create efficiencies. The creation of new information plays an important part in improving our societal standard of living. New medicines lead to healthier and more productive lives. New technologies improve efficiency, reducing the amount of resources consumed in production and, therefore, increasing the scope of activities that can be pursued. However, information is intangible and thus open to easy duplication. The rewards for discovering new information may then be short lived. In other words, despite the costs that one may incur in order to develop new information, the rewards may be easily confiscated by competitors. This fact may serve as a disincentive towards investment in discovering new information. In response to this disincentive, Congress has created statutory patent and copyright laws.

Patents and copyrights provide their holder with a time-constrained monopoly. This time constraint is just as economically important as the original monopoly grant. As discussed in Chapter I, the monopoly grant creates an incentive to invest time and other resources in the production of new information. Limiting the time that the monopoly is "good for" ensures that at some point this information will be handed over to a competitive market. This presents a delicate balancing act. On the one hand, it is desirable to encourage the innovation of new information; on the other hand, enormous benefits emerge from competitive markets.

---

8. 15 USC § 18a.

Broadcast Music, Inc. v. Columbia Broadcasting System, Inc.

Supreme Court of the United States

441 U.S. 1 (1979)

Mr. Justice WHITE delivered the opinion of the Court.

This case involves an action under the antitrust and copyright laws brought by respondent Columbia Broadcasting System, Inc. (CBS), against petitioners, American Society of Composers, Authors and Publishers (ASCAP) and Broadcast Music, Inc. (BMI), and their members and affiliates. The basic question presented is whether the issuance by ASCAP and BMI to CBS of blanket licenses to copyrighted musical compositions at fees negotiated by them is price fixing *per se* unlawful under the antitrust laws.

I

CBS operates one of three national commercial television networks, supplying programs to approximately 200 affiliated stations and telecasting approximately 7,500 network programs per year. Many, but not all, of these programs make use of copyrighted music recorded on the soundtrack. CBS also owns television and radio stations in various cities. . . .

Since 1897, the copyright laws have vested in the owner of a copyrighted musical composition the exclusive right to perform the work publicly for profit, but the legal right is not self-enforcing. In 1914, Victor Herbert and a handful of other composers organized ASCAP because those who performed copyrighted music for profit were so numerous and widespread, and most performances so fleeting, that as a practical matter it was impossible for the many individual copyright owners to negotiate with and license the users and to detect unauthorized uses. "ASCAP was organized as a 'clearing-house' for copyright owners and users to solve these problems" associated with the licensing of music. As ASCAP operates today, its 22,000 members grant it nonexclusive rights to license nondramatic performances of their works, and ASCAP issues licenses and distributes royalties to copyright owners in accordance with a schedule reflecting the nature and amount of the use of their music and other factors.

BMI, a nonprofit corporation owned by members of the broadcasting industry, was organized in 1939, is affiliated with or represents some 10,000 publishing companies and 20,000 authors and composers, and operates in much the same manner as ASCAP. Almost every domestic copyrighted composition is in the repertory either of ASCAP, with a total of three million compositions, or of BMI, with one million.

Both organizations operate primarily through blanket licenses, which give the licensees the right to perform any and all of the compositions owned by the members or affiliates as often as the licensees desire for a stated term. Fees for blanket licenses are ordinarily a percentage of total revenues or a flat dollar amount, and do not directly depend on the amount or type of music used. Radio and television broadcasters are the largest users of music, and almost all of them hold blanket licenses from both ASCAP and BMI. Until this litigation, CBS held blanket licenses from both organizations for its television network on a continuous basis since the late 1940's and had never attempted to secure any other form of license from either ASCAP[9] or any of its members.

