12 Market Power

12.A Introduction

In the competitive model, all consumers and producers are assumed to act as price takers, in effect behaving as if the demand or supply functions that they face are infinitely elastic at going market prices. However, this assumption may not be a good one when there are only a few agents on one side of a market, for these agents will often possess *market power*—the ability to alter profitably prices away from competitive levels.

The simplest example of market power arises when there is only a single seller, a *monopolist*, of some good. If this good's market demand is a continuous decreasing function of price, then the monopolist, recognizing that a small increase in its price above the competitive level leads to only a small reduction in its sales, will find it worthwhile to raise its price above the competitive level.

Similar effects can occur when there is more than one agent, but still not many, on one side of a market. Most often, these agents with market power are firms, whose fewness arises from nonconvexities in production technologies (recall the discussion of entry in Section 10.F).

In this chapter, we study the functioning of markets in which market power is present. We begin, in Section 12.B, by considering the case in which there is a monopolist seller of some good. We review the theory of monopoly pricing and identify the welfare loss that it creates.

The remaining sections focus on situations of *oligopoly*, in which a number of firms compete in a market. In Sections 12.C and 12.D, we discuss several models of oligopolistic pricing. Each incorporates different assumptions about the underlying structure of the market and behavior of firms. The discussion highlights the implicatons of these differing assumptions for market outcomes. In Section 12.C, we focus on static models of oligopolistic pricing, where competition is viewed as a one-shot, simultaneous event. In contrast, in Section 12.D, we study how repeated interaction among firms may affect pricing in oligopolistic markets. This discussion constitutes an application of the theory of repeated games, a subject that we discuss in greater generality in Appendix A.

The analysis in Sections 12.B to 12.D treats the number of firms in the market as exogeneously given. In reality, however, the number of active firms in a market is likely to be affected by factors such as the size of market demand and the nature of competition within the market. Sections 12.E and 12.F consider issues that arise when the number of active firms in a market is determined endogenously.

Section 12.E specifies a simple model of entry into an oligopolistic market and studies the determinants of the number of active firms. It offers an analysis that parallels that considered in Section 10.F for competitive markets.

Section 12.F returns to a theme raised in Chapter 10. We illustrate how the competitive (price-taking) model can be viewed as a limiting case of oligopoly in which the size of the market, and hence the number of firms that can profitably operate in it, grows large. In the model we study, an active firm's market power diminishes as the market size expands; in the limit, the equilibrium market price comes to approximate the competitive level.

In Section 12.G, we briefly consider how firms in oligopolistic markets can make strategic precommitments to affect the conditions of future competition in a manner favorable to themselves. This issue nicely illustrates the importance of credible commitments in strategic settings, an issue we studied extensively in Chapter 9. In Appendix B, we consider in greater detail a particularly striking example of strategic precommitment to affect future market conditions, the case of entry deterrence through capacity choice.

If you have not done so already, you should review the game theory chapters in Part II before studying Sections 12.C to 12.G (in particular, review all of Chapter 7, Sections 8.A to 8.D, and Sections 9.A and 9.B).

An excellent source for further study of the topics covered in this chapter is Tirole (1988).[1]

12.B Monopoly Pricing

In this section, we study the pricing behavior of a profit-maximizing *monopolist*, a firm that is the only producer of a good. The demand for this good at price p is given by the function $x(p)$, which we take to be continuous and strictly decreasing at all p such that $x(p) > 0$.[2] For convenience, we also assume that there exists a price $\bar{p} < \infty$ such that $x(p) = 0$ for all $p \geq \bar{p}$.[3] Throughout, we suppose that the monopolist knows the demand function for its product and can produce output level q at a cost of $c(q)$.

The monopolist's decision problem consists of choosing its price p so as to maximize its profits (in terms of the numeraire), or formally, of solving

$$\underset{p}{\text{Max}} \quad px(p) - c(x(p)). \tag{12.B.1}$$

1. See also the survey by Shapiro (1989) for the topics covered in Sections 12.C, 12.D, and 12.G.

2. Throughout this chapter we take a partial equilibrium approach; see Chapter 10 for a discussion of this approach.

3. This assumption helps to insure that an optimal solution to the monopolist's problem exists. (See Exercise 12.B.2 for an example in which the failure of this condition leads to nonexistence.)

