CHAPTER FIFTEEN

Imperfect Competition

This chapter discusses oligopoly markets, falling between the extremes of perfect competition and monopoly.

DEFINITION

Oligopoly. A market with relatively few firms but more than one.

Oligopolies raise the possibility of strategic interaction among firms. To analyze this strategic interaction rigorously, we will apply the concepts from game theory that were introduced in Chapter 8. Our game-theoretic analysis will show that small changes in details concerning the variables firms choose, the timing of their moves, or their information about market conditions or rival actions can have a dramatic effect on market outcomes. The first half of the chapter deals with short-term decisions such as pricing and output, and the second half covers longer-term decisions such as investment, advertising, and entry.

SHORT-RUN DECISIONS: PRICING AND OUTPUT

It is difficult to predict exactly the possible outcomes for price and output when there are few firms; prices depend on how aggressively firms compete, which in turn depends on which strategic variables firms choose, how much information firms have about rivals, and how often firms interact with each other in the market.

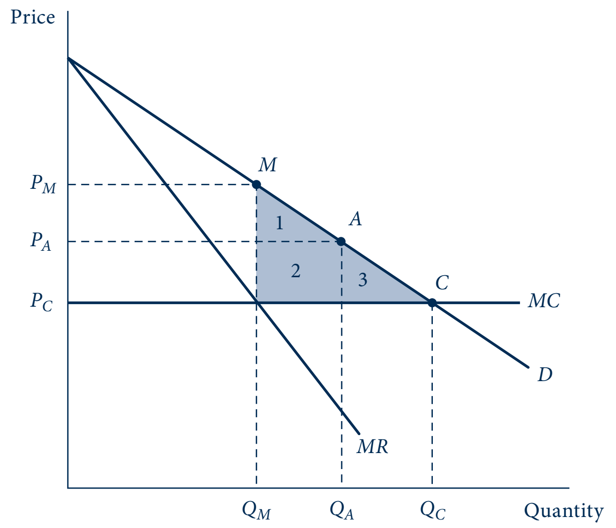
For example, consider the Bertrand game studied in the next section. The game involves two identical firms choosing prices simultaneously for their identical products in their one meeting in the market. The Bertrand game has a Nash equilibrium at point *C* in Figure 15.1. Even though there may be only two firms in the market, in this equilibrium they behave as though they were perfectly competitive, setting price equal to marginal cost and earning zero profit. We will discuss whether the Bertrand game is a realistic depiction of actual firm behavior, but an analysis of the model shows that it is possible to think up rigorous game-theoretic models in which one extreme—the competitive outcome—can emerge in concentrated markets with few firms.

At the other extreme, as indicated by point *M* in Figure 15.1, firms as a group may act as a cartel, recognizing that they can affect price and coordinate their decisions. Indeed, they may be able to act as a perfect cartel and achieve the highest possible profits—namely, the profit a monopoly would earn in the market. One way to maintain a cartel is to bind firms with explicit pricing rules. Such explicit pricing rules are often prohibited by antitrust law. But firms need not resort to explicit pricing rules if they interact on the market repeatedly; they can collude tacitly. High collusive prices can be maintained with

FIGURE 15.1

Pricing and Output under Imperfect Competition

Market equilibrium under imperfect competition can occur at many points on the demand curve. In the figure, which assumes that marginal costs are constant over all output ranges, the equilibrium of the Bertrand game occurs at point *C*, also corresponding to the perfectly competitive outcome. The perfect cartel outcome occurs at point *M*, also corresponding to the monopoly outcome. Many solutions may occur between points *M* and *C*, depending on the specific assumptions made about how firms compete. For example, the equilibrium of the Cournot game might occur at a point such as *A*. The deadweight loss given by the shaded triangle increases as one moves from point *C* to *M*.



the tacit threat of a price war if any firm undercuts. We will analyze this game formally and discuss the difficulty of maintaining collusion.

The Bertrand and cartel models determine the outer limits between which actual prices in an imperfectly competitive market are set (one such intermediate price is represented by point *A* in Figure 15.1). This band of outcomes may be wide, and given the plethora of available models there may be a model for nearly every point within the band. For example, in a later section we will show how the Cournot model, in which firms set quantities rather than prices as in the Bertrand model, leads to an outcome (such as point *A*) somewhere between *C* and *M* in Figure 15.1.

It is important to know where the industry is on the line between points *C* and *M* because total welfare (as measured by the sum of consumer surplus and firms' profits; see Chapter 12) depends on the location of this point. At point *C*, total welfare is as high as possible; at point *A*, total welfare is lower by the area of the shaded triangle 3. In Chapter 12, this shortfall in total welfare relative to the highest possible level was called *deadweight loss*. At point *M*, deadweight loss is even greater and is given by the area of shaded regions 1, 2, and 3. The closer the imperfectly competitive outcome to *C* and the farther from *M*, the higher is total welfare and the better off society will be.[1]

[1]Because this section deals with short-run decision variables (price and quantity), the discussion of total welfare in this paragraph focuses on short-run considerations. As discussed in a later section, an imperfectly competitive market may produce considerably more deadweight loss than a perfectly competitive one in the short run yet provide more innovation incentives, leading to lower production costs and new products and perhaps higher total welfare in the long run. The patent system intentionally impairs competition by granting a monopoly right to improve innovation incentives.

BERTRAND MODEL

The Bertrand model is named after the economist who first proposed it.[2] The model is a game involving two identical firms, labeled 1 and 2, producing identical products at a constant marginal cost (and constant average cost) c. The firms choose prices p_1 and p_2 simultaneously in a single period of competition. Because firms' products are perfect substitutes, all sales go to the firm with the lowest price. Sales are split evenly if $p_1 = p_2$. Let $D(p)$ be market demand.

We will look for the Nash equilibrium. The game has a continuum of actions, as does Example 8.5 (the Tragedy of the Commons) in Chapter 8. Unlike Example 8.5, we cannot use calculus to derive best-response functions because the profit functions are not differentiable here. Starting from equal prices, if one firm lowers its price by the smallest amount, then its sales and profit would essentially double. We will proceed by first guessing what the Nash equilibrium is and then spending some time to verify that our guess was in fact correct.

Nash equilibrium of the Bertrand game

The only pure-strategy Nash equilibrium of the Bertrand game is $p_1^* = p_2^* = c$. That is, the Nash equilibrium involves both firms charging marginal cost. In saying that this is the only Nash equilibrium, we are making two statements that need to be verified: This outcome is a Nash equilibrium, and there is no other Nash equilibrium.

To verify that this outcome is a Nash equilibrium, we need to show that both firms are playing a best response to each other—or, in other words, that neither firm has an incentive to deviate to some other strategy. In equilibrium, firms charge a price equal to marginal cost, which in turn is equal to average cost. But a price equal to average cost means firms earn zero profit in equilibrium. Can a firm earn more than the zero it earns in equilibrium by deviating to some other price? No. If it deviates to a higher price, then it will make no sales and therefore no profit, not strictly more than in equilibrium. If it deviates to a lower price, then it will make sales but will be earning a negative margin on each unit sold because price would be below marginal cost. Thus, the firm would earn negative profit, less than in equilibrium. Because there is no possible profitable deviation for the firm, we have succeeded in verifying that both firms' charging marginal cost is a Nash equilibrium.

It is clear that marginal cost pricing is the only pure-strategy Nash equilibrium. If prices exceeded marginal cost, the high-price firm would gain by undercutting the other slightly and capturing all the market demand. More formally, to verify that $p_1^* = p_2^* = c$ is the only Nash equilibrium, we will go one by one through an exhaustive list of cases for various values of p_1, p_2, and c, verifying that none besides $p_1 = p_2 = c$ is a Nash equilibrium. To reduce the number of cases, assume firm 1 is the low-price firm—that is, $p_1 \le p_2$. The same conclusions would be reached taking 2 to be the low-price firm.

There are three exhaustive cases: (i) $c > p_1$, (ii) $c < p_1$, and (iii) $c = p_1$. Case (i) cannot be a Nash equilibrium. Firm 1 earns a negative margin $p_1 - c$ on every unit it sells, and because it makes positive sales, it must earn negative profit. It could earn higher profit by deviating to a higher price. For example, firm 1 could guarantee itself zero profit by deviating to $p_1 = c$.

Case (ii) cannot be a Nash equilibrium either. At best, firm 2 gets only half of market demand (if $p_1 = p_2$) and at worst gets no demand (if $p_1 < p_2$). Firm 2 could capture all the market demand by undercutting firm 1's price by a tiny amount ε. This ε could be

[2] J. Bertrand, "Théorie Mathematique de la Richess Sociale," *Journal de Savants* (1883): 499–508.

chosen small enough that market price and total market profit are hardly affected. If $p_1 = p_2$ before the deviation, the deviation would essentially double firm 2's profit. If $p_1 < p_2$ before the deviation, the deviation would result in firm 2 moving from zero to positive profit. In either case, firm 2's deviation would be profitable.

Case (iii) includes the subcase of $p_1 = p_2 = c$, which we saw is a Nash equilibrium. The only remaining subcase in which $p_1 \leq p_2$ is $c = p_1 < p_2$. This subcase cannot be a Nash equilibrium: Firm 1 earns zero profit here but could earn positive profit by deviating to a price slightly above c but still below p_2.

Although the analysis focused on the game with two firms, it is clear that the same outcome would arise for any number of firms $n \geq 2$. The Nash equilibrium of the n-firm Bertrand game is $p_1^* = p_2^* = \cdots = p_n^* = c$.

Bertrand paradox

The Nash equilibrium of the Bertrand model is the same as the perfectly competitive outcome. Price is set to marginal cost, and firms earn zero profit. This result—that the Nash equilibrium in the Bertrand model is the same as in perfect competition even though there may be only two firms in the market—is called the *Bertrand paradox*. It is paradoxical that competition between as few as two firms would be so tough. The Bertrand paradox is a general result in the sense that we did not specify the marginal cost c or the demand curve; therefore, the result holds for any c and any downward-sloping demand curve.

In another sense, the Bertrand paradox is not general; it can be undone by changing various of the model's other assumptions. Each of the next several sections will present a different model generated by changing a different one of the Bertrand assumptions. In the next section, for example, we will assume that firms choose quantity rather than price, leading to what is called the *Cournot game*. We will see that firms do not end up charging marginal cost and earning zero profit in the Cournot game. In subsequent sections, we will show that the Bertrand paradox can also be avoided if still other assumptions are changed: if firms face capacity constraints rather than being able to produce an unlimited amount at cost c, if products are slightly differentiated rather than being perfect substitutes, or if firms engage in repeated interaction rather than one round of competition.

COURNOT MODEL

The Cournot model, named after the economist who proposed it,[3] is similar to the Bertrand model except that firms are assumed to simultaneously choose quantities rather than prices. As we will see, this simple change in strategic variable will lead to a big change in implications. Price will be above marginal cost, and firms will earn positive profit in the Nash equilibrium of the Cournot game. It is somewhat surprising (but nonetheless an important point to keep in mind) that this simple change in choice variable matters in the strategic setting of an oligopoly when it did not matter with a monopoly: The monopolist obtained the same profit-maximizing outcome whether it chose prices or quantities.

We will start with a general version of the Cournot game with n firms indexed by $i = 1, \ldots, n$. Each firm chooses its output q_i of an identical product simultaneously. The outputs are combined into a total industry output $Q = q_1 + q_2 + \cdots + q_n$,

[3] A. Cournot, *Researches into the Mathematical Principles of the Theory of Wealth*, trans. N. T. Bacon (New York: Macmillan, 1897). Although the Cournot model appears after Bertrand's in this chapter, Cournot's work, originally published in 1838, predates Bertrand's. Cournot's work is one of the first formal analyses of strategic behavior in oligopolies, and his solution concept anticipated Nash equilibrium.

resulting in market price $P(Q)$. Observe that $P(Q)$ is the inverse demand curve corresponding to the market demand curve $Q = D(P)$. Assume market demand is downward sloping and so inverse demand is, too; that is, $P'(Q) < 0$. Firm i's profit equals its total revenue, $P(Q)q_i$, minus its total cost, $C_i(q_i)$:

$$\pi_i = P(Q)q_i - C_i(q_i). \tag{15.1}$$

Nash equilibrium of the Cournot game

Unlike the Bertrand game, the profit function (15.1) in the Cournot game is differentiable; hence we can proceed to solve for the Nash equilibrium of this game just as we did in Example 8.5, the Tragedy of the Commons. That is, we find each firm i's best response by taking the first-order condition of the objective function (15.1) with respect to q_i:

$$\frac{\partial \pi_i}{\partial q_i} = \underbrace{P(Q) + P'(Q)q_i}_{MR} - \underbrace{C_i'(q_i)}_{MC} = 0. \tag{15.2}$$

Equation 15.2 must hold for all $i = 1, \ldots, n$ in the Nash equilibrium.

According to Equation 15.2, the familiar condition for profit maximization from Chapter 11—marginal revenue (MR) equals marginal cost (MC)—holds for the Cournot firm. As we will see from an analysis of the particular form that the marginal revenue term takes for the Cournot firm, price is above the perfectly competitive level (above marginal cost) but below the level in a perfect cartel that maximizes firms' joint profits.

In order for Equation 15.2 to equal 0, price must exceed marginal cost by the magnitude of the "wedge" term $P'(Q)q_i$. If the Cournot firm produces another unit on top of its existing production of q_i units, then, because demand is downward sloping, the additional unit causes market price to decrease by $P'(Q)$, leading to a loss of revenue of $P'(Q)q_i$ (the wedge term) from firm i's existing production.