The complaint filed by CBS charged various violations of the Sherman Act and the copyright laws. CBS argued that ASCAP and BMI are unlawful monopolies and that the

---

9. Unless the context indicates otherwise, references to ASCAP alone in this opinion usually apply to BMI as well.

blanket license is illegal price fixing, an unlawful tying arrangement, a concerted refusal to deal, and a misuse of copyrights. The District Court, though denying summary judgment to certain defendants, ruled that the practice did not fall within the *per se* rule. After an 8-week trial, limited to the issue of liability, the court dismissed the complaint, rejecting again the claim that the blanket license was price fixing and a *per se* violation of §1 of the Sherman Act, and holding that since direct negotiation with individual copyright owners is available and feasible there is no undue restraint of trade, illegal tying, misuse of copyrights, or monopolization.

Though agreeing with the District Court's factfinding and not disturbing its legal conclusions on the other antitrust theories of liability, the Court of Appeals held that the blanket license issued to television networks was a form of price fixing illegal *per se* under the Sherman Act. This conclusion, without more, settled the issue of liability under the Sherman Act ... and required reversal of the District Court's judgment, as well as a remand to consider the appropriate remedy.[10]

ASCAP and BMI petitioned for certiorari, presenting the question[] of the applicability of the *per se* rule.... CBS did not cross petition to challenge the failure to sustain its other antitrust claims. We granted certiorari because of the importance of the issues to the antitrust and copyright laws. Because we disagree with the Court of Appeals' conclusions with respect to the *per se* illegality of the blanket license, we reverse its judgment and remand the cause for further appropriate proceedings.

II

In construing and applying the Sherman Act's ban against contracts, conspiracies, and combinations in restraint of trade, the Court has held that certain agreements or practices are so "plainly anticompetitive," and so often "lack ... any redeeming virtue," that they are conclusively presumed illegal without further examination under the rule of reason generally applied in Sherman Act cases. This *per se* rule is a valid and useful tool of antitrust policy and enforcement. And agreements among competitors to fix prices on their individual goods or services are among those concerted activities that the Court has held to be within the *per se* category. But easy labels do not always supply ready answers.

A

To the Court of Appeals and CBS, the blanket license involves "price fixing" in the literal sense: the composers and publishing houses have joined together into an organization that sets its price for the blanket license it sells. But this is not a question simply of determining whether two or more potential competitors have literally "fixed" a "price." As generally used in the antitrust field, "price fixing" is a shorthand way of describing

---

10. The Court of Appeals went on to suggest some guidelines as to remedy, indicating that despite its conclusion on liability the blanket license was not totally forbidden. The Court of Appeals said: "Normally, after a finding of price-fixing, the remedy is an injunction against the price-fixing—in this case, the blanket license. We think, however, that if on remand a remedy can be fashioned which will ensure that the blanket license will not affect the price or negotiations for direct licenses, the blanket license need not be prohibited in all circumstances. The blanket license is not simply a 'naked restraint' ineluctably doomed to extinction. There is not enough evidence in the present record to compel a finding that the blanket license does not serve a market need for those who wish full protection against infringement suits or who, for some other business reason, deem the blanket license desirable.... Our objection to the blanket license is that it reduces price competition among the members and provides a disinclination to compete. We think that these objections may be removed if ASCAP itself is required to provide some form of per use licensing which will ensure competition among the individual members with respect to those networks which wish to engage in per use licensing."

certain categories of business behavior to which the *per se* rule has been held applicable. The Court of Appeals' literal approach does not alone establish that this particular practice is one of those types or that it is "plainly anticompetitive" and very likely without "redeeming virtue." Literalness is overly simplistic and often overbroad. When two partners set the price of their goods or services they are literally "price fixing," but they are not *per se* in violation of the Sherman Act. Thus, it is necessary to characterize the challenged conduct as falling within or without that category of behavior to which we apply the label "*per se* price fixing." That will often, but not always, be a simple matter.