An equivalent formulation in terms of quantity choices can be derived by thinking instead of the monopolist as deciding on the level of output that it desires to sell, $q \geq 0$, letting the price at which it can sell this output be given by the *inverse demand function* $p(\cdot) = x^{-1}(\cdot)$.[4] Using this inverse demand function, the monopolist's problem can then be stated as

$$\underset{q \geq 0}{\text{Max}} \quad p(q)q - c(q). \tag{12.B.2}$$

We shall focus our analysis on this quantity formulation of the monopolist's problem [identical conclusions could equally well be developed from problem (12.B.1)]. We assume throughout that $p(\cdot)$ and $c(\cdot)$ are continuous and twice differentiable at all $q \geq 0$, that $p(0) > c'(0)$, and that there exists a unique output level $q^\circ \in (0, \infty)$ such that $p(q^\circ) = c'(q^\circ)$. Thus, q° is the unique socially optimal (competitive) output level in this market (see Chapter 10).

Under these assumptions, a solution to problem (12.B.2) can be shown to exist.[5] Given the differentiability assumed, the monopolist's optimal quantity, which we denote by q^m, must satisfy the first-order condition[6]

$$p'(q^m)\, q^m + p(q^m) \leq c'(q^m), \quad \text{with equality if } q^m > 0. \tag{12.B.3}$$

The left-hand side of (12.B.3) is the *marginal revenue* from a differential increase in q at q^m, which is equal to the derivative of revenue $d[p(q)q]/dq$, while the right-hand side is the corresponding marginal cost at q^m. Since $p(0) > c'(0)$, condition (12.B.3) can be satisfied only at $q^m > 0$. Hence, under our assumptions, *marginal revenue must equal marginal cost* at the monopolist's optimal output level:

$$p'(q^m)q^m + p(q^m) = c'(q^m). \tag{12.B.4}$$

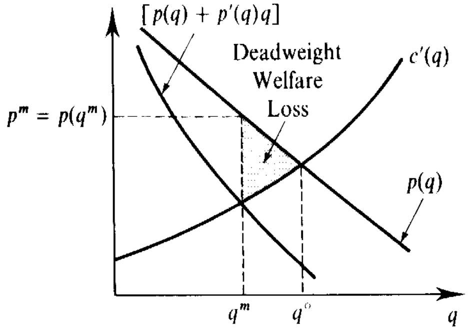
For the typical case in which $p'(q) < 0$ at all $q \geq 0$, condition (12.B.4) implies that we must have $p(q^m) > c'(q^m)$, and so *the price under monopoly exceeds marginal cost*. Correspondingly, the monopolist's optimal output q^m must be below the socially optimal (competitive) output level q°. The cause of this quantity distortion is the monopolist's recognition that a reduction in the quantity it sells allows it to increase the price charged on its remaining sales, an increase whose effect on profits is captured by the term $p'(q^m)q^m$ in condition (12.B.4).

The welfare loss from this quantity distortion, known as the *deadweight loss of monopoly*, can be measured using the change in Marshallian aggregate surplus

4. More precisely, to take account of the fact that $x(p) = 0$ for more than one value of p, we take $p(q) = \text{Min}\,\{p\colon x(p) = q\}$ at all $q \geq 0$. Thus, $p(0) = \bar{p}$, the lowest price at which $x(p) = 0$.

5. In particular, it follows from condition (12.B.3) and from the facts that $p'(q) \leq 0$ for all $q \geq 0$ and $p(q) < c'(q)$ for all $q > q^\circ$, that the monopolist's optimal choice must lie in the compact set $[0, q^\circ]$. Because the objective function in problem (12.B.2) is continuous, a solution must therefore exist (see Section M.F of the Mathematical Appendix).

6. Satisfaction of first-order condition (12.B.3) is sufficient for q^m to be an optimal choice if the objective function of problem (12.B.2) is concave on $[0, q^\circ]$. Note, however, that concavity of this objective function depends not only on the technology of the firm, as in the competitive model, but also on the shape of the inverse demand function. In particular, even with a convex cost function, the monopolist's profit function can violate this concavity condition if demand is a convex function of price.



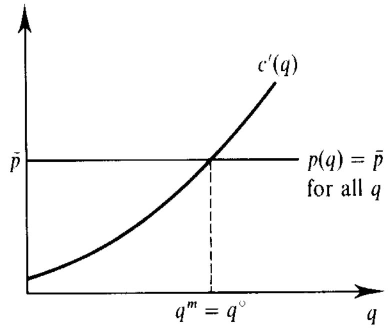


Figure 12.B.1 (left)
The monopoly solution and welfare loss when $p'(\cdot) < 0$.