To compare the Cournot outcome with the perfect cartel outcome, note that the objective for the cartel is to maximize joint profit:

$$\sum_{j=1}^{n} \pi_j = P(Q)\sum_{j=1}^{n} q_j - \sum_{j=1}^{n} C_j(q_j). \tag{15.3}$$

Taking the first-order condition of Equation 15.3 with respect to q_i gives

$$\frac{\partial}{\partial q_i}\left(\sum_{j=1}^{n} \pi_j\right) = \underbrace{P(Q) + P'(Q)\sum_{j=1}^{n} q_j}_{MR} - \underbrace{C_i'(q_i)}_{MC} = 0. \tag{15.4}$$

This first-order condition is similar to Equation 15.2 except that the wedge term,

$$P'(Q)\sum_{j=1}^{n} q_j = P'(Q)Q, \tag{15.5}$$

is larger in magnitude with a perfect cartel than with Cournot firms. In maximizing joint profits, the cartel accounts for the fact that an additional unit of firm i's output, by reducing market price, reduces the revenue earned on *all* firms' existing output. Hence $P'(Q)$ is multiplied by total cartel output Q in Equation 15.5. The Cournot firm accounts for the reduction in revenue only from its own existing output q_i. Hence Cournot firms will end up overproducing relative to the joint profit-maximizing outcome. That is, the extra production in the Cournot outcome relative to a perfect cartel will end up in lower joint profit for the firms. What firms would regard as overproduction is good for society because it means that the Cournot outcome (point A, referring back to Figure 15.1) will involve more total welfare than the perfect cartel outcome (point M in Figure 15.1).

EXAMPLE 15.1 Natural-Spring Duopoly

As a numerical example of some of these ideas, we will consider a case with just two firms and simple demand and cost functions. Following Cournot's nineteenth-century example of two natural springs, we assume that each spring owner has a large supply of (possibly healthful) water and faces the problem of how much to provide the market. A firm's cost of pumping and bottling q_i liters is $C_i(q_i) = cq_i$, implying that marginal costs are a constant c per liter. Inverse demand for spring water is

$$P(Q) = a - Q, \tag{15.6}$$

where a is the demand intercept (measuring the strength of spring water demand) and $Q = q_1 + q_2$ is total spring water output. We will now examine various models of how this market might operate.

Bertrand model. In the Nash equilibrium of the Bertrand game, the two firms set price equal to marginal cost. Hence market price is $P^* = c$, total output is $Q^* = a - c$, firm profit is $\pi_i^* = 0$, and total profit for all firms is $\Pi^* = 0$. For the Bertrand quantity to be positive we must have $a > c$, which we will assume throughout the problem.

Cournot model. The solution for the Nash equilibrium follows Example 8.6 closely. Profits for the two Cournot firms are

$$\begin{aligned} \pi_1 &= P(Q)q_1 - cq_1 = (a - q_1 - q_2 - c)q_1, \\ \pi_2 &= P(Q)q_2 - cq_2 = (a - q_1 - q_2 - c)q_2. \end{aligned} \tag{15.7}$$

Using the first-order conditions to solve for the best-response functions, we obtain

$$q_1 = \frac{a - q_2 - c}{2}, \qquad q_2 = \frac{a - q_1 - c}{2}. \tag{15.8}$$

Solving Equations 15.8 simultaneously yields the Nash equilibrium

$$q_1^* = q_2^* = \frac{a - c}{3}. \tag{15.9}$$

Thus, total output is $Q^* = (2/3)(a - c)$. Substituting total output into the inverse demand curve implies an equilibrium price of $P^* = (a + 2c)/3$. Substituting price and outputs into the profit functions (Equations 15.7) implies $\pi_1^* = \pi_2^* = (1/9)(a - c)^2$, so total market profit equals $\Pi^* = \pi_1^* = \pi_2^* = (2/9)(a - c)^2$.

Perfect cartel. The objective function for a perfect cartel involves joint profits

$$\pi_1 + \pi_2 = (a - q_1 - q_2 - c)q_1 + (a - q_1 - q_2 - c)q_2. \tag{15.10}$$

The two first-order conditions for maximizing Equation 15.10 with respect to q_1 and q_2 are the same:

$$\frac{\partial}{\partial q_1}(\pi_1 + \pi_2) = \frac{\partial}{\partial q_2}(\pi_1 + \pi_2) = a - 2q_1 - 2q_2 - c = 0. \tag{15.11}$$

The first-order conditions do not pin down market shares for firms in a perfect cartel because they produce identical products at constant marginal cost. But Equation 15.11 does pin down total output: $q_1^* + q_2^* = Q^* = (1/2)(a - c)$. Substituting total output into inverse demand implies that the cartel price is $P^* = (1/2)(a + c)$. Substituting price and quantities into Equation 15.10 implies a total cartel profit of $\Pi^* = (1/4)(a - c)^2$.

Comparison. Moving from the Bertrand model to the Cournot model to a perfect cartel, because $a > c$ we can show that quantity Q^* decreases from $a - c$ to $(2/3)(a - c)$ to $(1/2)(a - c)$. It can also be shown that price P^* and industry profit Π^* increase. For example, if $a = 120$ and $c = 0$ (implying that inverse demand is $P(Q) = 120 - Q$ and that production is costless), then market quantity is 120 with Bertrand competition, 80 with Cournot competition, and 60 with a perfect cartel. Price increases from 0 to 40 to 60 across the cases, and industry profit increases from 0 to 3,200 to 3,600.

QUERY: In a perfect cartel, do firms play a best response to each other's quantities? If not, in which direction would they like to change their outputs? What does this say about the stability of cartels?

EXAMPLE 15.2 Cournot Best-Response Diagrams

Continuing with the natural-spring duopoly from Example 15.1, it is instructive to solve for the Nash equilibrium using graphical methods. We will graph the best-response functions given in Equation 15.8; the intersection between the best responses is the Nash equilibrium. As background, you may want to review a similar diagram (Figure 8.8) for the Tragedy of the Commons.

The linear best-response functions are most easily graphed by plotting their intercepts, as shown in Figure 15.2. The best-response functions intersect at the point $q_1^* = q_2^* = (a - c)/3$, which was the Nash equilibrium of the Cournot game computed using algebraic methods in Example 15.1.

FIGURE 15.2 Best-Response Diagram for Cournot Duopoly

Firms' best responses are drawn as thick lines; their intersection (E) is the Nash equilibrium of the Cournot game. Isoprofit curves for firm 1 increase until point M is reached, which is the monopoly outcome for firm 1.

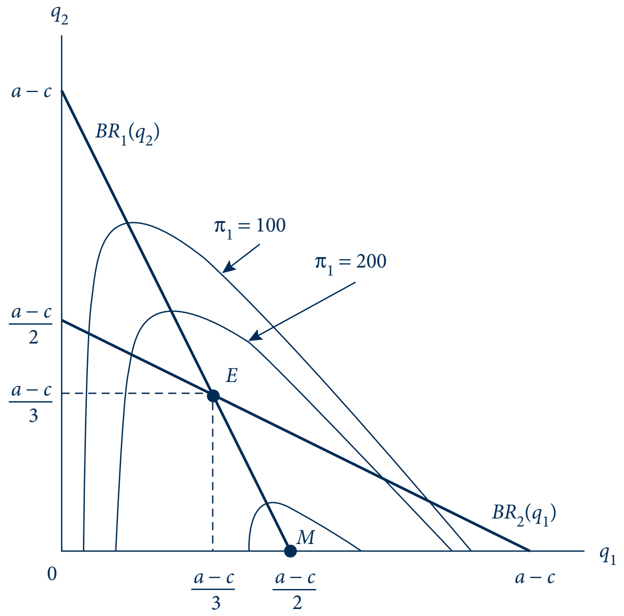


Figure 15.2 displays firms' isoprofit curves. An *isoprofit curve* for firm 1 is the locus of quantity pairs providing it with the same profit level. To compute the isoprofit curve associated with a profit level of (say) 100, we start by setting Equation 15.7 equal to 100:

$$\pi_1 = (a - q_1 - q_2 - c)q_1 = 100. \tag{15.12}$$

Then we solve for q_2 to facilitate graphing the isoprofit:

$$q_2 = a - c - q_1 - \frac{100}{q_1}. \tag{15.13}$$

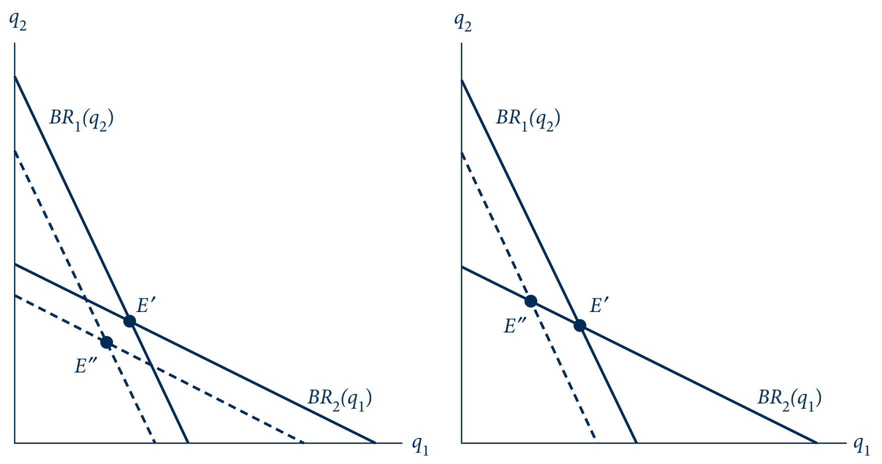
Several example isoprofits for firm 1 are shown in the figure. As profit increases from 100 to 200 to yet higher levels, the associated isoprofits shrink down to the monopoly point, which is the highest isoprofit on the diagram. To understand why the individual isoprofits are shaped like frowns, refer back to Equation 15.13. As q_1 approaches 0, the last term $(-100/q_1)$ dominates, causing the left side of the frown to turn down. As q_1 increases, the $-q_1$ term in Equation 15.13 begins to dominate, causing the right side of the frown to turn down.

Figure 15.3 shows how to use best-response diagrams to quickly tell how changes in such underlying parameters as the demand intercept a or marginal cost c would affect the equilibrium. Figure 15.3a depicts an increase in both firms' marginal cost c. The best responses shift inward, resulting in a new equilibrium that involves lower output for both. Although firms have the same marginal cost in this example, one can imagine a model in which firms have different marginal cost parameters and so can be varied independently. Figure 15.3b depicts an increase in just firm 1's marginal cost; only firm 1's best response shifts. The new equilibrium involves lower output for firm 1 and higher output for firm 2. Although firm 2's best response does not shift, it still increases its output as it anticipates a reduction in firm 1's output and best responds to this anticipated output reduction.

QUERY: Explain why firm 1's individual isoprofits reach a peak on its best-response function in Figure 15.2. What would firm 2's isoprofits look like in Figure 15.2? How would you represent an increase in demand intercept a in Figure 15.3?

FIGURE 15.3 Shifting Cournot Best Responses

Firms' initial best responses are drawn as solid lines, resulting in a Nash equilibrium at point E'. Panel (a) depicts an increase in both firms' marginal costs, shifting their best responses—now given by the dashed lines—inward. The new intersection point, and thus the new equilibrium, is point E''. Panel (b) depicts an increase in just firm 1's marginal cost.



(a) Increase in both firms' marginal costs

(b) Increase in firm 1's marginal cost

Varying the number of Cournot firms

The Cournot model is particularly useful for policy analysis because it can represent the whole range of outcomes from perfect competition to perfect cartel/monopoly (i.e., the whole range of points between C and M in Figure 15.1) by varying the number of firms n from $n = \infty$ to $n = 1$. For simplicity, consider the case of identical firms, which here means the n firms sharing the same cost function $C(q_i)$. In equilibrium, firms will produce the same share of total output: $q_i = Q/n$. Substituting $q_i = Q/n$ into Equation 15.12, the wedge term becomes $P'(Q)Q/n$. The wedge term disappears as n grows large; firms become infinitesimally small. An infinitesimally small firm effectively becomes a price-taker because it produces so little that any decrease in market price from an increase in output hardly affects its revenue. Price approaches marginal cost and the market outcome approaches the perfectly competitive one. As n decreases to 1, the wedge term approaches that in Equation 15.5, implying the Cournot outcome approaches that of a perfect cartel. As the Cournot firm's market share grows, it internalizes the revenue loss from a decrease in market price to a greater extent.

EXAMPLE 15.3 Natural-Spring Oligopoly

Return to the natural springs in Example 15.1, but now consider a variable number n of firms rather than just two. The profit of one of them, firm i, is

$$\pi_i = P(Q)q_i - cq_i = (a - Q - c)q_i = (a - q_i - Q_{-i} - c)q_i. \tag{15.14}$$

It is convenient to express total output as $Q = q_i + Q_{-i}$, where $Q_{-i} = Q - q_i$ is the output of all firms except for i. Taking the first-order condition of Equation 15.14 with respect to q_i, we recognize that firm i takes Q_{-i} as a given and thus treats it as a constant in the differentiation,

$$\frac{\partial \pi_i}{\partial q_i} = a - 2q_i - Q_{-i} - c = 0, \tag{15.15}$$

which holds for all $i = 1, 2, \ldots, n$.

The key to solving the system of n equations for the n equilibrium quantities is to recognize that the Nash equilibrium involves equal quantities because firms are symmetric. Symmetry implies that

$$Q^*_{-i} = Q^* - q_i^* = nq_i^* - q_i^* = (n-1)q_i^*. \tag{15.16}$$

Substituting Equation 15.16 into 15.15 yields

$$a - 2q_i^* - (n-1)q_i^* - c = 0, \tag{15.17}$$

or $q_i^* = (a - c)/(n + 1)$.

Total market output is

$$Q^* = nq_i^* = \left(\frac{n}{n+1}\right)(a - c), \tag{15.18}$$

and market price is

$$P^* = a - Q^* = \left(\frac{1}{n+1}\right)a + \left(\frac{n}{n+1}\right)c. \tag{15.19}$$

Substituting for q_i^*, Q^*, and P^* into the firm's profit Equation 15.14, we have that total profit for all firms is

$$\Pi^* = n\pi_i^* = n\left(\frac{a-c}{n+1}\right)^2. \tag{15.20}$$

Setting $n = 1$ in Equations 15.18–15.20 gives the monopoly outcome, which gives the same price, total output, and profit as in the perfect cartel case computed in Example 15.1. Letting n grow without bound in Equations 15.18–15.20 gives the perfectly competitive outcome, the same outcome computed in Example 15.1 for the Bertrand case.

QUERY: We used the trick of imposing symmetry after taking the first-order condition for firm i's quantity choice. It might seem simpler to impose symmetry *before* taking the first-order condition. Why would this be a mistake? How would the incorrect expressions for quantity, price, and profit compare with the correct ones here?

Prices or quantities?