Consequently, as we recognized in *United States v. Topco Associates, Inc.,*"[i]t is only after considerable experience with certain business relationships that courts classify them as *per se* violations...." We have never examined a practice like this one before; indeed, the Court of Appeals recognized that "[i]n dealing with performing rights in the music industry we confront conditions both in copyright law and in antitrust law which are *sui generis*." And though there has been rather intensive antitrust scrutiny of ASCAP and its blanket licenses, that experience hardly counsels that we should outlaw the blanket license as a *per se* restraint of trade.

B

* * *

The Department of Justice first investigated allegations of anticompetitive conduct by ASCAP over 50 years ago. A criminal complaint was filed in 1934, but the Government was granted a midtrial continuance and never returned to the courtroom. In separate complaints in 1941, the United States charged that the blanket license, which was then the only license offered by ASCAP and BMI, was an illegal restraint of trade and that arbitrary prices were being charged as the result of an illegal copyright pool. The Government sought to enjoin ASCAP's exclusive licensing powers and to require a different form of licensing by that organization. The case was settled by a consent decree that imposed tight restrictions on ASCAP's operations. Following complaints relating to the television industry, successful private litigation against ASCAP by movie theaters, and a Government challenge to ASCAP's arrangements with similar foreign organizations, the 1941 decree was reopened and extensively amended in 1950.

Under the amended decree, which still substantially controls the activities of ASCAP, members may grant ASCAP only nonexclusive rights to license their works for public performance. Members, therefore, retain the rights individually to license public performances, along with the rights to license the use of their compositions for other purposes. ASCAP itself is forbidden to grant any license to perform one or more specified compositions in the ASCAP repertory unless both the user and the owner have requested it in writing to do so. ASCAP is required to grant to any user making written application a nonexclusive license to perform all ASCAP compositions either for a period of time or on a per-program basis. ASCAP may not insist on the blanket license, and the fee for the per-program license, which is to be based on the revenues for the program on which ASCAP music is played, must offer the applicant a genuine economic choice between the per-program license and the more common blanket license. If ASCAP and a putative licensee are unable to agree on a fee within 60 days, the applicant may apply to the District Court for a determination of a reasonable fee, with ASCAP having the burden of proving reasonableness.

The 1950 decree, as amended from time to time, continues in effect, and the blanket license continues to be the primary instrument through which ASCAP conducts its business under the decree. The courts have twice construed the decree not to require ASCAP to issue licenses for selected portions of its repertory. It also remains true that the decree

guarantees the legal availability of direct licensing of performance rights by ASCAP members; and the District Court found, and in this respect the Court of Appeals agreed, that there are no practical impediments preventing direct dealing by the television networks if they so desire. Historically, they have not done so. Since 1946, CBS and other television networks have taken blanket licenses from ASCAP and BMI. It was not until this suit arose that the CBS network demanded any other kind of license.

Of course, a consent judgment, even one entered at the behest of the Antitrust Division, does not immunize the defendant from liability for actions, including those contemplated by the decree, that violate the rights of nonparties. But it cannot be ignored that the Federal Executive and Judiciary have carefully scrutinized ASCAP and the challenged conduct, have imposed restrictions on various of ASCAP's practices, and, by the terms of the decree, stand ready to provide further consideration, supervision, and perhaps invalidation of asserted anticompetitive practices. In these circumstances, we have a unique indicator that the challenged practice may have redeeming competitive virtues and that the search for those values is not almost sure to be in vain. Thus, although CBS is not bound by the Antitrust Division's actions, the decree is a fact of economic and legal life in this industry, and the Court of Appeals should not have ignored it completely in analyzing the practice....

* * *

III

Of course, we are no more bound than is CBS by the views of the Department of Justice, the results in the prior lower court cases, or the opinions of various experts about the merits of the blanket license. But while we must independently examine this practice, all those factors should caution us against too easily finding blanket licensing subject to *per se* invalidation.

A

As a preliminary matter, we are mindful that the Court of Appeals' holding would appear to be quite difficult to contain....

Although the Court of Appeals apparently thought the blanket license could be saved in some or even many applications, it seems to us that the *per se* rule does not accommodate itself to such flexibility and that the observations of the Court of Appeals with respect to remedy tend to impeach the *per se* basis for the holding of liability.