Figure 12.B.2 (right)
The monopoly solution when $p'(q) = 0$ for all q.

(see Section 10.E),

$$\int_{q^m}^{q^\circ} [p(s) - c'(s)]\, ds > 0,$$

where q° is the socially optimal (competitive) output level.

Figure 12.B.1 illustrates the monopoly outcome in this case. The monopolist's quantity q^m is determined by the intersection of the graphs of marginal revenue $p'(q)q + p(q)$ and marginal cost $c'(q)$. The monopoly price $p(q^m)$ can then be determined from the inverse demand curve. The deadweight welfare loss is equal to the area of the shaded region.

Note from condition (12.B.4) that the monopoly quantity distortion is absent in the special case in which $p'(q) = 0$ for all q. In this case, where $p(q)$ equals some constant $\bar{p}$ at all $q > 0$, the monopolist sells the same quantity as a price-taking competitive firm because it perceives that any increase in its price above the competitive price $\bar{p}$ causes it to lose all its sales.[7] Figure 12.B.2 depicts this special case.

Example 12.B.1: *Monopoly Pricing with a Linear Inverse Demand Function and Constant Returns to Scale.* Suppose that the inverse demand function in a monopolized market is $p(q) = a - bq$ and that the monopolist's cost function is $c(q) = cq$, where $a > c \geq 0$ [so that $p(0) > c'(0)$] and $b > 0$. In this case, the objective function of the monopolist's problem (12.B.2) is concave, and so condition (12.B.4) is both necessary and sufficient for a solution to the monopolist's problem. From condition (12.B.4), we can calculate the monopolist's optimal quantity and price to be $q^m = (a - c)/2b$ and $p^m = (a + c)/2$. In contrast, the socially optimal (competitive) output level and price are $q^\circ = (a - c)/b$ and $p^\circ = p(q^\circ) = c$. ■

Although we do not discuss these issues here, we point out that the behavioral distortions arising under monopoly are not limited to pricing decisions. (Exercises 12.B.9 and 12.B.10 ask you to investigate two examples.)

The monopoly quantity distortion is fundamentally linked to the fact that if the monopolist wants to increase the quantity it sells, it must lower its price on *all* its existing sales. In fact,

7. This inverse demand function arises, for example, when each consumer i has quasilinear preferences of the form $u_i(q_i) + m_i$ with $u_i(q_i) = \bar{p}q_i$, where q_i is consumer i's consumption of the good under study and m_i is his consumption of the numeraire commodity. [Strictly speaking, with these preferences we now have a multivalued demand correspondence rather than a demand function, but $p(\cdot)$ is nevertheless a function as before.]

if the monopolist were able to *perfectly discriminate* among its customers in the sense that it could make a distinct offer to each consumer, knowing the consumer's preferences for its product, then the monopoly quantity distortion would disappear.

To see this formally, let each consumer i have a quasilinear utility function of the form $u_i(q_i) + m_i$ over the amount q_i of the monopolist's good that he consumes and the amount m_i that he consumes of the numeraire good, and normalize $u_i(0) = 0$. Suppose that the monopolist makes a take-it-or-leave-it offer to each consumer i of the form (q_i, T_i), where q_i is the quantity offered to consumer i and T_i is the total payment that the consumer must make in return. Given offer (q_i, T_i), consumer i will accept the monopolist's offer if and only if $u_i(q_i) - T_i \geq 0$. As a result, the monopolist can extract a payment of exactly $u_i(q_i)$ from consumer i in return for q_i units of its product, leaving the consumer with a surplus of exactly zero from consumption of the good. Given this fact, the monopolist will choose the quantities it sells to the I consumers $(q_1, \ldots, q_I)$ to solve

$$\underset{(q_1,\ldots,q_I)\geq 0}{\text{Max}} \quad \sum_{i=1}^{I} u_i(q_i) - c(\textstyle\sum_i q_i). \tag{12.B.5}$$

Note, however, that any solution to problem (12.B.5) maximizes the aggregate surplus in the market, and so the monopolist will sell each consumer exactly the socially optimal (competitive) quantity. Of course, the distributional properties of this outcome would not be terribly attractive in the absence of wealth redistribution: The monopolist would get all the aggregate surplus generated by its product, and each consumer i would receive a surplus of zero (i.e., each consumer i's welfare would be exactly equal to the level he would achieve if he consumed none of the monopolist's product). But in principle, these distributional problems can be corrected through lump-sum redistribution of the numeraire.