Moving from price competition in the Bertrand model to quantity competition in the Cournot model changes the market outcome dramatically. This change is surprising on first thought. After all, the monopoly outcome from Chapter 14 is the same whether we assume the monopolist sets price or quantity. Further thought suggests why price and quantity are such different strategic variables. Starting from equal prices, a small reduction in one firm's price allows it to steal all the market demand from its competitors. This sharp benefit from undercutting makes price competition extremely "tough." Quantity competition is "softer." Starting from equal quantities, a small increase in one firm's quantity has only a marginal effect on the revenue that other firms receive from their existing output. Firms have less of an incentive to outproduce each other with quantity competition than to undercut each other with price competition.

An advantage of the Cournot model is its realistic implication that the industry grows more competitive as the number n of firms entering the market increases from monopoly to perfect competition. In the Bertrand model there is a discontinuous jump from monopoly to perfect competition if just two firms enter, and additional entry beyond two has no additional effect on the market outcome.

An apparent disadvantage of the Cournot model is that firms in real-world markets tend to set prices rather than quantities, contrary to the Cournot assumption that firms choose quantities. For example, grocers advertise prices for orange juice, say, $3.00 a container, in newpaper circulars rather than the number of containers it stocks. As we will see in the next section, the Cournot model applies even to the orange juice market if we reinterpret quantity to be the firm's *capacity*, defined as the most the firm can sell given the capital it has in place and other available inputs in the short run.

CAPACITY CONSTRAINTS

For the Bertrand model to generate the Bertrand paradox (the result that two firms essentially behave as perfect competitors), firms must have unlimited capacities. Starting from equal prices, if a firm lowers its price the slightest amount, then its demand essentially doubles. The firm can satisfy this increased demand because it has no capacity constraints, giving firms a big incentive to undercut. If the undercutting firm could not serve all the demand at its lower price because of capacity constraints, that would leave some residual demand for the higher-priced firm and would decrease the incentive to undercut.

Consider a two-stage game in which firms build capacity in the first stage and firms choose prices p_1 and p_2 in the second stage.[4] Firms cannot sell more in the second stage

[4]The model is due to D. Kreps and J. Scheinkman, "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes," *Bell Journal of Economics* (Autumn 1983): 326–37.

than the capacity built in the first stage. If the cost of building capacity is sufficiently high, it turns out that the subgame-perfect equilibrium of this sequential game leads to the same outcome as the Nash equilibrium of the Cournot model.

To see this result, we will analyze the game using backward induction. Consider the second-stage pricing game supposing the firms have already built capacities $\bar{q}_1$ and $\bar{q}_2$ in the first stage. Let $\bar{p}$ be the price that would prevail when production is at capacity for both firms. A situation in which

$$p_1 = p_2 < \bar{p} \tag{15.21}$$

is not a Nash equilibrium. At this price, total quantity demanded exceeds total capacity; therefore, firm 1 could increase its profits by raising price slightly and continuing to sell $\bar{q}_1$. Similarly,

$$p_1 = p_2 > \bar{p} \tag{15.22}$$

is not a Nash equilibrium because now total sales fall short of capacity. At least one firm (say, firm 1) is selling less than its capacity. By cutting price slightly, firm 1 can increase its profits by selling up to its capacity, $\bar{q}_1$. Hence the Nash equilibrium of this second-stage game is for firms to choose the price at which quantity demanded exactly equals the total capacity built in the first stage:[5]

$$p_1 = p_2 = \bar{p}. \tag{15.23}$$

Anticipating that the price will be set such that firms sell all their capacity, the first-stage capacity choice game is essentially the same as the Cournot game. Therefore, the equilibrium quantities, price, and profits will be the same as in the Cournot game. Thus, even in markets (such as orange juice sold in grocery stores) where it looks like firms are setting prices, the Cournot model may prove more realistic than it first seems.

PRODUCT DIFFERENTIATION

Another way to avoid the Bertrand paradox is to replace the assumption that the firms' products are identical with the assumption that firms produce differentiated products. Many (if not most) real-world markets exhibit product differentiation. For example, toothpaste brands vary somewhat from supplier to supplier—differing in flavor, fluoride content, whitening agents, endorsement from the American Dental Association, and so forth. Even if suppliers' product attributes are similar, suppliers may still be differentiated in another dimension: physical location. Because demanders will be closer to some suppliers than to others, they may prefer nearby sellers because buying from them involves less travel time.

Meaning of "the market"

The possibility of product differentiation introduces some fuzziness into what we mean by the market for a good. With identical products, demanders were assumed to be indifferent about which firm's output they bought; hence they shop at the lowest-price firm, leading to the law of one price. The law of one price no longer holds if demanders strictly

[5]For completeness, it should be noted that there is no pure-strategy Nash equilibrium of the second-stage game with unequal prices ($p_1 \neq p_2$). The low-price firm would have an incentive to increase its price and/or the high-price firm would have an incentive to lower its price. For large capacities, there may be a complicated mixed-strategy Nash equilibrium, but this can be ruled out by supposing the cost of building capacity is sufficiently high.

prefer one supplier to another at equal prices. Are green-gel and white-paste toothpastes in the same market or in two different ones? Is a pizza parlor at the outskirts of town in the same market as one in the middle of town?

With differentiated products, we will take *the market* to be a group of closely related products that are more substitutable among each other (as measured by cross-price elasticities) than with goods outside the group. We will be somewhat loose with this definition, avoiding precise thresholds for how high the cross-price elasticity must be between goods within the group (and how low with outside goods). Arguments about which goods should be included in a product group often dominate antitrust proceedings, and we will try to avoid this contention here.

Bertrand competition with differentiated products

Return to the Bertrand model but now suppose there are n firms that simultaneously choose prices $p_i\,(i = 1, \ldots, n)$ for their differentiated products. Product i has its own specific attributes a_i, possibly reflecting special options, quality, brand advertising, or location. A product may be endowed with the attribute (orange juice is by definition made from oranges and cranberry juice from cranberries), or the attribute may be the result of the firm's choice and spending level (the orange juice supplier can spend more and make its juice from fresh oranges rather than from frozen concentrate). The various attributes serve to differentiate the products. Firm i's demand is

$$q_i(p_i, P_{-i}, a_i, A_{-i}), \tag{15.24}$$

where P_{-i} is a list of all other firms' prices besides i's, and A_{-i} is a list of all other firms' attributes besides i's. Firm i's total cost is

$$C_i(q_i, a_i) \tag{15.25}$$

and profit is thus

$$\pi_i = p_i q_i - C_i(q_i, a_i). \tag{15.26}$$

With differentiated products, the profit function (Equation 15.26) is differentiable, so we do not need to solve for the Nash equilibrium on a case-by-case basis as we did in the Bertrand model with identical products. We can solve for the Nash equilibrium as in the Cournot model, solving for best-response functions by taking each firm's first-order condition (here with respect to price rather than quantity). The first-order condition from Equation 15.26 with respect to p_i is

$$\frac{\partial \pi_i}{\partial p_i} = \underbrace{q_i + p_i \frac{\partial q_i}{\partial p_i}}_{A} - \underbrace{\frac{\partial C_i}{\partial q_i} \cdot \frac{\partial q_i}{\partial p_i}}_{B} = 0. \tag{15.27}$$

The first two terms (labeled A) on the right side of Equation 15.27 are a sort of marginal revenue—not the usual marginal revenue from an increase in quantity, but rather the marginal revenue from an increase in price. The increase in price increases revenue on existing sales of q_i units, but we must also consider the negative effect of the reduction in sales ($\partial q_i/\partial p_i$ multiplied by the price p_i) that would have been earned on these sales. The last term, labeled B, is the cost savings associated with the reduced sales that accompany an increased price.

The Nash equilibrium can be found by simultaneously solving the system of first-order conditions in Equation 15.27 for all $i = 1, \ldots, n$. If the attributes a_i are also choice variables (rather than just endowments), there will be another set of first-order conditions to consider. For firm i, the first-order condition with respect to a_i has the form

$$\frac{\partial \pi_i}{\partial a_i} = p_i \frac{\partial q_i}{\partial a_i} - \frac{\partial C_i}{\partial a_i} - \frac{\partial C_i}{\partial q_i} \cdot \frac{\partial q_i}{\partial a_i} = 0. \tag{15.28}$$

The simultaneous solution of these first-order conditions can be complex, and they yield few definitive conclusions about the nature of market equilibrium. Some insights from particular cases will be developed in the next two examples.

EXAMPLE 15.4 Toothpaste as a Differentiated Product

Suppose that two firms produce toothpaste, one a green gel and the other a white paste. To simplify the calculations, suppose that production is costless. Demand for product i is

$$q_i = a_i - p_i + \frac{p_j}{2}. \tag{15.29}$$

The positive coefficient on p_j, the other good's price, indicates that the goods are gross substitutes. Firm i's demand is increasing in the attribute a_i, which we will take to be demanders' inherent preference for the variety in question; we will suppose that this is an endowment rather than a choice variable for the firm (and so will abstract from the role of advertising to promote preferences for a variety).

Algebraic solution. Firm i's profit is

$$\pi_i = p_i q_i - C_i(q_i) = p_i\left(a_i - p_i + \frac{p_j}{2}\right), \tag{15.30}$$

where $C_i(q_i) = 0$ because i's production is costless. The first-order condition for profit maximization with respect to p_i is

$$\frac{\partial \pi_i}{\partial p_i} = a_i - 2p_i + \frac{p_j}{2} = 0. \tag{15.31}$$

Solving for p_i gives the following best-response functions for $i = 1, 2$:

$$p_1 = \frac{1}{2}\left(a_1 + \frac{p_2}{2}\right), \quad p_2 = \frac{1}{2}\left(a_2 + \frac{p_1}{2}\right). \tag{15.32}$$

Solving Equations 15.32 simultaneously gives the Nash equilibrium prices

$$p_i^* = \frac{8}{15}a_i + \frac{2}{15}a_j. \tag{15.33}$$

The associated profits are

$$\pi_i^* = \left(\frac{8}{15}a_i + \frac{2}{15}a_j\right)^2. \tag{15.34}$$

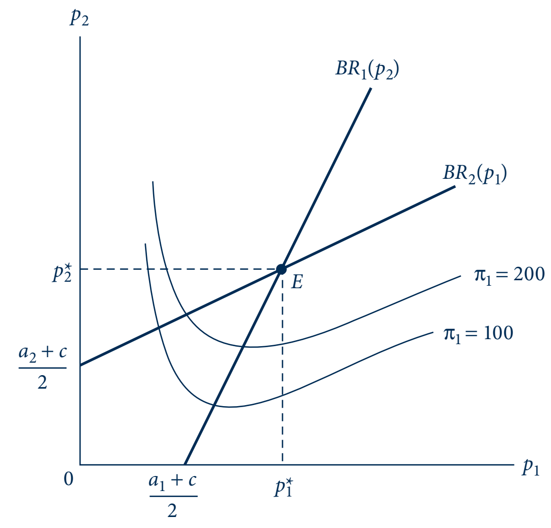
Firm i's equilibrium price is not only increasing in its own attribute, a_i, but also in the other product's attribute, a_j. An increase in a_j causes firm j to increase its price, which increases firm i's demand and thus the price i charges.

Graphical solution. We could also have solved for equilibrium prices graphically, as in Figure 15.4. The best responses in Equation 15.32 are upward sloping. They intersect at the Nash equilibrium, point E. The isoprofit curves for firm 1 are smile-shaped. To see this, take the expression for firm 1's profit in Equation 15.30, set it equal to a certain profit level (say, 100), and solve for p_2 to facilitate graphing it on the best-response diagram. We have

$$p_2 = \frac{100}{p_1} + p_1 - a_1. \tag{15.35}$$

FIGURE 15.4 Best Responses for Bertrand Model with Differentiated Products

Firm' best responses are drawn as thick lines; their intersection (E) is the Nash equilibrium. Isoprofit curves for firm 1 increase moving out along firm 1's best-response function.



The smile turns up as p_1 approaches 0 because the denominator of $100/p_1$ approaches 0. The smile turns up as p_1 grows large because then the second term on the right side of Equation 15.35 grows large. Isoprofit curves for firm 1 increase as one moves away from the origin along its best-response function.

QUERY: How would a change in the demand intercepts be represented on the diagram?

EXAMPLE 15.5 Hotelling's Beach

A simple model in which identical products are differentiated because of the location of their suppliers (spatial differentiation) was provided by H. Hotelling in the 1920s.[6] As shown in Figure 15.5, two ice cream stands, labeled A and B, are located along a beach of length L. The stands make identical ice cream cones, which for simplicity are assumed to be costless to produce. Let a and b represent the firms' locations on the beach. (We will take the locations of the ice cream stands as given; in a later example we will revisit firms' equilibrium location choices.) Assume that demanders are located uniformly along the beach, one at each unit of length. Carrying ice cream a distance d back to one's beach umbrella costs td^2 because ice cream melts more the higher the temperature t and the further one must walk.[7] Consistent with the Bertrand assumption, firms choose prices p_A and p_B simultaneously.

Determining demands. Let x be the location of the consumer who is indifferent between buying from the two ice cream stands. The following condition must be satisfied by x:

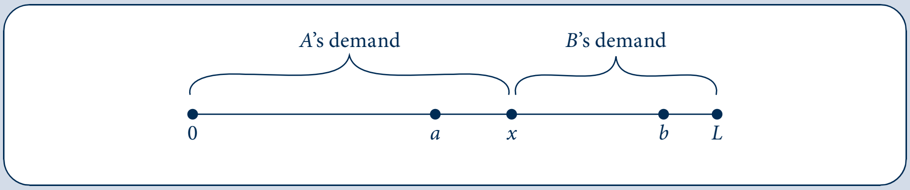
$$p_A + t(x - a)^2 = p_B + t(b - x)^2. \qquad \textbf{(15.36)}$$

[6]H. Hotelling, "Stability in Competition," *Economic Journal* 39 (1929): 41–57.

[7]The assumption of quadratic "transportation costs" turns out to simplify later work, when we compute firms' equilibrium locations in the model.

FIGURE 15.5 Hotelling's Beach

Ice cream stands A and B are located at points a and b along a beach of length L. The consumer who is indifferent between buying from the two stands is located at x. Consumers to the left of x buy from A and to the right buy from B.