CBS would prefer that ASCAP be authorized, indeed directed, to make all its compositions available at standard per-use rates within negotiated categories of use. But if this in itself or in conjunction with blanket licensing constitutes illegal price fixing by copyright owners, CBS urges that an injunction issue forbidding ASCAP to issue any blanket license or to negotiate any fee except on behalf of an individual member for the use of his own copyrighted work or works. Thus, we are called upon to determine that blanket licensing is unlawful across the board. We are quite sure, however, that the *per se* rule does not require any such holding.

B

In the first place, the line of commerce allegedly being restrained, the performing rights to copyrighted music, exists at all only because of the copyright laws.... Although the copyright laws confer no rights on copyright owners to fix prices among themselves or otherwise to violate the antitrust laws, we would not expect that any market arrangements reasonably necessary to effectuate the rights that are granted would be deemed a *per se* violation of the Sherman Act. Otherwise, the commerce anticipated by the Copyright

Act and protected against restraint by the Sherman Act would not exist at all or would exist only as a pale reminder of what Congress envisioned.

<div align="center">C</div>

More generally, in characterizing this conduct under the *per se* rule, our inquiry must focus on whether the effect and, here because it tends to show effect, the purpose of the practice are to threaten the proper operation of our predominantly free-market economy — that is, whether the practice facially appears to be one that would always or almost always tend to restrict competition and decrease output, and in what portion of the market, or instead one designed to "increase economic efficiency and render markets more, rather than less, competitive."

The blanket license, as we see it, is not a "naked restrain[t] of trade with no purpose except stifling of competition," *White Motor Co. v. United States*, 372 U.S. 253, 263 (1963), but rather accompanies the integration of sales, monitoring, and enforcement against unauthorized copyright use. As we have already indicated, ASCAP and the blanket license developed together out of the practical situation in the marketplace: thousands of users, thousands of copyright owners, and millions of compositions. Most users want unplanned, rapid, and indemnified access to any and all of the repertory of compositions, and the owners want a reliable method of collecting for the use of their copyrights. Individual sales transactions in this industry are quite expensive, as would be individual monitoring and enforcement, especially in light of the resources of single composers. Indeed, as both the Court of Appeals and CBS recognize, the costs are prohibitive for licenses with individual radio stations, nightclubs, and restaurants, and it was in that milieu that the blanket license arose.

A middleman with a blanket license was an obvious necessity if the thousands of individual negotiations, a virtual impossibility, were to be avoided. Also, individual fees for the use of individual compositions would presuppose an intricate schedule of fees and uses, as well as a difficult and expensive reporting problem for the user and policing task for the copyright owner. Historically, the market for public-performance rights organized itself largely around the single-fee blanket license, which gave unlimited access to the repertory and reliable protection against infringement. When ASCAP's major and user-created competitor, BMI, came on the scene, it also turned to the blanket license.

With the advent of radio and television networks, market conditions changed, and the necessity for and advantages of a blanket license for those users may be far less obvious than is the case when the potential users are individual television or radio stations, or the thousands of other individuals and organizations performing copyrighted compositions in public. But even for television network licenses, ASCAP reduces costs absolutely by creating a blanket license that is sold only a few, instead of thousands, of times, and that obviates the need for closely monitoring the networks to see that they do not use more than they pay for. ASCAP also provides the necessary resources for blanket sales and enforcement, resources unavailable to the vast majority of composers and publishing houses. Moreover, a bulk license of some type is a necessary consequence of the integration necessary to achieve these efficiencies, and a necessary consequence of an aggregate license is that its price must be established.

<div align="center">D</div>

This substantial lowering of costs, which is of course potentially beneficial to both sellers and buyers, differentiates the blanket license from individual use licenses. The blanket license is composed of the individual compositions plus the aggregating service.