Thus, the welfare loss from monopoly pricing can be seen as arising from constraints that prevent the monopolist from charging fully discriminatory prices. In practice, however, these constraints can be significant. They may include the costs of assessing separate charges for different consumers, the monopolist's lack of information about consumer preferences, and the possibility of consumer resale. Exercise 12.B.5 explores some of these factors. It provides conditions under which the best the monopolist can do is to name a single per-unit price, as we assumed at the beginning of this section.

12.C Static Models of Oligopoly

We now turn to cases in which more than one, but still not many, firms compete in a market. These are known as situations of *oligopoly*. Competition among firms in an oligopolistic market is inherently a setting of strategic interaction. For this reason, the appropriate tool for its analysis is game theory. Because this discussion constitutes our first application of the theory of games, we focus on relatively simple *static* models of oligopoly, in which there is only one period of competitive interaction and firms take their actions simultaneously.

We begin by studying a model of simultaneous price choices by firms with constant returns to scale technologies, known as the *Bertrand model*. This model displays a striking feature: With just two firms in a market, we obtain a perfectly competitive outcome. Motivated by this finding, we then consider three alterations of this model that weaken its strong and often implausible conclusion: a change in the firm's strategy from choosing its price to choosing its quantity of output (the *Cournot model*); the introduction of capacity constraints (or, more generally, decreasing returns to scale); and the presence of product differentiation.[8]

One lesson of this analysis is that a critical part of game-theoretic modeling goes into choosing the strategies and payoff functions of the players. In the context of oligopolistic markets, this choice requires that considerable thought be given both to the demand and technological features of the market and to the underlying processes of competition.

Unless otherwise noted, we restrict our attention to pure strategy equilibria of the models we study.

The Bertrand Model of Price Competition

We begin by considering the model of oligopolistic competition proposed by Bertrand (1883). There are two profit-maximizing firms, firms 1 and 2 (a *duopoly*), in a market whose demand function is given by $x(p)$. As in Section 10.B, we assume that $x(\cdot)$ is continuous and strictly decreasing at all p such that $x(p) > 0$ and that there exists a $\bar{p} < \infty$ such that $x(p) = 0$ for all $p \geq \bar{p}$. The two firms have constant returns to scale technologies with the same cost, $c > 0$, per unit produced. We assume that $x(c) \in (0, \infty)$, which implies that the socially optimal (competitive) output level in this market is strictly positive and finite (see Chapter 10).

Competition takes place as follows: The two firms simultaneously name their prices p_1 and p_2. Sales for firm j are then given by

$$x_j(p_j, p_k) = \begin{cases} x(p_j) & \text{if } p_j < p_k \\ \frac{1}{2}x(p_j) & \text{if } p_j = p_k \\ 0 & \text{if } p_j > p_k. \end{cases}$$

The firms produce to order and so they incur production costs only for an output level equal to their actual sales. Given prices p_j and p_k, firm j's profits are therefore equal to $(p_j - c)x_j(p_j, p_k)$.

The Bertrand model constitutes a well-defined simultaneous-move game to which we can apply the concepts developed in Chapter 8. In fact, the Nash equilibrium outcome of this model, presented in Proposition 12.C.1, is relatively simple to discern.

Proposition 12.C.1: There is a unique Nash equilibrium (p_1^*, p_2^*) in the Bertrand duopoly model. In this equilibrium, both firms set their prices equal to cost: $p_1^* = p_2^* = c$.

Proof: To begin, note that both firms setting their prices equal to c is indeed a Nash equilibrium. At these prices, both firms earn zero profits. Neither firm can gain by raising its price because it will then make no sales (thereby still earning zero); and by lowering its price below c a firm increases its sales but incurs losses. What remains is to show that there can be no other Nash equilibrium.[9] Suppose, first, that the lower of the two prices named is less than c. In this case, the firm naming this price

8. Section 12.D studies a fourth variation that involves repeated interaction among firms.

9. Recall that we restrict attention to pure strategy equilibria here. See Exercise 12.C.2 for a consideration of mixed strategy equilibria. There you are asked to show that under the conditions assumed here, Proposition 12.C.1 continues to hold: $p_1^* = p_2^* = c$ is the unique Nash equilibrium, pure or mixed, of the Bertrand model.

incurs losses. But by raising its price above c, the worst it can do is earn zero. Thus, these price choices could not constitute a Nash equilibrium.