The left side of Equation 15.36 is the generalized cost of buying from A (including the price paid and the cost of transporting the ice cream the distance $x - a$). Similarly, the right side is the generalized cost of buying from B. Solving Equation 15.36 for x yields

$$x = \frac{b+a}{2} + \frac{p_B - p_A}{2t(b-a)}. \tag{15.37}$$

If prices are equal, the indifferent consumer is located midway between a and b. If A's price is less than B's, then x shifts toward endpoint L. (This is the case shown in Figure 15.5.)

Because all demanders between 0 and x buy from A and because there is one consumer per unit distance, it follows that A's demand equals x:

$$q_A(p_A, p_B, a, b) = x = \frac{b+a}{2} + \frac{p_B - p_A}{2t(b-a)}. \tag{15.38}$$

The remaining $L - x$ consumers constitute B's demand:

$$q_B(p_B, p_A, b, a) = L - x = L - \frac{b+a}{2} + \frac{p_A - p_B}{2t(b-a)}. \tag{15.39}$$

Solving for Nash equilibrium. The Nash equilibrium is found in the same way as in Example 15.4 except that, for demands, we use Equations 15.38 and 15.39 in place of Equation 15.29. Skipping the details of the calculations, the Nash equilibrium prices are

$$\begin{aligned} p_A^* &= \frac{t}{3}(b-a)(2L+a+b), \\ p_B^* &= \frac{t}{3}(b-a)(4L-a-b). \end{aligned} \tag{15.40}$$

These prices will depend on the precise location of the two stands and will differ from each other. For example, if we assume that the beach is $L = 100$ yards long, $a = 40$ yards, $b = 70$ yards, and $t = \$0.001$ (one tenth of a penny), then $p_A^* = \$3.10$ and $p_B^* = \$2.90$. These price differences arise only from the locational aspects of this problem—the cones themselves are identical and costless to produce. Because A is somewhat more favorably located than B, it can charge a higher price for its cones without losing too much business to B. Using Equation 15.38 shows that

$$x = \frac{110}{2} + \frac{3.10 - 2.90}{(2)(0.001)(110)} \approx 52, \tag{15.41}$$

so stand A sells 52 cones, whereas B sells only 48 despite its lower price. At point x, the consumer is indifferent between walking the 12 yards to A and paying \$3.10 or walking 18 yards to B and paying \$2.90. The equilibrium is inefficient in that a consumer slightly to the right of x would incur a shorter walk by patronizing A but still chooses B because of A's power to set higher prices.

Equilibrium profits are

$$\pi_A^* = \frac{t}{18}(b-a)(2L+a+b)^2,$$
$$\pi_B^* = \frac{t}{18}(b-a)(4L-a-b)^2. \tag{15.42}$$

Somewhat surprisingly, the ice cream stands benefit from faster melting, as measured here by the transportation cost t. For example, if we take $L = 100$, $a = 40$, $b = 70$, and $t = \$0.001$ as in the previous paragraph, then $\pi_A^* = \$160$ and $\pi_B^* = \$140$ (rounding to the nearest dollar). If transportation costs doubled to $t = \$0.002$, then profits would double to $\pi_A^* = \$320$ and $\pi_B^* = \$280$.

The transportation/melting cost is the only source of differentiation in the model. If $t = 0$, then we can see from Equation 15.40 that prices equal 0 (which is marginal cost given that production is costless) and from Equation 15.42 that profits equal 0—in other words, the Bertrand paradox results.

QUERY: What happens to prices and profits if ice cream stands locate in the same spot? If they locate at the opposite ends of the beach?

Consumer search and price dispersion

Hotelling's model analyzed in Example 15.5 suggests the possibility that competitors may have some ability to charge prices above marginal cost and earn positive profits even if the physical characteristics of the goods they sell are identical. Firms' various locations—closer to some demanders and farther from others—may lead to spatial differentiation. The Internet makes the physical location of stores less relevant to consumers, especially if shipping charges are independent of distance (or are not assessed). Even in this setting, firms can avoid the Bertrand paradox if we drop the assumption that demanders know every firm's price in the market. Instead we will assume that demanders face a small cost s, called a *search cost*, to visit the store (or click to its website) to find its price.

Peter Diamond, winner of the Nobel Prize in economics in 2010, developed a model in which demanders search by picking one of the n stores at random and learning its price. Demanders know the equilibrium distribution of prices but not which store is charging which price. Demanders get their first price search for free but then must pay s for additional searches. They need at most one unit of the good, and they all have the same gross surplus v for the one unit.[8]

Not only do stores manage to avoid the Bertrand paradox in this model, they obtain the polar opposite outcome: All charge the monopoly price v, which extracts all consumer surplus! This outcome holds no matter how small the search cost s is—as long as s is positive (say, a penny). It is easy to see that all stores charging v is an equilibrium. If all charge the same price v, then demanders may as well buy from the first store they search because additional searches are costly and do not end up revealing a lower price. It can also be seen that this is the only equilibrium. Consider any outcome in which at least one store charges less than v, and consider the lowest-price store (label it i) in this outcome.

[8] P. Diamond, "A Model of Price Adjustment," *Journal of Economic Theory* 3 (1971): 156–68.

Store i could raise its price p_i by as much as s and still make all the sales it did before. The lowest price a demander could expect to pay elsewhere is no less than p_t, and the demander would have to pay the cost s to find this other price.

Less extreme equilibria are found in models where consumers have different search costs.[9] For example, suppose one group of consumers can search for free and another group has to pay s per search. In equilibrium, there will be some price dispersion across stores. One set of stores serves the low–search-cost demanders (and the lucky high–search-cost consumers who happen to stumble on a bargain). These bargain stores sell at marginal cost. The other stores serve the high–search-cost demanders at a price that makes these demanders indifferent between buying immediately and taking a chance that the next price search will uncover a bargain store.

TACIT COLLUSION

In Chapter 8, we showed that players may be able to earn higher payoffs in the subgame-perfect equilibrium of an infinitely repeated game than from simply repeating the Nash equilibrium from the single-period game indefinitely. For example, we saw that, if players are patient enough, they can cooperate on playing silent in the infinitely repeated version of the Prisoners' Dilemma rather than finking on each other each period. From the perspective of oligopoly theory, the issue is whether firms must endure the Bertrand paradox (marginal cost pricing and zero profits) in each period of a repeated game or whether they might instead achieve more profitable outcomes through tacit collusion.

A distinction should be drawn between tacit collusion and the formation of an explicit cartel. An explicit cartel involves legal agreements enforced with external sanctions if the agreements (e.g., to sustain high prices or low outputs) are violated. Tacit collusion can only be enforced through punishments internal to the market—that is, only those that can be generated within a subgame-perfect equilibrium of a repeated game. Antitrust laws generally forbid the formation of explicit cartels, so tacit collusion is usually the only way for firms to raise prices above the static level.

Finitely repeated game

Taking the Bertrand game to be the stage game, Selten's theorem from Chapter 8 tells us that repeating the stage game any finite number of times T does not change the outcome. The only subgame-perfect equilibrium of the finitely repeated Bertrand game is to repeat the stage-game Nash equilibrium—marginal cost pricing—in each of the T periods. The game unravels through backward induction. In any subgame starting in period T, the unique Nash equilibrium will be played regardless of what happened before. Because the outcome in period $T - 1$ does not affect the outcome in the next period, it is as though period $T - 1$ is the last period, and the unique Nash equilibrium must be played then, too. Applying backward induction, the game unravels in this manner all the way back to the first period.

Infinitely repeated game

If the stage game is repeated infinitely many periods, however, the folk theorem applies. The folk theorem indicates that any feasible and individually rational payoff can be sustained each period in an infinitely repeated game as long as the discount factor, δ, is close enough to unity. Recall that the discount factor is the value in the present period of one

[9]The following model is due to S. Salop and J. Stiglitz, "Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion," *Review of Economic Studies* 44 (1977): 493–510.

dollar earned one period in the future—a measure, roughly speaking, of how patient players are. Because the monopoly outcome (with profits divided among the firms) is a feasible and individually rational outcome, the folk theorem implies that the monopoly outcome must be sustainable in a subgame-perfect equilibrium for δ close enough to 1. Let's investigate the threshold value of δ needed.

First suppose there are two firms competing in a Bertrand game each period. Let Π_M denote the monopoly profit and P_M the monopoly price in the stage game. The firms may collude tacitly to sustain the monopoly price—with each firm earning an equal share of the monopoly profit—by using the grim trigger strategy of continuing to collude as long as no firm has undercut P_M in the past but reverting to the stage-game Nash equilibrium of marginal cost pricing every period from then on if any firm deviates by undercutting. Successful tacit collusion provides the profit stream

$$\begin{aligned} V^{\text{collude}} &= \frac{\Pi_M}{2} + \delta \cdot \frac{\Pi_M}{2} + \delta^2 \cdot \frac{\Pi_M}{2} + \cdots \\ &= \frac{\Pi_M}{2}(1 + \delta + \delta^2 + \cdots) \\ &= \left(\frac{\Pi_M}{2}\right)\left(\frac{1}{1-\delta}\right). \end{aligned} \tag{15.43}$$

Refer to Chapter 8 for a discussion of adding up a series of discount factors $1 + \delta + \delta^2 + \cdots$. We need to check that a firm has no incentive to deviate. By undercutting the collusive price P_M slightly, a firm can obtain essentially all the monopoly profit for itself in the current period. This deviation would trigger the grim strategy punishment of marginal cost pricing in the second and all future periods, so all firms would earn zero profit from there on. Hence the stream of profits from deviating is $V^{\text{deviate}} = \Pi_M$.

For this deviation not to be profitable we must have $V^{\text{collude}} \geq V^{\text{deviate}}$ or, on substituting,

$$\left(\frac{\Pi_M}{2}\right)\left(\frac{1}{1-\delta}\right) \geq \Pi_M. \tag{15.44}$$

Rearranging Equation 15.44, the condition reduces to $\delta \geq 1/2$. To prevent deviation, firms must value the future enough that the threat of losing profits by reverting to the one-period Nash equilibrium outweighs the benefit of undercutting and taking the whole monopoly profit in the present period.

EXAMPLE 15.6 Tacit Collusion in a Bertrand Model

Bertrand duopoly. Suppose only two firms produce a certain medical device used in surgery. The medical device is produced at constant average and marginal cost of \$10, and the demand for the device is given by

$$Q = 5{,}000 - 100P. \tag{15.45}$$

If the Bertrand game is played in a single period, then each firm will charge \$10 and a total of 4,000 devices will be sold. Because the monopoly price in this market is \$30, firms have a clear incentive to consider collusive strategies. At the monopoly price, total profits each period are \$40,000, and each firm's share of total profits is \$20,000. According to Equation 15.44, collusion at the monopoly price is sustainable if

$$20{,}000\left(\frac{1}{1-\delta}\right) \geq 40{,}000 \tag{15.46}$$

or if $\delta \geq 1/2$, as we saw.

Is the condition $\delta \geq 1/2$ likely to be met in this market? That depends on what factors we consider in computing δ, including the interest rate and possible uncertainty about whether the game will continue. Leave aside uncertainty for a moment and consider only the interest rate. If the period length is one year, then it might be reasonable to assume an annual interest rate of $r = 10\%$. As shown in the Appendix to Chapter 17, $\delta = 1/(1 + r)$; therefore, if $r = 10\%$, then $\delta = 0.91$. This value of δ clearly exceeds the threshold of 1/2 needed to sustain collusion. For δ to be less than the 1/2 threshold for collusion, we must incorporate uncertainty into the discount factor. There must be a significant chance that the market will not continue into the next period—perhaps because a new surgical procedure is developed that renders the medical device obsolete.

We focused on the best possible collusive outcome: the monopoly price of \$30. Would collusion be easier to sustain at a lower price, say \$20? No. At a price of \$20, total profits each period are \$30,000, and each firm's share is \$15,000. Substituting into Equation 15.44, collusion can be sustained if

$$15{,}000\left(\frac{1}{1-\delta}\right) \geq 30{,}000, \tag{15.47}$$

again implying $\delta \geq 1/2$. Whatever collusive profit the firms try to sustain will cancel out from both sides of Equation 15.44, leaving the condition $\delta \geq 1/2$. Therefore, we get a discrete jump in firms' ability to collude as they become more patient—that is, as δ increases from 0 to 1.[10] For δ below 1/2, no collusion is possible. For δ above 1/2, any price between marginal cost and the monopoly price can be sustained as a collusive outcome. In the face of this multiplicity of subgame-perfect equilibria, economists often focus on the one that is most profitable for the firms, but the formal theory as to why firms would play one or another of the equilibria is still unsettled.

Bertrand oligopoly. Now suppose n firms produce the medical device. The monopoly profit continues to be \$40,000, but each firm's share is now only \$40,000/n. By undercutting the monopoly price slightly, a firm can still obtain the whole monopoly profit for itself regardless of how many other firms there are. Replacing the collusive profit of \$20,000 in Equation 15.46 with \$40,000/n, we have that the n firms can successfully collude on the monopoly price if

$$\frac{40{,}000}{n}\left(\frac{1}{1-\delta}\right) \geq 40{,}000, \tag{15.48}$$

or

$$\delta \geq 1 - \frac{1}{n}. \tag{15.49}$$

Taking the "reasonable" discount factor of $\delta = 0.91$ used previously, collusion is possible when 11 or fewer firms are in the market and impossible with 12 or more. With 12 or more firms, the only subgame-perfect equilibrium involves marginal cost pricing and zero profits.

Equation 15.49 shows that tacit collusion is easier the more patient are firms (as we saw before) and the fewer of them there are. One rationale used by antitrust authorities to challenge certain mergers is that a merger may reduce n to a level such that Equation 15.49 begins to be satisfied and collusion becomes possible, resulting in higher prices and lower total welfare.

QUERY: A period can be interpreted as the length of time it takes for firms to recognize and respond to undercutting by a rival. What would be the relevant period for competing gasoline stations in a small town? In what industries would a year be a reasonable period?

[10]The discrete jump in firms' ability to collude is a feature of the Bertrand model; the ability to collude increases continuously with δ in the Cournot model of Example 15.7.