Here, the whole is truly greater than the sum of its parts; it is, to some extent, a different product. The blanket license has certain unique characteristics: It allows the licensee immediate use of covered compositions, without the delay of prior individual negotiations and great flexibility in the choice of musical material. Many consumers clearly prefer the characteristics and cost advantages of this marketable package, and even small-performing rights societies that have occasionally arisen to compete with ASCAP and BMI have offered blanket licenses. Thus, to the extent the blanket license is a different product, ASCAP is not really a joint sales agency offering the individual goods of many sellers, but is a separate seller offering its blanket license, of which the individual compositions are raw material.[11] ASCAP, in short, made a market in which individual composers are inherently unable to compete fully effectively.

E

Finally, we have some doubt — enough to counsel against application of the *per se* rule — about the extent to which this practice threatens the "central nervous system of the economy," *United States v. Socony-Vacuum Oil Co.*, 310 U.S. 150, 226 n.59 (1940), that is, competitive pricing as the free market's means of allocating resources. Not all arrangements among actual or potential competitors that have an impact on price are *per se* violations of the Sherman Act or even unreasonable restraints. Mergers among competitors eliminate competition, including price competition, but they are not *per se* illegal, and many of them withstand attack under any existing antitrust standard. Joint ventures and other cooperative arrangements are also not usually unlawful, at least not as price-fixing schemes, where the agreement on price is necessary to market the product at all.

Here, the blanket-license fee is not set by competition among individual copyright owners, and it is a fee for the use of any of the compositions covered by the license. But the blanket license cannot be wholly equated with a simple horizontal arrangement among competitors. ASCAP does set the price for its blanket license, but that license is quite different from anything any individual owner could issue. The individual composers and authors have neither agreed not to sell individually in any other market nor use the blanket license to mask price fixing in such other markets. Moreover, the substantial restraints placed on ASCAP and its members by the consent decree must not be ignored. The District Court found that there was no legal, practical, or conspiratorial impediment to CBS's obtaining individual licenses; CBS, in short, had a real choice.

With this background in mind, which plainly enough indicates that over the years, and in the face of available alternatives, the blanket license has provided an acceptable mechanism for at least a large part of the market for the performing rights to copyrighted musical compositions, we cannot agree that it should automatically be declared illegal in all of its many manifestations. Rather, when attacked, it should be subjected to a more discriminating examination under the rule of reason. It may not ultimately survive that attack, but that is not the issue before us today.

IV

As we have noted, the enigmatic remarks of the Court of Appeals with respect to remedy appear to have departed from the court's strict, *per se* approach and to have invited

---

11. Moreover, because of the nature of the product — a composition can be simultaneously "consumed" by many users — composers have numerous markets and numerous incentives to produce, so the blanket license is unlikely to cause a decreased output, one of the normal undesirable effects of a cartel. And since popular songs get an increased share of ASCAP's revenue distributions, composers compete even within the blanket license in terms of productivity and consumer satisfaction.

a more careful analysis. But this left the general import of its judgment that the licensing practices of ASCAP and BMI under the consent decree are *per se* violations of the Sherman Act. We reverse that judgment, ... and remand for further proceedings to consider any unresolved issues that CBS may have properly brought to the Court of Appeals. Of course, this will include an assessment under the rule of reason of the blanket license as employed in the television industry, if that issue was preserved by CBS in the Court of Appeals.

The judgment of the Court of Appeals is reversed, and the cases are remanded to that court for further proceedings consistent with this opinion.

It is so ordered.

Mr. Justice STEVENS, dissenting.

The Court holds that ASCAP's blanket license is not a species of price fixing categorically forbidden by the Sherman Act. I agree with that holding. The Court remands the cases to the Court of Appeals, leaving open the question whether the blanket license as employed by ASCAP and BMI is unlawful under a rule-of-reason inquiry. I think that question is properly before us now and should be answered affirmatively.