Now suppose that one firm's price is equal to c and that the other's price is strictly greater than c: $p_j = c$, $p_k > c$. In this case, firm j is selling to the entire market but making zero profits. By raising its price a little, say to $\hat{p}_j = c + (p_k - c)/2$, firm j would still make all the sales in the market, but at a strictly positive profit. Thus, these price choices also could not constitute an equilibrium.

Finally, suppose that both price choices are strictly greater than c: $p_j > c$, $p_k > c$. Without loss of generality, assume that $p_j \leq p_k$. In this case, firm k can be earning at most $\frac{1}{2}(p_j - c)x(p_j)$. But by setting its price equal to $p_j - \varepsilon$ for $\varepsilon > 0$, that is, by undercutting firm j's price, firm k will get the entire market and earn $(p_j - \varepsilon - c)x(p_j - \varepsilon)$. Since $(p_j - \varepsilon - c)x(p_j - \varepsilon) > \frac{1}{2}(p_j - c)x(p_j)$ for small-enough $\varepsilon > 0$, firm k can strictly increase its profits by doing so. Thus, these price choices are also not an equilibrium.

The three types of price configurations that we have just ruled out constitute all the possible price configurations other than $p_1 = p_2 = c$, and so we are done. ■

The striking implication of Proposition 12.C.1 is that with only two firms we get the perfectly competitive outcome. In effect, competition between the two firms makes each firm face an infinitely elastic demand curve at the price charged by its rival.

The basic idea of Proposition 12.C.1 can also be readily extended to any number of firms greater than two. [In this case, if firm j names the lowest price in the market, say $\tilde{p}$, along with $\tilde{J} - 1$ other firms, it earns $(1/\tilde{J})x(\tilde{p})$.] You are asked to show this in Exercise 12.C.1.

Exercise 12.C.1: Show that in any Nash equilibrium of the Bertrand model with $J > 2$ firms, all sales take place at a price equal to cost.

Thus, the Bertrand model predicts that the distortions arising from the exercise of market power are limited to the special case of monopoly. Notable as this result is, it also seems an unrealistic conclusion in many (although not all) settings. In the remainder of this section, we examine three changes in the Bertrand model that considerably weaken this strong conclusion: First, we make *quantity* the firms' strategic variable. Second, we introduce *capacity constraints* (or, more generally, decreasing returns to scale). Third, we allow for *product differentiation.*

Quantity Competition (The Cournot Model)

Suppose now that competition between the two firms takes a somewhat different form: The two firms simultaneously decide how much to produce, q_1 and q_2. Given these quantity choices, price adjusts to the level that clears the market, $p(q_1 + q_2)$, where $p(\cdot) = x^{-1}(\cdot)$ is the inverse demand function. This model is known as the *Cournot model*, after Cournot (1838). You can imagine farmers deciding how much of a perishable crop to pick each morning and send to a market. Once they have done so, the price at the market ends up being the level at which all the crops that have been sent are sold.[10] In this discussion, we assume that $p(\cdot)$ is differentiable

10. One scenario that will lead to this outcome arises when buyers bid for the crops sent that day (very much like sellers in the Bertrand model; see Exercise 12.C.5).

with $p'(q) < 0$ at all $q \geq 0$. As before, both firms produce output at a cost of $c > 0$ per unit. We also assume that $p(0) > c$ and that there exists a unique output level $q^\circ \in (0, \infty)$ such that $p(q^\circ) = c$ [in terms of the demand function $x(\cdot)$, $q^\circ = x(c)$]. Quantity q° is therefore the socially optimal (competitive) output level in this market.

To find a (pure strategy) Nash equilibrium of this model, consider firm j's maximization problem given an output level $\bar{q}_k$ of the other firm, $k \neq j$:

$$\operatorname*{Max}_{q_j \geq 0} \quad p(q_j + \bar{q}_k)q_j - cq_j. \tag{12.C.1}$$

In solving problem (12.C.1), firm j acts exactly like a monopolist who faces inverse demand function $\tilde{p}(q_j) = p(q_j + \bar{q}_k)$. An optimal quantity choice for firm j given its rival's output $\bar{q}_k$ must therefore satisfy the first-order condition

$$p'(q_j + \bar{q}_k)\, q_j + p(q_j + \bar{q}_k) \leq c, \quad \text{with equality if } q_j > 0. \tag{12.C.2}$$

For each $\bar{q}_k$, we let $b_j(\bar{q}_k)$ denote firm j's set of optimal quantity choices; $b_j(\cdot)$ is firm j's *best-response correspondence* (or *function* if it is single-valued).