EXAMPLE 15.7 Tacit Collusion in a Cournot Model

Suppose that there are again two firms producing medical devices but that each period they now engage in quantity (Cournot) rather than price (Bertrand) competition. We will again investigate the conditions under which firms can collude on the monopoly outcome. To generate the monopoly outcome in a period, firms need to produce 1,000 each; this leads to a price of \$30, total profits of \$40,000, and firm profits of \$20,000. The present discounted value of the stream of these collusive profits is

$$V^{\text{collude}} = 20{,}000\left(\frac{1}{1-\delta}\right). \tag{15.50}$$

Computing the present discounted value of the stream of profits from deviating is somewhat complicated. The optimal deviation is not as simple as producing the whole monopoly output oneself and having the other firm produce nothing. The other firm's 1,000 units would be provided to the market. The optimal deviation (by firm 1, say) would be to best respond to firm 2's output of 1,000. To compute this best response, first note that if demand is given by Equation 15.45, then inverse demand is given by

$$P = 50 - \frac{Q}{100}. \tag{15.51}$$

Firm 1's profit is

$$\pi_1 = Pq_1 - cq_1 = q_1\left(40 - \frac{q_1 + q_2}{100}\right). \tag{15.52}$$

Taking the first-order condition with respect to q_1 and solving for q_1 yields the best-response function

$$q_1 = 2{,}000 - \frac{q_2}{2}. \tag{15.53}$$

Firm 1's optimal deviation when firm 2 produces 1,000 units is to increase its output from 1,000 to 1,500. Substituting these quantities into Equation 15.52 implies that firm 1 earns \$22,500 in the period in which it deviates.

How much firm 1 earns in the second and later periods following a deviation depends on the trigger strategies firms use to punish deviation. Assume that firms use the grim strategy of reverting to the Nash equilibrium of the stage game—in this case, the Nash equilibrium of the Cournot game—every period from then on. In the Nash equilibrium of the Cournot game, each firm best responds to the other in accordance with the best-response function in Equation 15.53 (switching subscripts in the case of firm 2). Solving these best-response equations simultaneously implies that the Nash equilibrium outputs are $q_1^* = q_2^* = 4{,}000/3$ and that profits are $\pi_1^* = \pi_2^* = \$17{,}778$. Firm 1's present discounted value of the stream of profits from deviation is

$$\begin{aligned} V^{\text{deviate}} &= 22{,}500 + 17{,}778 \cdot \delta + 17{,}778 \cdot \delta^2 + 17{,}778 \cdot \delta^3 + \cdots \\ &= 22{,}500 + (17{,}778 \cdot \delta)(1 + \delta + \delta^2 + \cdots) \\ &= \$22{,}500 + \$17{,}778\left(\frac{\delta}{1-\delta}\right). \end{aligned} \tag{15.54}$$

We have $V^{\text{collude}} \geq V^{\text{deviate}}$ if

$$\$20{,}000\left(\frac{1}{1-\delta}\right) \geq \$22{,}500 + \$17{,}778\left(\frac{\delta}{1-\delta}\right) \tag{15.55}$$

or, after some algebra, if $\delta \geq 0.53$.

Unlike with the Bertrand stage game, with the Cournot stage game there is a possibility of some collusion for discount factors below 0.53. However, the outcome would have to involve higher outputs and lower profits than monopoly.

QUERY: The benefit to deviating is lower with the Cournot stage game than with the Bertrand stage game because the Cournot firm cannot steal all the monopoly profit with a small deviation. Why then is a more stringent condition ($\delta \geq 0.53$ rather than $\delta \geq 0.5$) needed to collude on the monopoly outcome in the Cournot duopoly compared with the Bertrand duopoly?

LONGER-RUN DECISIONS: INVESTMENT, ENTRY, AND EXIT

The chapter has so far focused on the most basic short-run decisions regarding what price or quantity to set. The scope for strategic interaction expands when we introduce longer-run decisions. Take the case of the market for cars. Longer-run decisions include whether to update the basic design of the car, a process that might take up to two years to complete. Longer-run decisions may also include investing in robotics to lower production costs, moving manufacturing plants closer to consumers and cheap inputs, engaging in a new advertising campaign, and entering or exiting certain product lines (say, ceasing the production of station wagons or starting production of hybrid cars). In making such decisions, an oligopolist must consider how rivals will respond to it. Will competition with existing rivals become tougher or milder? Will the decision lead to the exit of current rivals or encourage new ones to enter? Is it better to be the first to make such a decision or to wait until after rivals move?

Flexibility versus commitment

Crucial to our analysis of longer-run decisions such as investment, entry, and exit is how easy it is to reverse a decision once it has been made. On first thought, it might seem that it is better for a firm to be able to easily reverse decisions because this would give the firm more flexibility in responding to changing circumstances. For example, a car manufacturer might be more willing to invest in developing a hybrid-electric car if it could easily change the design back to a standard gasoline-powered one should the price of gasoline (and the demand for hybrid cars along with it) decrease unexpectedly. Absent strategic considerations—and so for the case of a monopolist—a firm would always value flexibility and reversibility. The "option value" provided by flexibility is discussed in further detail in Chapter 7.

Surprisingly, the strategic considerations that arise in an oligopoly setting may lead a firm to prefer its decision be irreversible. What the firm loses in terms of flexibility may be offset by the value of being able to commit to the decision. We will see a number of instances of the value of commitment in the next several sections. If a firm can commit to an action before others move, the firm may gain a first-mover advantage. A firm may use its first-mover advantage to stake out a claim to a market by making a commitment to serve it and in the process limit the kinds of actions its rivals find profitable. Commitment is essential for a first-mover advantage. If the first mover could secretly reverse its decision, then its rival would anticipate the reversal and the firms would be back in the game with no first-mover advantage.

We already encountered a simple example of the value of commitment in the Battle of the Sexes game from Chapter 8. In the simultaneous version of the model, there were three Nash equilibria. In one pure-strategy equilibrium, the wife obtains her highest payoff by attending her favorite event with her husband, but she obtains lower payoffs in the other two equilibria (a pure-strategy equilibrium in which she attends her less favored event and a mixed-strategy equilibrium giving her the lowest payoff of all three). In the sequential version of the game, if a player were given the choice between being the first mover and having the ability to commit to attending an event or being the second mover and having the flexibility to be able to meet up with the first wherever he or she showed up, a player would always choose the ability to commit. The first mover can guarantee his or her preferred outcome as the unique subgame-perfect equilibrium by committing to attend his or her favorite event.

Sunk costs

Expenditures on irreversible investments are called *sunk costs.*

DEFINITION

Sunk cost. A *sunk cost* is an expenditure on an investment that cannot be reversed and has no resale value.

Sunk costs include expenditures on unique types of equipment (e.g., a newsprint-making machine) or job-specific training for workers (developing the skills to use the newsprint machine). There is sometimes confusion between sunk costs and what we have called *fixed costs.* They are similar in that they do not vary with the firm's output level in a production period and are incurred even if no output is produced in that period. But instead of being incurred periodically, as are many fixed costs (heat for the factory, salaries for secretaries and other administrators), sunk costs are incurred only once in connection with a single investment.[11] Some fixed costs may be avoided over a sufficiently long run—say, by reselling the plant and equipment involved—but sunk costs can never be recovered because the investments involved cannot be moved to a different use. When the firm makes a sunk investment, it has committed itself to that investment, and this may have important consequences for its strategic behavior.

First-mover advantage in the Stackelberg model

The simplest setting to illustrate the first-mover advantage is in the Stackelberg model, named after the economist who first analyzed it.[12] The model is similar to a duopoly version of the Cournot model except that—rather than simultaneously choosing the quantities of their identical outputs—firms move sequentially, with firm 1 (the leader) choosing its output first and then firm 2 (the follower) choosing after observing firm 1's output.

We use backward induction to solve for the subgame-perfect equilibrium of this sequential game. Begin with the follower's output choice. Firm 2 chooses the output q_2 that maximizes its own profit, taking firm 1's output as given. In other words, firm 2 best responds to firm 1's output. This results in the same best-response function for firm 2 as we computed in the Cournot game from the first-order condition (Equation 15.2). Label this best-response function $BR_2(q_1)$.

Turn then to the leader's output choice. Firm 1 recognizes that it can influence the follower's action because the follower best responds to 1's observed output. Substituting $BR_2(q_1)$ into the profit function for firm 1 given by Equation 15.1, we have

$$\pi_1 = P\left(q_1 + BR_2(q_1)\right)q_1 - C_1(q_1). \tag{15.56}$$

[11]Mathematically, the notion of sunk costs can be integrated into the per-period total cost function as

$$C_t(q_t) = S + F_t + cq_t,$$

where S is the per-period amortization of sunk costs (e.g., the interest paid for funds used to finance capital investments), F_t is the per-period fixed costs, c is marginal cost, and q_t is per-period output. If $q_t = 0$, then $C_t = S + F_t$; but if the production period is long enough, then some or all of F_t may also be avoidable. No portion of S is avoidable, however.

[12]H. von Stackelberg, *The Theory of the Market Economy*, trans. A. T. Peacock (New York: Oxford University Press, 1952).

The first-order condition with respect to q_1 is

$$\frac{\partial \pi_1}{\partial q_1} = P(Q) + P'(Q)q_1 + \underbrace{P'(Q)BR_2'(q_1)q_1}_{S} - C_i'(q_i) = 0. \tag{15.57}$$

This is the same first-order condition computed in the Cournot model (see Equation 15.2) except for the addition of the term S, which accounts for the strategic effect of firm 1's output on firm 2's. The strategic effect S will lead firm 1 to produce more than it would have in a Cournot model. By overproducing, firm 1 leads firm 2 to reduce q_2 by the amount $BR_2'(q_1)$; the fall in firm 2's output increases market price, thus increasing the revenue that firm 1 earns on its existing sales. We know that q_2 decreases with an increase in q_1 because best-response functions under quantity competition are generally downward sloping; see Figure 15.2 for an illustration.

The strategic effect would be absent if the leader's output choice were unobservable to the follower or if the leader could reverse its output choice in secret. The leader must be able to commit to an observable output choice or else firms are back in the Cournot game. It is easy to see that the leader prefers the Stackelberg game to the Cournot game. The leader could always reproduce the outcome from the Cournot game by choosing its Cournot output in the Stackelberg game. The leader can do even better by producing more than its Cournot output, thereby taking advantage of the strategic effect S.

EXAMPLE 15.8 Stackelberg Springs

Recall the two natural-spring owners from Example 15.1. Now, rather than having them choose outputs simultaneously as in the Cournot game, assume that they choose outputs sequentially as in the Stackelberg game, with firm 1 being the leader and firm 2 the follower.

Firm 2's output. We will solve for the subgame-perfect equilibrium using backward induction, starting with firm 2's output choice. We already found firm 2's best-response function in Equation 15.8, repeated here:

$$q_2 = \frac{a - q_1 - c}{2}. \tag{15.58}$$

Firm 1's output. Now fold the game back to solve for firm 1's output choice. Substituting firm 2's best response from Equation 15.58 into firm 1's profit function from Equation 15.56 yields

$$\pi_1 = \left[a - q_1 - \left(\frac{a - q_1 - c}{2}\right) - c\right]q_1 = \frac{1}{2}(a - q_1 - c)q_1. \tag{15.59}$$

Taking the first-order condition,

$$\frac{\partial \pi_1}{\partial q_1} = \frac{1}{2}(a - 2q_1 - c) = 0, \tag{15.60}$$

and solving gives $q_1^* = (a - c)/2$. Substituting q_1^* back into firm 2's best-response function gives $q_2^* = (a - c)/4$. Profits are $\pi_1^* = (1/8)(a - c)^2$ and $\pi_2^* = (1/16)(a - c)^2$.

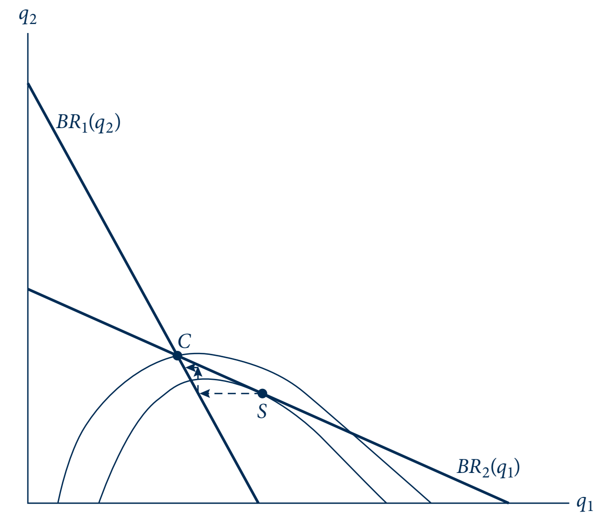
To provide a numerical example, suppose $a = 120$ and $c = 0$. Then $q_1^* = 60$, $q_2^* = 30$, $\pi_1^* = \$1{,}800$, and $\pi_2^* = \$900$. Firm 1 produces twice as much and earns twice as much as firm 2. Recall from the simultaneous Cournot game in Example 15.1 that, for these numerical values, total market output was 80 and total industry profit was 3,200, implying that each of the two firms produced $80/2 = 40$ units and earned $\$3{,}200/2 = \$1{,}600$. Therefore, when firm 1 is the first mover in a sequential game, it produces (60 − 40)/40 = 50% more and earns (1,800 − 1,600)/1,600 = 12.5% more than in the simultaneous game.

Graphing the Stackelberg outcome. Figure 15.6 illustrates the Stackelberg equilibrium on a best-response function diagram. The leader realizes that the follower will always best respond, so the resulting outcome will always be on the follower's best-response function. The leader effectively picks the point on the follower's best-response function that maximizes the leader's profit. The highest isoprofit (highest in terms of profit level, but recall from Figure 15.2 that higher profit levels are reached as one moves down toward the horizontal axis) is reached at the point *S* of tangency between firm 1's isoprofit and firm 2's best-response function. This is the Stackelberg equilibrium. Compared with the Cournot equilibrium at point *C*, the Stackelberg equilibrium involves higher output and profit for firm 1. Firm 1's profit is higher because, by committing to the high output level, firm 2 is forced to respond by reducing its output.

Commitment is required for the outcome to stray from firm 1's best-response function, as happens at point *S*. If firm 1 could secretly reduce q_1 (perhaps because q_1 is actual capacity that can be secretly reduced by reselling capital equipment for close to its purchase price to a manufacturer of another product that uses similar capital equipment), then it would move back to its best response, firm 2 would best respond to this lower quantity, and so on, following the dotted arrows from *S* back to *C*.