* * *

I

* * *

[I]t cannot seriously be questioned ... that ASCAP and BMI have steadfastly adhered to the policy of only offering overall blanket or per-program licenses, notwithstanding requests for more limited authorizations.... It is the refusal to license anything less than the entire repertoire — rather than the decision to offer blanket licenses themselves — that raises the serious antitrust questions in this case.

* * *

Under our prior cases, there would be no question about the illegality of the blanket-only licensing policy if ASCAP and BMI were the exclusive sources of all licenses....

* * *

... But ... ASCAP does not have exclusive control of the copyrights in its portfolio, and it is perfectly possible — at least as a legal matter — for a user of music to negotiate directly with composers and publishers for whatever rights he may desire. The availability of a practical alternative alters the competitive effect of a ... blanket-licensing policy....

[Whether such a policy is unlawful] depends on an evaluation of the effect of the practice on competition in the relevant market. And, of course, it is well settled that a sales practice that is permissible for a small vendor, at least when no coercion is present, may be unreasonable when employed by a company that dominates the market. We therefore must consider what the record tells us about the competitive character of this market.

III

The market for music at issue here is wholly dominated by ASCAP-issued blanket licenses.[12] Virtually every domestic copyrighted composition is in the repertoire of either ASCAP or BMI. And again, virtually without exception, the only means that has been used to secure authority to perform such compositions is the blanket license.

The blanket all-or-nothing license is patently discriminatory. The user purchases full access to ASCAP's entire repertoire, even though his needs could be satisfied by a far more

---

12. As in the majority opinion, my references to ASCAP generally encompass BMI as well.

limited selection. The price he pays for this access is unrelated either to the quantity or the quality of the music he actually uses, or, indeed, to what he would probably use in a competitive system. Rather, in this unique all-or-nothing system, the price is based on a percentage of the user's advertising revenues, a measure that reflects the customer's ability to pay but is totally unrelated to factors—such as the cost, quality, or quantity of the product—that normally affect price in a competitive market. The ASCAP system requires users to buy more music than they want at a price which, while not beyond their ability to pay and perhaps not even beyond what is "reasonable" for the access they are getting, may well be far higher than what they would choose to spend for music in a competitive system. It is a classic example of economic discrimination.

The record plainly establishes that there is no price competition between separate musical compositions. Under a blanket license, it is no more expensive for a network to play the most popular current hit in prime time than it is to use an unknown composition as background music in a soap opera. Because the cost to the user is unaffected by the amount used on any program or on all programs, the user has no incentive to economize by, for example, substituting what would otherwise be less expensive songs for established favorites or by reducing the quantity of music used on a program. The blanket license thereby tends to encourage the use of more music, and also of a larger share of what is really more valuable music, than would be expected in a competitive system characterized by separate licenses. And since revenues are passed on to composers on a basis reflecting the character and frequency of the use of their music, the tendency is to increase the rewards of the established composers at the expense of those less well known. Perhaps the prospect is in any event unlikely, but the blanket license does not present a new songwriter with any opportunity to try to break into the market by offering his product for sale at an unusually low price. The absence of that opportunity, however unlikely it may be, is characteristic of a cartelized rather than a competitive market.

The current state of the market cannot be explained on the ground that it could not operate competitively, or that issuance of more limited—and thus less restrictive—licenses by ASCAP is not feasible. The District Court's findings disclose no reason why music-performing rights could not be negotiated on a per-composition or per-use basis, either with the composer or publisher directly or with an agent such as ASCAP. In fact, ASCAP now compensates composers and publishers on precisely those bases. If distributions of royalties can be calculated on a per-use and per-composition basis, it is difficult to see why royalties could not also be collected in the same way. Moreover, the record also shows that where ASCAP's blanket-license scheme does not govern, competitive markets do. A competitive market for "synch" rights exists,[13] and after the use of blanket licenses in the motion picture industry was discontinued, such a market promptly developed in that industry. In sum, the record demonstrates that the market at issue here is one that could be highly competitive, but is not competitive at all.