A pair of quantity choices (q_1^*, q_2^*) is a Nash equilibrium if and only if $q_j^* \in b_j(q_k^*)$ for $k \neq j$ and $j = 1, 2$. Hence, if (q_1^*, q_2^*) is a Nash equilibrium, these quantities must satisfy[11]

$$p'(q_1^* + q_2^*)q_1^* + p(q_1^* + q_2^*) \leq c, \quad \text{with equality if } q_1^* > 0 \tag{12.C.3}$$

and

$$p'(q_1^* + q_2^*)q_2^* + p(q_1^* + q_2^*) \leq c, \quad \text{with equality if } q_2^* > 0. \tag{12.C.4}$$

It can be shown that under our assumptions we must have $(q_1^*, q_2^*) \gg 0$, and so conditions (12.C.3) and (12.C.4) must both hold with equality in any Nash equilibrium.[12] Adding these two equalities tells us that in any Nash equilibrium we must have

$$p'(q_1^* + q_2^*)\left(\frac{q_1^* + q_2^*}{2}\right) + p(q_1^* + q_2^*) = c. \tag{12.C.5}$$

Condition (12.C.5) allows us to reach the conclusion presented in Proposition 12.C.2.

Proposition 12.C.2: In any Nash equilibrium of the Cournot duopoly model with cost $c > 0$ per unit for the two firms and an inverse demand function $p(\cdot)$ satisfying $p'(q) < 0$ for all $q \geq 0$ and $p(0) > c$, the market price is greater than c (the competitive price) and smaller than the monopoly price.

11. Note that this method of analysis, which relies on the use of first-order conditions to calculate best responses, differs from the method used in the analysis of the Bertrand model. The reason is that in the Bertrand model each firm's objective function is discontinuous in its decision variable, so that differential optimization techniques cannot be used. Fortunately, the determination of the Nash equilibrium in the Bertrand model turned out, nevertheless, to be quite simple.

12. To see this, suppose that $q_1^* = 0$. Condition (12.C.3) then implies that $p(q_2^*) \leq c$. By condition (12.C.4) and the fact that $p'(\cdot) < 0$, this implies that were $q_2^* > 0$ we would have $p'(q_2^*)q_2^* + p(q_2^*) < c$, and so $q_2^* = 0$. But this means that $p(0) \leq c$, contradicting the assumption that $p(0) > c$. Hence, we must have $q_1^* > 0$. A similar argument shows that $q_2^* > 0$. Note, however, that this conclusion depends on our assumption of equal costs for the two firms. For example, a firm might set its output equal to zero if it is much less efficient than its rival. Exercise 12.C.9 considers some of the issues that arise when firms have differing costs.

Proof: That the equilibrium price is above c (the competitive price) follows immediately from condition (12.C.5) and the facts that $q_1^* + q_2^* > 0$ and $p'(q) < 0$ at all $q \geq 0$. We next argue that $(q_1^* + q_2^*) > q^m$, that is, that the equilibrium duopoly price $p(q_1^* + q_2^*)$ is strictly less than the monopoly price $p(q^m)$. The argument is in two parts.

First, we argue that $(q_1^* + q_2^*) \geq q^m$. To see this, suppose that $q^m > (q_1^* + q_2^*)$. By increasing its quantity to $\hat{q}_j = q^m - q_k^*$, firm j would (weakly) increase the joint profit of the two firms (the firms' joint profit then equals the monopoly profit level, its largest possible level). In addition, because aggregate quantity increases, price must fall, and so firm k is strictly worse off. This implies that firm j is strictly better off, and so firm j would have a profitable deviation if $q^m > (q_1^* + q_2^*)$. We conclude that we must have $(q_1^* + q_2^*) \geq q^m$.

Second, condition (12.C.5) implies that we cannot have $(q_1^* + q_2^*) = q^m$ because then

$$p'(q^m)\frac{q^m}{2} + p(q^m) = c,$$

in violation of the monopoly first-order condition (12.B.4). Thus, we must in fact have $(q_1^* + q_2^*) > q^m$. ■

Proposition 12.C.2 tells us that the presence of two firms is *not* sufficient to obtain a competitive outcome in the Cournot model, in contrast with the prediction of the Bertrand model. The reason is straightforward. In this model, a firm no longer sees itself as facing an infinitely elastic demand. Rather, if the firm reduces its quantity by a (differential) unit, it increases the market price by $-p'(q_1 + q_2)$. If the firms found themselves jointly producing the competitive quantity and consequently earning zero profits, either one could do strictly better by reducing its output slightly.