FIGURE 15.6 Stackelberg Game

Best-response functions from the Cournot game are drawn as thick lines. Frown-shaped curves are firm 1's isoprofits. Point *C* is the Nash equilibrium of the Cournot game (invoking simultaneous output choices). The Stackelberg equilibrium is point *S*, the point at which the highest isoprofit for firm 1 is reached on firm 2's best-response function. At *S*, firm 1's isoprofit is tangent to firm 2's best-response function. If firm 1 cannot commit to its output, then the outcome function unravels, following the dotted line from *S* back to *C*.



QUERY: What would be the outcome if the identity of the first mover were not given and instead firms had to compete to be the first? How would firms vie for this position? Do these considerations help explain overinvestment in Internet firms and telecommunications during the "dot-com bubble?"

Contrast with price leadership

In the Stackelberg game, the leader uses what has been called a "top dog" strategy,[13] aggressively overproducing to force the follower to scale back its production. The leader earns more than in the associated simultaneous game (Cournot), whereas the follower earns less. Although it is generally true that the leader prefers the sequential game to the simultaneous game (the leader can do at least as well, and generally better, by playing its Nash equilibrium strategy from the simultaneous game), it is not generally true that the leader harms the follower by behaving as a "top dog." Sometimes the leader benefits by behaving as a "puppy dog," as illustrated in Example 15.9.

EXAMPLE 15.9 Price-Leadership Game

Return to Example 15.4, in which two firms chose price for differentiated toothpaste brands simultaneously. So that the following calculations do not become too tedious, we make the simplifying assumptions that $a_1 = a_2 = 1$ and $c = 0$. Substituting these parameters back into Example 15.4 shows that equilibrium prices are $2/3 \approx 0.667$ and profits are $4/9 \approx 0.444$ for each firm.

Now consider the game in which firm 1 chooses price before firm 2.[14] We will solve for the subgame-perfect equilibrium using backward induction, starting with firm 2's move. Firm 2's best response to its rival's choice p_1 is the same as computed in Example 15.4—which, on substituting $a_2 = 1$ and $c = 0$ into Equation 15.32, is

$$p_2 = \frac{1}{2} + \frac{p_1}{4}. \tag{15.61}$$

Fold the game back to firm 1's move. Substituting firm 2's best response into firm 1's profit function from Equation 15.30 gives

$$\pi_1 = p_1\left[1 - p_1 + \frac{1}{2}\left(\frac{1}{2} + \frac{p_1}{4}\right)\right] = \frac{p_1}{8}(10 - 7p_1). \tag{15.62}$$

Taking the first-order condition and solving for the equilibrium price, we obtain $p_1^* \approx 0.714$. Substituting into Equation 15.61 gives $p_2^* \approx 0.679$. Equilibrium profits are $\pi_1^* \approx 0.446$ and $\pi_2^* \approx 0.460$. Both firms' prices and profits are higher in this sequential game than in the simultaneous one, but now the follower earns even more than the leader.

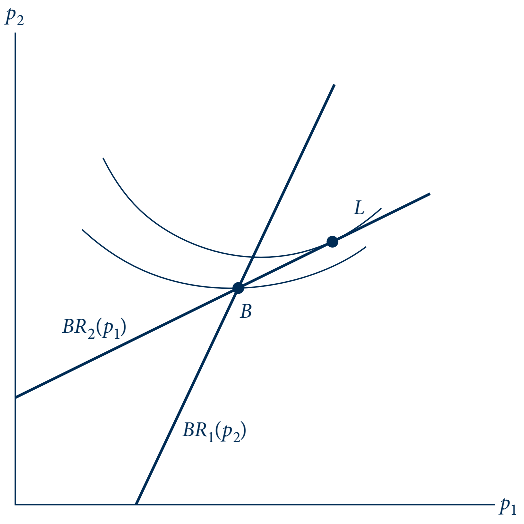
As illustrated in the best-response function diagram in Figure 15.7, firm 1 commits to a high price to induce firm 2 to raise its price also, essentially "softening" the competition between them.

[13]"Top dog," "puppy dog," and other colorful labels for strategies are due to D. Fudenberg and J. Tirole, "The Fat Cat Effect, the Puppy Dog Ploy, and the Lean and Hungry Look," *American Economic Review Papers and Proceedings* 74 (1984): 361–68.

[14]Sometimes this game is called the Stackelberg price game, although technically the original Stackelberg game involved quantity competition.

FIGURE 15.7 Price-Leadership Game

Thick lines are best-response functions from the game in which firms choose prices for differentiated products. U-shaped curves are firm 1's isoprofits. Point *B* is the Nash equilibrium of the simultaneous game, and *L* is the subgame-perfect equilibrium of the sequential game in which firm 1 moves first. At *L*, firm 1's isoprofit is tangent to firm 2's best response.



The leader needs a moderate price increase (from 0.667 to 0.714) to induce the follower to increase its price slightly (from 0.667 to 0.679), so the leader's profits do not increase as much as the follower's.

QUERY: What choice variable realistically is easier to commit to, prices or quantities? What business strategies do firms use to increase their commitment to their list prices?

We say that the first mover is playing a "puppy dog" strategy in Example 15.9 because it increases its price relative to the simultaneous-move game; when translated into outputs, this means that the first mover ends up producing less than in the simultaneous-move game. It is as though the first mover strikes a less aggressive posture in the market and so leads its rival to compete less aggressively.

A comparison of Figures 15.6 and 15.7 suggests the crucial difference between the games that leads the first mover to play a "top dog" strategy in the quantity game and a "puppy dog" strategy in the price game: The best-response functions have different slopes. The goal is to induce the follower to compete less aggressively. The slopes of the best-response functions determine whether the leader can best do that by playing aggressively itself or by softening its strategy. The first mover plays a "top dog" strategy in the sequential quantity game or indeed any game in which best responses slope down. When best responses slope down, playing more aggressively induces a rival to respond by competing less aggressively. Conversely, the first mover plays a "puppy dog" strategy in the price game or any game in which best responses slope up. When best responses slope up, playing less aggressively induces a rival to respond by competing less aggressively.

Therefore, knowing the slope of firms' best responses provides considerable insight into the sort of strategies firms will choose if they have commitment power. The Extensions at the end of this chapter provide further technical details, including shortcuts for determining the slope of a firm's best-response function just by looking at its profit function.

STRATEGIC ENTRY DETERRENCE

We saw that, by committing to an action, a first mover may be able to manipulate the second mover into being a less aggressive competitor. In this section we will see that the first mover may be able to prevent the entry of the second mover entirely, leaving the first mover as the sole firm in the market. In this case, the firm may not behave as an unconstrained monopolist because it may have distorted its actions to fend off the rival's entry.

In deciding whether to deter the second mover's entry, the first mover must weigh the costs and benefits relative to accommodating entry—that is, allowing entry to happen. Accommodating entry does not mean behaving nonstrategically. The first mover would move off its best-response function to manipulate the second mover into being less competitive, as described in the previous section. The cost of deterring entry is that the first mover would have to move off its best-response function even further than it would if it accommodates entry. The benefit is that it operates alone in the market and has market demand to itself. Deterring entry is relatively easy for the first mover if the second mover must pay a substantial sunk cost to enter the market.

EXAMPLE 15.10 Deterring Entry of a Natural Spring

Recall Example 15.8, where two natural-spring owners choose outputs sequentially. We now add an entry stage: In particular, after observing firm 1's initial quantity choice, firm 2 decides whether to enter the market. Entry requires the expenditure of sunk cost K_2, after which firm 2 can choose output. Market demand and cost are as in Example 15.8. To simplify the calculations, we will take the specific numerical values $a = 120$ and $c = 0$ [implying that inverse demand is $P(Q) = 120 - Q$, and that production is costless]. To further simplify, we will abstract from firm 1's entry decision and assume that it has already sunk any cost needed to enter before the start of the game. We will look for conditions under which firm 1 prefers to deter rather than accommodate firm 2's entry.

Accommodating entry. Start by computing firm 1's profit if it accommodates firm 2's entry, denoted π_1^{acc}. This has already been done in Example 15.8, in which there was no issue of deterring firm 2's entry. There we found firm 1's equilibrium output to be $(a - c)/2 = q_1^{acc}$ and its profit to be $(a - c)^2/8 = \pi_1^{acc}$. Substituting the specific numerical values $a = 120$ and $c = 0$, we have $q_1^{acc} = 60$ and $\pi_1^{acc} = (120 - 0)^2/8 = 1{,}800$.

Deterring entry. Next, compute firm 1's profit if it deters firm 2's entry, denoted π_1^{det}. To deter entry, firm 1 needs to produce an amount q_1^{det} high enough that, even if firm 2 best responds to q_1^{det}, it cannot earn enough profit to cover its sunk cost K_2. We know from Equation 15.58 that firm 2's best-response function is

$$q_2 = \frac{120 - q_1}{2}. \tag{15.63}$$

Substituting for q_2 in firm 2's profit function (Equation 15.7) and simplifying gives

$$\pi_2 = \left(\frac{120 - q_1^{det}}{2}\right)^2 - K_2. \tag{15.64}$$

Setting firm 2's profit in Equation 15.64 equal to 0 and solving yields

$$q_1^{det} = 120 - 2\sqrt{K_2}\,; \tag{15.65}$$

q_1^{det} is the firm 1 output needed to keep firm 2 out of the market. At this output level, firm 1's profit is

$$\pi_1^{det} = 2\sqrt{K_2}\left(120 - 2\sqrt{K_2}\right), \tag{15.66}$$

which we found by substituting q_1^{det}, $a = 120$, and $c = 0$ into firm 1's profit function from Equation 15.7. We also set $q_2 = 0$ because, if firm 1 is successful in deterring entry, it operates alone in the market.

Comparison. The final step is to juxtapose π_1^{acc} and π_1^{det} to find the condition under which firm 1 prefers deterring to accommodating entry. To simplify the algebra, let $x = 2\sqrt{K_2}$. Then $\pi_1^{det} = \pi_1^{acc}$ if

$$x^2 - 120x + 1{,}800 = 0. \tag{15.67}$$

Applying the quadratic formula yields

$$x = \frac{120 \pm \sqrt{7{,}200}}{2}. \tag{15.68}$$

Taking the smaller root (because we will be looking for a minimum threshold), we have $x = 17.6$ (rounding to the nearest decimal). Substituting $x = 17.6$ into $x = 2\sqrt{K_2}$ and solving for K_2 yields

$$K_2 = \left(\frac{x}{2}\right)^2 = \left(\frac{17.6}{2}\right)^2 \approx 77. \tag{15.69}$$

If $K_2 = 77$, then entry is so cheap for firm 2 that firm 1 would have to increase its output all the way to $q_1^{det} = 102$ in order to deter entry. This is a significant distortion above what it would produce when accommodating entry: $q_1^{acc} = 60$. If $K_2 < 77$, then the output distortion needed to deter entry wastes so much profit that firm 1 prefers to accommodate entry. If $K_2 > 77$, output need not be distorted as much to deter entry; thus, firm 1 prefers to deter entry.

QUERY: Suppose the first mover must pay the same entry cost as the second, $K_1 = K_2 = K$. Suppose further that K is high enough that the first mover prefers to deter rather than accommodate the second mover's entry. Would this sunk cost not be high enough to keep the first mover out of the market, too? Why or why not?

A real-world example of overproduction (or overcapacity) to deter entry is provided by the 1945 antitrust case against Alcoa, a U.S. aluminum manufacturer. A U.S. federal court ruled that Alcoa maintained much higher capacity than was needed to serve the market as a strategy to deter rivals' entry, and it held that Alcoa was in violation of antitrust laws.

To recap what we have learned in the last two sections: with quantity competition, the first mover plays a "top dog" strategy regardless of whether it deters or accommodates the second mover's entry. True, the entry-deterring strategy is more aggressive than the entry-accommodating one, but this difference is one of degree rather than kind. However, with price competition (as in Example 15.9), the first mover's entry-deterring strategy would differ in kind from its entry-accommodating strategy. It would play a "puppy dog" strategy if it wished to accommodate entry because this is how it manipulates the second mover into playing less aggressively. It plays a "top dog" strategy of lowering its price relative to the simultaneous game if it wants to deter entry. Two general principles emerge.

- Entry deterrence is always accomplished by a "top dog" strategy whether competition is in quantities or prices, or (more generally) whether best-response functions slope down or up. The first mover simply wants to create an inhospitable environment for the second mover.
- If firm 1 wants to accommodate entry, whether it should play a "puppy dog" or "top dog" strategy depends on the nature of competition—in particular, on the slope of the best-response functions.

SIGNALING

The preceding sections have shown that the first mover's ability to commit may afford it a big strategic advantage. In this section we will analyze another possible first-mover advantage: the ability to signal. If the second mover has incomplete information about market conditions (e.g., costs, demand), then it may try to learn about these conditions by observing how the first mover behaves. The first mover may try to distort its actions to manipulate what the second learns. The analysis in this section is closely tied to the material on signaling games in Chapter 8, and the reader may want to review that material before proceeding with this section.

The ability to signal may be a plausible benefit of being a first mover in some settings in which the benefit we studied earlier—commitment—is implausible. For example, in industries where the capital equipment is readily adapted to manufacture other products, costs are not very "sunk"; thus, capacity commitments may not be especially credible. The first mover can reduce its capacity with little loss. For another example, the price–leadership game involved a commitment to price. It is hard to see what sunk costs are involved in setting a price and thus what commitment value it has.[15] Yet even in the absence of commitment value, prices may have strategic, signaling value.

Entry-deterrence model

Consider the incomplete information game in Figure 15.8. The game involves a first mover (firm 1) and a second mover (firm 2) that choose prices for their differentiated products. Firm 1 has private information about its marginal cost, which can take on one of two values: high with probability $\Pr(H)$ or low with probability $\Pr(L) = 1 - \Pr(H)$. In period 1, firm 1 serves the market alone. At the end of the period, firm 2 observes firm 1's price and decides whether to enter the market. If it enters, it sinks an entry cost K_2 and learns the true level of firm 1's costs; then firms compete as duopolists in the second period, choosing prices for differentiated products as in Example 15.4 or 15.5. (We do not need to be specific about the exact form of demands.) If firm 2 does not enter, it obtains a payoff of zero, and firm 1 again operates alone in the market. Assume there is no discounting between periods.