IV

Since the record describes a market that could be competitive and is not, and since that market is dominated by two firms engaged in a single, blanket method of dealing, it surely seems logical to conclude that trade has been restrained unreasonably. ASCAP argues, however, that at least as to CBS, there has been no restraint at all since the network is free to deal directly with copyright holders.

---

13. The "synch" right is the right to record a copyrighted song in synchronization with the film or videotape, and is obtained separately from the right to perform the music. It is the latter which is controlled by ASCAP and BMI.

The District Court found that CBS had failed to establish that it was compelled to take a blanket license from ASCAP. While CBS introduced evidence suggesting that a significant number of composers and publishers, satisfied as they are with the ASCAP system, would be "disinclined" to deal directly with the network, the court found such evidence unpersuasive in light of CBS's substantial market power in the music industry and the importance to copyright holders of network television exposure. Moreover, it is arguable that CBS could go further and, along with the other television networks, use its economic resources to exploit destructive competition among purveyors of music by driving the price of performance rights down to a far lower level. But none of this demonstrates that ASCAP's practices are lawful, or that ASCAP cannot be held liable for injunctive relief at CBS's request.

* * *

Far from establishing ASCAP's immunity from liability, these District Court findings, in my judgment, confirm the illegality of its conduct. Neither CBS nor any other user has been willing to assume the costs and risks associated with an attempt to purchase music on a competitive basis. The fact that an attempt by CBS to break down the ASCAP monopoly might well succeed does not preclude the conclusion that smaller and less powerful buyers are totally foreclosed from a competitive market. Despite its size, CBS itself may not obtain music on a competitive basis without incurring unprecedented costs and risks. The fear of unpredictable consequences, coupled with the certain and predictable costs and delays associated with a change in its method of purchasing music, unquestionably inhibits any CBS management decision to embark on a competitive crusade. Even if ASCAP offered CBS a special bargain to forestall any such crusade, that special arrangement would not cure the marketwide restraint.

Whatever management decision CBS should or might have made, it is perfectly clear that the question whether competition in the market has been unduly restrained is not one that any single company's management is authorized to answer.... Even without judicial intervention, the ASCAP monopoly might eventually be broken by CBS, if the benefits of doing so outweigh the significant costs and risks involved in commencing direct dealing. But that hardly means that the blanket-licensing policy at issue here is lawful. An arrangement that produces marketwide price discrimination and significant barriers to entry unreasonably restrains trade even if the discrimination and the barriers have only a limited life expectancy. History suggests, however, that these restraints have an enduring character.

Antitrust policy requires that great aggregations of economic power be closely scrutinized. That duty is especially important when the aggregation is composed of statutory monopoly privileges. Our cases have repeatedly stressed the need to limit the privileges conferred by patent and copyright strictly to the scope of the statutory grant. The record in this case plainly discloses that the limits have been exceeded and that ASCAP and BMI exercise monopoly powers that far exceed the sum of the privileges of the individual copyright holders. Indeed, ASCAP itself argues that its blanket license constitutes a product that is significantly different from the sum of its component parts. I agree with that premise, but I conclude that the aggregate is a monopolistic restraint of trade proscribed by the Sherman Act.

Notes and Questions

1. The Competitive Virtues of Blanket Licensing: Have no doubt that blanket licensing, where a single price is set on access to millions of songs, represents a price fixing agreement