At the same time, competition does lower the price below the monopoly level, the price that would maximize the firms' joint profit. This occurs because when each firm determines the profitability of selling an additional unit it fails to consider the reduction in its rival's profit that is caused by the ensuing decrease in the market price [note that in firm j's first-order condition (12.C.2), only q_j multiplies the term $p'(\cdot)$, whereas in the first-order condition for joint profit maximization $(q_1 + q_2)$ does].

Example 12.C.1: *Cournot Duopoly with a Linear Inverse Demand Function and Constant Returns to Scale.* Consider a Cournot duopoly in which the firms have a cost per unit produced of c and the inverse demand function is $p(q) = a - bq$, with $a > c \geq 0$ and $b > 0$. Recall that the monopoly quantity and price are $q^m = (a - c)/2b$ and $p^m = (a + c)/2$ and that the socially optimal (competitive) output and price are $q^\circ = (a - c)/b$ and $p^\circ = p(q^\circ) = c$. Using the first-order condition (12.C.2), we find that firm j's best-response function in this Cournot model is given by $b_j(q_k) = \text{Max}\,\{0, (a - c - bq_k)/2b\}$.

Firm 1's best-response function $b_1(q_2)$ is depicted graphically in Figure 12.C.1. Since $b_1(0) = (a - c)/2b$, its graph hits the q_1 axis at the monopoly output level $(a - c)/2b$. This makes sense: Firm 1's best response to firm 2 producing no output is to produce exactly its monopoly output level. Similarly, since $b_1(q_2) = 0$ for all

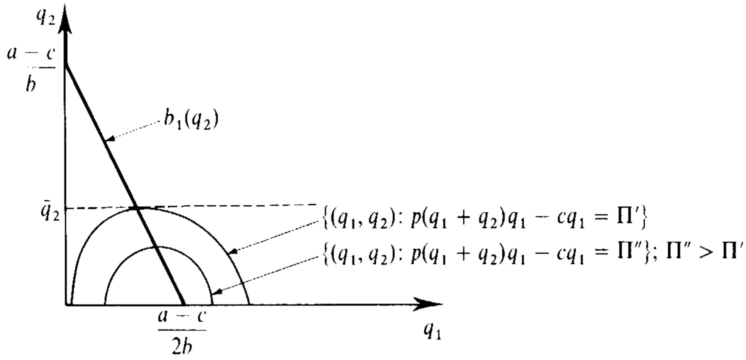


Figure 12.C.1
Firm 1's best-response function in the Cournot duopoly model of Example 12.C.1.

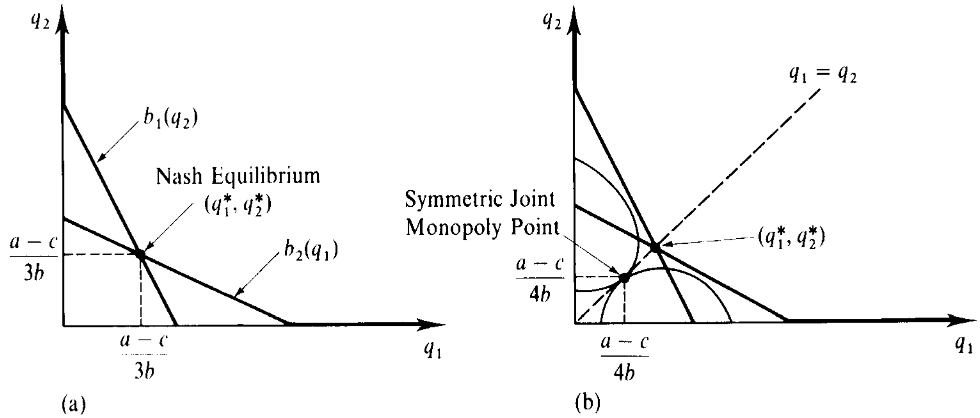


Figure 12.C.2
Nash equilibrium in the Cournot duopoly model of Example 12.C.1.