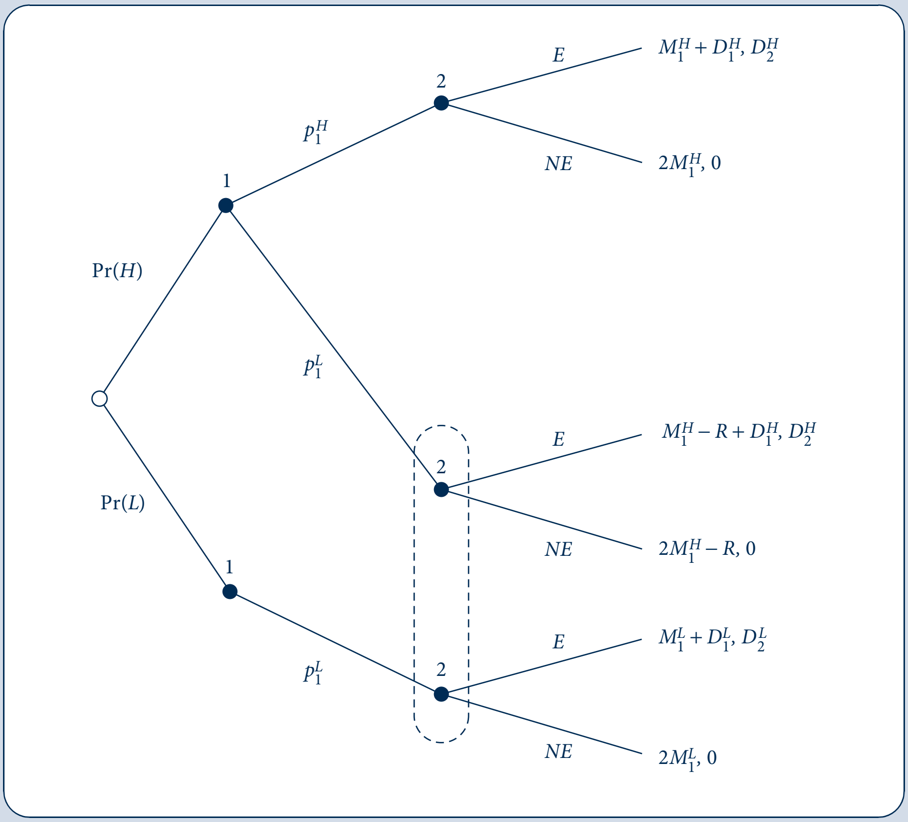
Firm 2 draws inferences about firm 1's cost from the price that firm 1 charges in the first period. Firm 2 earns more if it competes against the high-cost type because the

[15]The Query in Example 15.9 asks you to consider reasons why a firm may be able to commit to a price. The firm may gain commitment power by using contracts (e.g., long-term supply contracts with customers or a most-favored customer clause, which ensures that if the firm lowers price in the future to other customers, then the favored customer gets a rebate on the price difference). The firm may advertise a price through an expensive national advertising campaign. The firm may have established a valuable reputation as charging "everyday low prices."

FIGURE 15.8

Signaling for Entry Deterrence

Firm 1 signals its private information about its cost (high H or low L) through the price it sets in the first period. Firm 2 observes firm 1's price and then decides whether to enter. If firm 2 enters, the firms compete as duopolists; otherwise, firm 1 operates alone on the market again in the second period. Firm 2 earns positive profit if and only if it enters against a high cost rival.



high-cost type's price will be higher, and as we saw in Examples 15.4 and 15.5, the higher the rival's price for a differentiated product, the higher the firm's own demand and profit. Let D_i^t be the duopoly profit (not including entry costs) for firm $i \in \{1, 2\}$ if firm 1 is of type $t \in \{L, H\}$. To make the model interesting, we will suppose $D_2^L < K_2 < D_2^H$, so that firm 2 earns more than its entry cost if it faces the high-cost type but not if it faces the low-cost type. Otherwise, the information in firm 1's signal would be useless because firm 2 would always enter or always stay out regardless of firm 1's type.

To simplify the model, we will suppose that the low-cost type only has one relevant action in the first period—namely, setting its monopoly price p_1^L. The high-cost type can choose one of two prices: can set the monopoly price associated with its type, p_1^H, or it can choose the same price as the low type, p_1^L. Presumably, the optimal monopoly price is increasing in marginal cost; thus, $p_1^L < p_1^H$. Let M_1^t be firm 1's monopoly profit if it is of type $t \in \{L, H\}$ (the profit if it is alone and charges its optimal monopoly price p_1^H if it is the high type and p_1^L if it is the low type). Let R be the high type's loss relative to the optimal monopoly profit in the first period if it charges p_1^L rather than its optimal monopoly price p_1^H. Thus, if the high type charges p_1^H in the first period, then it earns M_1^H in that period, but if it charges p_1^L, it earns $M_1^H - R$.

Separating equilibrium

We will look for two kinds of perfect Bayesian equilibria: separating and pooling. In a separating equilibrium, the different types of the first mover must choose different actions. Here, there is only one such possibility for firm 1: The low-cost type chooses p_1^L and the high-cost type chooses p_1^H. Firm 2 learns firm 1's type from these actions perfectly and stays out on seeing p_1^l and enters on seeing p_1^H. It remains to check whether the high-cost type would prefer to deviate to p_1^L. In equilibrium, the high type earns a total profit of $M_1^H + D_1^H$: M_1^H in the first period because it charges its optimal monopoly price, and D_1^H in the second because firm 2 enters and the firms compete as duopolists. If the high type were to deviate to p_1^L, then it would earn $M_1^H - R$ in the first period, the loss R coming from charging a price other than its first-period optimum, but firm 2 would think it is the low type and would not enter. Hence firm 1 would earn M_1^H in the second period, for a total of $2M_1^H - R$ across periods. For deviation to be unprofitable we must have

$$M_1^H + D_1^H \geq 2M_1^H - R \tag{15.70}$$

or (after rearranging)

$$R \geq M_1^H - D_1^H. \tag{15.71}$$

That is, the high-type's loss from distorting its price from its monopoly optimum in the first period exceeds its gain from deterring firm 2's entry in the second period.

If the condition in Equation 15.71 does not hold, there still may be a separating equilibrium in an expanded game in which the low type can charge other prices besides p_1^L. The high type could distort its price downward below p_1^L, increasing the first-period loss the high type would suffer from pooling with the low type to such an extent that the high type would rather charge p_1^H even if this results in firm 2's entry.

Pooling equilibrium

If the condition in Equation 15.71 does not hold, then the high type would prefer to pool with the low type if pooling deters entry. Pooling deters entry if firm 2's prior belief that firm 1 is the high type, $\Pr(H)$—which is equal to its posterior belief in a pooling equilibrium—is low enough that firm 2's expected payoff from entering,

$$\Pr(H)D_2^H + [1 - \Pr(H)]D_2^L - K_2, \tag{15.72}$$

is less than its payoff of zero from staying out of the market.

Predatory pricing

The incomplete-information model of entry deterrence has been used to explain why a rational firm might want to engage in *predatory pricing*, the practice of charging an artificially low price to prevent potential rivals from entering or to force existing rivals to exit. The predatory firm sacrifices profits in the short run to gain a monopoly position in future periods.

Predatory pricing is prohibited by antitrust laws. In the most famous antitrust case, dating back to 1911, John D. Rockefeller—owner of the Standard Oil Company that controlled a substantial majority of refined oil in the United States—was accused of attempting to monopolize the oil market by cutting prices dramatically to drive rivals out and then raising prices after rivals had exited the market or sold out to Standard Oil. Predatory pricing remains a controversial antitrust issue because of the difficulty in distinguishing between predatory conduct, which authorities would like to prevent, and competitive conduct, which authorities would like to promote. In addition, economists initially had trouble developing game-theoretic models in which predatory pricing is rational and credible.

Suitably interpreted, predatory pricing may emerge as a rational strategy in the incomplete-information model of entry deterrence. Predatory pricing can show up in a separating equilibrium—in particular, in the expanded model where the low-cost type can separate only by reducing price below its monopoly optimum. Total welfare is actually higher in this separating equilibrium than it would be in its full-information counterpart. Firm 2's entry decision is the same in both outcomes, but the low-cost type's price may be lower (to signal its type) in the predatory outcome.

Predatory pricing can also show up in a pooling equilibrium. In this case it is the high-cost type that charges an artificially low price, pricing below its first-period optimum to keep firm 2 out of the market. Whether social welfare is lower in the pooling equilibrium than in a full-information setting is unclear. In the first period, price is lower (and total welfare presumably higher) in the pooling equilibrium than in a full-information setting. On the other hand, deterring firm 2's entry results in higher second-period prices and lower welfare. Weighing the first-period gain against the second-period loss would require detailed knowledge of demand curves, discount factors, and so forth.

The incomplete-information model of entry deterrence is not the only model of predatory pricing that economists have developed. Another model involves frictions in the market for financial capital that stem perhaps from informational problems (between borrowers and lenders) of the sort we will discuss in Chapter 18. With limits on borrowing, firms may only have limited resources to "make a go" in a market. A larger firm may force financially strapped rivals to endure losses until their resources are exhausted and they are forced to exit the market.

HOW MANY FIRMS ENTER?

To this point, we have taken the number of firms in the market as given, often assuming that there are at most two firms (as in Examples 15.1, 15.3, and 15.10). We did allow for a general number of firms, n, in some of our analysis (as in Examples 15.3 and 15.7) but were silent about how this number n was determined. In this section, we provide a game-theoretic analysis of the number of firms by introducing a first stage in which a large number of potential entrants can each choose whether to enter. We will abstract from first-mover advantages, entry deterrence, and other strategic considerations by assuming that firms make their entry decisions simultaneously. Strategic considerations are interesting and important, but we have already developed some insights into strategic considerations from the previous sections and—by abstracting from them—we can simplify the analysis here.

Barriers to entry

For the market to be oligopolistic with a finite number of firms rather than perfectly competitive with an infinite number of infinitesimal firms, some factors, called *barriers to entry*, must eventually make entry unattractive or impossible. We discussed many of these factors at length in the previous chapter on monopoly. If a sunk cost is required to enter the market, then—even if firms can freely choose whether to enter—only a limited number of firms will choose to enter in equilibrium because competition among more than that number would drive profits below the level needed to recoup the sunk entry cost. Government intervention in the form of patents or licensing requirements may prevent firms from entering even if it would be profitable for them to do so.

Some of the new concepts discussed in this chapter may introduce additional barriers to entry. Search costs may prevent consumers from finding new entrants with lower prices and/or higher quality than existing firms. Product differentiation may raise entry barriers because of strong brand loyalty. Existing firms may bolster brand loyalty through expensive advertising campaigns, and softening this brand loyalty may require entrants to conduct similarly expensive advertising campaigns. Existing firms may take other strategic measures to deter entry, such as committing to a high capacity or output level, engaging in predatory pricing, or other measures discussed in previous sections.

Long-run equilibrium

Consider the following game-theoretic model of entry in the long run. A large number of symmetric firms are potential entrants into a market. Firms make their entry decisions simultaneously. Entry requires the expenditure of sunk cost K. Let n be the number of firms that decide to enter. In the next stage, the n firms engage in some form of competition over a sequence of periods during which they earn the present discounted value of some constant profit stream. To simplify, we will usually collapse the sequence of periods of competition into a single period. Let $g(n)$ be the profit earned by an individual firm in this competition subgame [not including the sunk cost, so $g(n)$ is a gross profit]. Presumably, the more firms in the market, the more competitive the market is and the less an individual firm earns, so $g'(n) < 0$.

We will look for a subgame-perfect equilibrium in pure strategies.[16] This will be the number of firms, n^*, satisfying two conditions. First, the n^* entering firms earn enough to cover their entry cost: $g(n^*) \geq K$. Otherwise, at least one of them would have preferred to have deviated to not entering. Second, an additional firm cannot earn enough to cover its entry cost: $g(n^* + 1) \leq K$. Otherwise, a firm that remained out of the market could have profitably deviated by entering. Given that $g'(n) < 0$, we can put these two conditions together and say that n^* is the greatest integer satisfying $g(n^*) \geq K$.

This condition is reminiscent of the zero-profit condition for long-run equilibrium under perfect competition. The slight nuance here is that active firms are allowed to earn positive profits. Especially if K is large relative to the size of the market, there may only be a few long-run entrants (thus, the market looks like a canonical oligopoly) earning well above what they need to cover their sunk costs, yet an additional firm does not enter because its entry would depress individual profit enough that the entrant could not cover its large sunk cost.

Is the long-run equilibrium efficient? Does the oligopoly involve too few or too many firms relative to what a benevolent social planner would choose for the market? Suppose the social planner can choose the number of firms (restricting entry through licenses and promoting entry through subsidizing the entry cost) but cannot regulate prices or other competitive conduct of the firms once in the market. The social planner would choose n to maximize

$$CS(n) + ng(n) - nK, \tag{15.73}$$

where $CS(n)$ is equilibrium consumer surplus in an oligopoly with n firms, $ng(n)$ is total equilibrium profit (gross of sunk entry costs) across all firms, and nK is the total expenditure on sunk entry costs. Let n^{**} be the social planner's optimum.

In general, the long-run equilibrium number of firms, n^*, may be greater or less than the social optimum, n^{**}, depending on two offsetting effects: the *appropriability effect* and the *business-stealing effect*.

[16]A symmetric mixed-strategy equilibrium also exists in which sometimes more and sometimes fewer firms enter than can cover their sunk costs. There are multiple pure-strategy equilibria depending on the identity of the n^* entrants, but n^* is uniquely identified.

- The social planner takes account of the benefit of increased consumer surplus from lower prices, but firms do not appropriate consumer surplus and so do not take into account this benefit. This appropriability effect would lead a social planner to choose more entry than in the long-run equilibrium: $n^{**} > n^*$.
- Working in the opposite direction is that entry causes the profits of existing firms to decrease, as indicated by the derivative $g'(n) < 0$. Entry increases the competitiveness of the market, destroying some of firms' profits. In addition, the entrant "steals" some market share from existing firms—hence the term *business-stealing effect*. The marginal firm does not take other firms' loss in profits when making its entry decision, whereas the social planner would. The business-stealing effect biases long-run equilibrium toward more entry than a social planner would choose: $n^{**} < n^*$.