between, not one or two competitors, but as the Court put it "some 10,000 publishing companies and 20,000 authors and composers." However the Court, instead of condemning this behavior as one of the largest price fixing schemes in history, endorsed the behavior as beneficial to consumers. In deciding whether to find blanket licenses *per se* unlawful, the Court takes note of the following quote from *United States v. Topco Associates Inc*: "[i]t is only after considerable experience with certain business relationships that courts classify them as *per se* violations." The Court concludes that it has never dealt with practices like those of ASCAP and BMI. However, it finds that the fact that the Antitrust Division of the Department of Justice has issued a consent decree governing the now-challenged practices "[is] a unique indicator that the challenged practice may have some redeeming competitive virtues." What are the redeeming virtues of ASCAP and BMI's challenged practices? The free market is suppressed in favor of some other structural method for making price and output decisions. The blanket licensing agreement in BMI provided an alternative to free market transactions between each individual copyright holder and each firm that wanted to use the copyrighted material. It is possible and in fact likely that ASCAP and BMI were developed and are maintainable because their method for making price and output decisions for copyrighted material is superior to a free market. How can this be?

2. *A New Product Created with the Least Restrictive Means:* The Court notes that in this scheme "the whole is truly greater than the sum of its parts; it is, to some extent, a different product." Why should this matter? The Court draws a line between fixing the price of pre-existing products and creating an entirely new one. So long as the pricing agreement creates something new then the application of the antitrust laws is less strict and uses the more liberal rule of reason. Why should this be so?

The Court also finds it important that "there was no legal, practical, or conspiratorial impediment to CBS's obtaining individual licenses" directly from the artists or publishing companies. If the blanket license had restricted the ability of the artists to sell licenses to their copyrights individually would this case have come out differently?

3. *Transaction Costs Economics:* The field of **transaction costs economics** studies the costs of carrying out certain transactions based upon the type of transaction and the way such transactions are organized. **Transaction costs** are the costs of time, information, contracting and monitoring that are required to execute an exchange. If the goal is to maximize the efficiency by which transactions are made, then it makes sense to utilize the structure that minimizes transaction costs. In this light, when transaction costs are minimized by allocation through a free market, then that structure will be used. However, when transaction costs are minimized by using some other formal structure, then a free market may no longer be superior. The Court finds that a blanket fee lowers transaction costs because the costs of negotiating with thousands of different artists would be prohibitive. What are the transaction costs involved in buying the right to use copyrighted material? Would market exchange or ASCAP's current method minimize these transaction costs? Can you think of an even more efficient method for the copyright use market?

4. *Developing the Rights of Copyright Holders:* The benefits of intellectual property rights were discussed just prior to the *BMI* case. As a result of these benefits, Congress has enacted patent and copyright laws. These laws are supposed to encourage the development of new creations and information by granting a time constrained monopoly. Moreover, these laws also protect the rights of the creators to their monopoly. Part of the Court's decision in *BMI* relied on the fact that, absent ASCAP and BMI's blanket licensing agreement, copyright holders would have a difficult time marketing their products and enforcing their rights against infringement. That is, individual negotiation and enforcement

would be very costly both to the copyright holders and users. Thus, the blanket licensing agreement seems to encourage the sale and use of copyrighted material in conjunction with a cost effective method for enforcing the copyrights. Moreover, those copyright holders or potential users who found individual negotiation to be more cost efficient were not foreclosed from using market mechanisms. In other words, the Court found the blanket licensing price restraint to be a necessary component in further developing the goals of copyright laws. Can you think of an alternative for marketing individual copyrights that is more in line with the model of pure competition? Will technology have an effect on the necessity of the blanket licensing agreement in supporting trade? Is there any other way for individual copyright holders to enforce their monopoly rights?

5. "Inherently Suspect" Restraints and the Rule of Reason: Some arrangements among competitors fall short of the naked price fixing or market allocations subject to per se condemnation, but nonetheless pose serious competitive consequences. In these circumstances, courts require defendants to explain the efficiency justifications for the agreement before the court proceeds to a rule of reason balancing. If the defendant cannot carry its burden of persuasion, the practice is condemned. Courts have employed this analysis in restraints involving advertising. See Polygram Holding, Inc. v. F.T.C., 416 F.3d 29 (D.C. Cir. 2005) and Realcomp II, Ltd. v. F.T.C., 635 F.3d 815 (6th Cir. 2011). Does this approach make sense?