Depending on the functional forms for demand and costs, the appropriability effect dominates in some cases, and there is less entry in long-run equilibrium than is efficient. In other cases, the business-stealing dominates, and there is more entry in long-run equilibrium than is efficient, as in Example 15.11.

EXAMPLE 15.11 Cournot in the Long Run

Long-run equilibrium. Return to Example 15.3 of a Cournot oligopoly. We will determine the long-run equilibrium number of firms in the market. Let K be the sunk cost a firm must pay to enter the market in an initial entry stage. Suppose there is one period of Cournot competition after entry. To further simplify the calculations, assume that $a = 1$ and $c = 0$. Substituting these values back into Example 15.3, we have that an individual firm's gross profit is

$$g(n) = \left(\frac{1}{n+1}\right)^2. \tag{15.74}$$

The long-run equilibrium number of firms is the greatest integer n^* satisfying $g(n^*) \geq K$. Ignoring integer problems, n^* satisfies

$$n^* = \frac{1}{\sqrt{K}} - 1. \tag{15.75}$$

Social planner's problem. We first compute the individual terms in the social planner's objective function (Equation 15.73). Consumer surplus equals the area of the shaded triangle in Figure 15.9, which, using the formula for the area of a triangle, is

$$CS(n) = \frac{1}{2}Q(n)[a - P(n)] = \frac{n^2}{2(n+1)^2}; \tag{15.76}$$

here the last equality comes from substituting for price and quantity from Equations 15.18 and 15.19. Total profits for all firms (gross of sunk costs) equal the area of the shaded rectangle:

$$ng(n) = Q(n)P(n) = \frac{n}{(n+1)^2}. \tag{15.77}$$

Substituting from Equations 15.76 and 15.77 into the social planner's objective function (Equation 15.73) gives

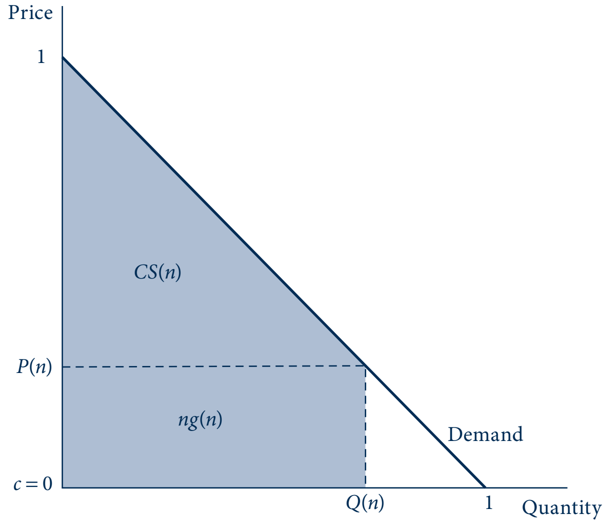
$$\frac{n^2}{2(n+1)^2} + \frac{n}{(n+1)^2} - nK. \tag{15.78}$$

After removing positive constants, the first-order condition with respect to n is

$$1 - K(n+1)^3 = 0, \tag{15.79}$$

FIGURE 15.9 Profit and Consumer Surplus in Example 15.11

Equilibrium for n firms drawn for the demand and cost assumptions in Example 15.11. Consumer surplus, $CS(n)$, is the area of the shaded triangle. Total profits $ng(n)$ for all firms (gross of sunk costs) is the area of the shaded rectangle.



implying that

$$n^{**} = \frac{1}{K^{1/3}} - 1. \qquad \textbf{(15.80)}$$

Ignoring integer problems, this is the optimal number of firms for a social planner.

Comparison. If $K < 1$ (a condition required for there to be any entry), then $n^{**} < n^*$, and so there is more entry in long-run equilibrium than a social planner would choose. To take a particular numerical example, let $K = 0.1$. Then $n^* = 2.16$ and $n^{**} = 1.15$, implying that the market would be a duopoly in long-run equilibrium, but a social planner would have preferred a monopoly.

QUERY: If the social planner could set both the number of firms and the price in this example, what choices would he or she make? How would these compare to long-run equilibrium?

Feedback effect

We found that certain factors decreased the stringency of competition and increased firms' profits (e.g., quantity rather than price competition, product differentiation, search costs, discount factors sufficient to sustain collusion). A feedback effect is that the more profitable the market is for a given number of firms, the more firms will enter the market, making the market more competitive and less profitable than it would be if the number of firms were fixed.

To take an extreme example, compare the Bertrand and Cournot games. Taking as given that the market involves two identical producers, we would say that the Bertrand game is much more competitive and less profitable than the Cournot game. This conclusion would be reversed if firms facing a sunk entry cost were allowed to make rational entry decisions. Only one firm would choose to enter the Bertrand market. A second firm would drive gross profit to zero, and so its entry cost would not be covered. The long-run equilibrium outcome would involve a monopolist and thus the highest prices and profits possible, exactly the opposite of our conclusions when the number of firms was fixed! On the other hand, the Cournot market may have space for several entrants driving prices and profits below their monopoly levels in the Bertrand market.

The moderating effect of entry should lead economists to be careful when drawing conclusions about oligopoly outcomes. Product differentiation, search costs, collusion, and other factors may reduce competition and increase profits in the short run, but they may also lead to increased entry and competition in the long run and thus have ambiguous effects overall on prices and profits. Perhaps the only truly robust conclusions about prices and profits in the long run involve sunk costs. Greater sunk costs constrain entry even in the long run, so we can confidently say that prices and profits will tend to be higher in industries requiring higher sunk costs (as a percentage of sales) to enter.[17]

INNOVATION

At the end of the previous chapter, we asked which market structure—monopoly or perfect competition—leads to more innovation in new products and cost-reducing processes. If monopoly is more innovative, will the long-run benefits of innovation offset the short-run deadweight loss of monopoly? The same questions can be asked in the context of oligopoly. Do concentrated market structures, with few firms perhaps charging high prices, provide better incentives for innovation? Which is more innovative, a large or a small firm? An established firm or an entrant? Answers to these questions can help inform policy toward mergers, entry regulation, and small-firm subsidies.

As we will see with the aid of some simple models, there is no definite answer as to what level of concentration is best for long-run total welfare. We will derive some general trade-offs, but quantifying these trade-offs to determine whether a particular market would be more innovative if it were concentrated or unconcentrated will depend on the nature of competition for innovation, the nature of competition for consumers, and the specification of demand and cost functions. The same can be said for determining what firm size or age is most innovative.

The models we introduce here are of *product* innovations, the invention of a product (e.g., plasma televisions) that did not exist before. Another class of innovations is that of *process* innovations, which reduce the cost of producing existing products—for example, the use of robot technology in automobile manufacture.

Monopoly on innovation

Begin by supposing that only a single firm, call it firm 1, has the capacity to innovate. For example, a pharmaceutical manufacturer may have an idea for a malaria vaccine that no other firm is aware of. How much would the firm be willing to complete research and development for the vaccine and to test it with large-scale clinical trials? How does this willingness to spend (which we will take as a measure of the innovativeness of the firm) depend on concentration of firms in the market?

[17]For more on robust conclusions regarding industry structure and competitiveness, see J. Sutton, *Sunk Costs and Market Structure* (Cambridge, MA: MIT Press, 1991).

Suppose first that there is currently no other vaccine available for malaria. If firm 1 successfully develops the vaccine, then it will be a monopolist. Letting Π_M be the monopoly profit, firm 1 would be willing to spend as much as Π_M to develop the vaccine. Next, to examine the case of a less concentrated market, suppose that another firm (firm 2) already has a vaccine on the market for which firm 1's would be a therapeutic substitute. If firm 1 also develops its vaccine, the firms compete as duopolists. Let π_D be the duopoly profit. In a Bertrand model with identical products, $\pi_D = 0$, but $\pi_D > 0$ in other models—for example, models involving quantity competition or collusion. Firm 1 would be willing to spend as much as π_D to develop the vaccine in this case. Comparing the two cases, because $\Pi_M > \pi_D$, it follows that firm 1 would be willing to spend more (and, by this measure, would be more innovative) in a more concentrated market. The general principle here can be labeled a *dissipation effect*: Competition dissipates some of the profit from innovation and thus reduces incentives to innovate. The dissipation effect is part of the rationale behind the patent system. A patent grants monopoly rights to an inventor, intentionally restricting competition to ensure higher profits and greater innovation incentives.

Another comparison that can be made is to see which firm, 1 or 2, has more of an incentive to innovate given that it has a monopoly on the initial idea. Firm 1 is initially out of the market and must develop the new vaccine to enter. Firm 2 is already in the malaria market with its first vaccine but can consider developing a second one as well, which we will continue to assume is a perfect substitute. As shown in the previous paragraph, firm 1 would be willing to pay up to π_D for the innovation. Firm 2 would not be willing to pay anything because it is currently a monopolist in the malaria vaccine market and would continue as a monopolist whether or not it developed the second medicine. (Crucial to this conclusion is that the firm with the initial idea can decline to develop it but still not worry that the other firm will take the idea; we will change this assumption in the next subsection.) Therefore, the potential competitor (firm 1) is more innovative by our measure than the existing monopolist (firm 2). The general principle here has been labeled a *replacement effect*: Firms gain less incremental profit and thus have less incentive to innovate if the new product replaces an existing product already making profit than if the firm is a new entrant in the market. The replacement effect can explain turnover in certain industries where old firms become increasingly conservative and are eventually displaced by innovative and quickly growing startups, as Microsoft displaced IBM as the dominant company in the computer industry and as Google now threatens to replace Microsoft.

Competition for innovation

New firms are not always more innovative than existing firms. The dissipation effect may counteract the replacement effect, leading old firms to be more innovative. To see this trade-off requires yet another variant of the model. Suppose now that more than one firm has an initial idea for a possible innovation and that they compete to see which can develop the idea into a viable product. For example, the idea for a new malaria vaccine may have occurred to scientists in two firms' laboratories at about the same time, and the firms may engage in a race to see who can produce a viable vaccine from this initial idea. Continue to assume that firm 2 already has a malaria vaccine on the market and that this new vaccine would be a perfect substitute for it.

The difference between the models in this and the previous section is that if firm 2 does not win the race to develop the idea, then the idea does not simply fall by the wayside but rather is developed by the competitor, firm 1. Firm 2 has an incentive to win the innovation competition to prevent firm 1 from becoming a competitor. Formally, if firm 1 wins the innovation competition, then it enters the market and is a competitor with firm 2, earning duopoly profit π_D. As we have repeatedly seen, this is the maximum that firm 1 would pay for the innovation. Firm 2's profit is Π_M if it wins the competition for the innovation but π_D if it loses and firm 1 wins. Firm 2 would pay up to the difference, $\Pi_M - \pi_D$, for the innovation. If $\Pi_M > 2\pi_D$—that is, if industry profit under a monopoly is greater than under a duopoly, which it is when some of the monopoly profit is dissipated by duopoly competition—then $\Pi_M - \pi_D > \pi_D$, and firm 2 will have more incentive to innovate than firm 1.

This model explains the puzzling phenomenon of dominant firms filing for "sleeping patents": patents that are never implemented. Dominant firms have a substantial incentive—as we have seen, possibly greater than entrants'—to file for patents to prevent entry and preserve their dominant position. Whereas the replacement effect may lead to turnover in the market and innovation by new firms, the dissipation effect may help preserve the position of dominant firms and retard the pace of innovation.

Summary

Many markets fall between the polar extremes of perfect competition and monopoly. In such imperfectly competitive markets, determining market price and quantity is complicated because equilibrium involves strategic interaction among the firms. In this chapter, we used the tools of game theory developed in Chapter 8 to study strategic interaction in oligopoly markets. We first analyzed oligopoly firms' short-run choices such as prices and quantities and then went on to analyze firms' longer-run decisions such as product location, innovation, entry, and the deterrence of entry. We found that seemingly small changes in modeling assumptions may lead to big changes in equilibrium outcomes. Therefore, predicting behavior in oligopoly markets may be difficult based on theory alone and may require knowledge of particular industries and careful empirical analysis. Still, some general principles did emerge from our theoretical analysis that aid in understanding oligopoly markets.

- One of the most basic oligopoly models, the Bertrand model involves two identical firms that set prices simultaneously. The equilibrium resulted in the Bertrand paradox: Even though the oligopoly is the most concentrated possible, firms behave as perfect competitors, pricing at marginal cost and earning zero profit.
- The Bertrand paradox is not the inevitable outcome in an oligopoly but can be escaped by changing assumptions underlying the Bertrand model—for example, allowing for quantity competition, differentiated products, search costs, capacity constraints, or repeated play leading to collusion.
- As in the Prisoners' Dilemma, firms could profit by coordinating on a less competitive outcome, but this outcome will be unstable unless firms can explicitly collude by forming a legal cartel or tacitly collude in a repeated game.
- For tacit collusion to sustain supercompetitive profits, firms must be patient enough that the loss from a price war in future periods to punish undercutting exceeds the benefit from undercutting in the current period.
- Whereas a nonstrategic monopolist prefers flexibility to respond to changing market conditions, a strategic oligopolist may prefer to commit to a single choice. The firm can commit to the choice if it involves a sunk cost that cannot be recovered if the choice is later reversed.
- A first mover can gain an advantage by committing to a different action from what it would choose in the Nash equilibrium of the simultaneous game. To deter entry, the first mover should commit to reducing the entrant's profits using an aggressive "top dog" strategy (high output or low price). If it does not deter entry, the first mover should commit to a strategy leading its rival to compete less aggressively. This is sometimes a "top dog" and sometimes a "puppy dog" strategy, depending on the slope of firms' best responses.
- Holding the number of firms in an oligopoly constant in the short run, the introduction of a factor that softens competition (e.g., product differentiation, search costs, collusion) will increase firms' profit, but an offsetting effect in the long run is that entry—which tends to reduce oligopoly profit—will be more attractive.
- Innovation may be even more important than low prices for total welfare in the long run. Determining which oligopoly structure is the most innovative is difficult because offsetting effects (dissipation and replacement) are involved.