$q_2 \geq (a - c)/b$, the graph of firm 1's best-response function hits the q_2 axis at the socially optimal (competitive) output level $(a - c)/b$. Again, this makes sense: If firm 2 chooses an output level of at least $(a - c)/b$, any attempt by firm 1 to make sales results in a price below c. Two isoprofit loci of firm 1 are also drawn in the figure; these are sets of the form $\{(q_1, q_2)\colon p(q_1 + q_2)q_1 - cq_1 = \Pi\}$ for some profit level Π. The profit levels associated with these loci increase as we move toward firm 1's monopoly point $(q_1, q_2) = ((a - c)/2b, 0)$. Observe that firm 1's isoprofit loci have a zero slope where they cross the graph of firm 1's best-response function. This is because the best response $b_1(\bar{q}_2)$ identifies firm 1's maximal profit point on the line $q_2 = \bar{q}_2$ and must therefore correspond to a point of tangency between this line and an isoprofit locus. Firm 2's best-response function can be depicted similarly; given the symmetry of the firms, it is located symmetrically with respect to firm 1's best-response function in (q_1, q_2)-space [i.e., it hits the q_2 axis at $(a - c)/2b$ and hits the q_1 axis at $(a - c)/b$].

The Nash equilibrium, which in this example is unique, can be computed by finding the output pair (q_1^*, q_2^*) at which the graphs of the two best-response functions intersect, that is, at which $q_1^* = b_1(q_2^*)$ and $q_2^* = b_2(q_1^*)$. It is depicted in Figure 12.C.2(a) and corresponds to individual outputs of $q_1^* = q_2^* = \frac{1}{3}[(a - c)/b]$, total output of $\frac{2}{3}[(a - c)/b]$, and a market price of $p(q_1^* + q_2^*) = \frac{1}{3}(a + 2c) \in (c, p^m)$.

Also shown in Figure 12.C.2(b) is the symmetric joint monopoly point $(q^m/2, q^m/2) = ((a - c)/4b, (a - c)/4b)$. It can be seen that this point, at which each

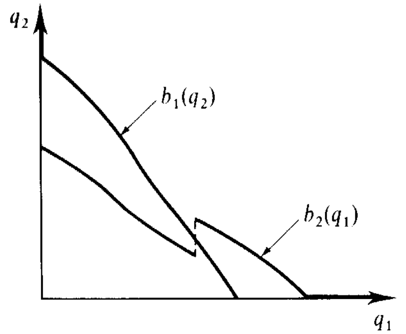


Figure 12.C.3
Nonexistence of (pure strategy) Nash equilibrium in the Cournot model.

firm produces half of the monopoly output of $(a - c)/2b$, is each firm's most profitable point on the $q_1 = q_2$ ray. ■

Exercise 12.C.6: Verify the computations and other claims in Example 12.C.1.

Up to this point we have not made any assumptions about the quasiconcavity in q_j of each firm j's objective function in problem (12.C.1). Without quasiconcavity of these functions, however, a pure strategy Nash equilibrium of this quantity game may not exist. For example, as happens in Figure 12.C.3, the best-response function of a firm lacking a quasiconcave objective function may "jump," leading to the possibility of nonexistence. (Strictly speaking, for a situation like the one depicted in Figure 12.C.3 to arise, the two firms must have different cost functions; see Exercise 12.C.8.) With quasiconcavity, we can use Proposition 8.D.3 to show that a pure strategy Nash equilibrium necessarily exists.

Suppose now that we have $J > 2$ identical firms facing the same cost and demand functions as above. Letting Q_J^* be aggregate output at equilibrium, an argument parallel to that above leads to the following generalization of condition (12.C.5):

$$p'(Q_J^*)\frac{Q_J^*}{J} + p(Q_J^*) = c. \tag{12.C.6}$$

At one extreme, when $J = 1$, condition (12.C.6) coincides with the monopoly first-order condition that we have seen in Section 12.B. At the other extreme, we must have $p(Q_J^*) \to c$ as $J \to \infty$. To see this, note that since Q_J^* is always less than the socially optimal (competitive) quantity q^o, it must be the case that $p'(Q_J^*)(Q_J^*/J) \to 0$ as $J \to \infty$. Hence, condition (12.C.6) implies that price must approach marginal cost as the number of firms grows infinitely large. This provides us with our first taste of a "competitive limit" result, a topic we shall return to in Section 12.F. Exercise 12.C.7 asks you to verify these claims for the model of Example 12.C.1.

Exercise 12.C.7: Derive the Nash equilibrium price and quantity levels in the Cournot model with J firms where each firm has a constant unit production cost of c and the inverse demand function in the market is $p(q) = a - bq$, with $a > c \geq 0$ and $b > 0$. Verify that when $J = 1$, we get the monopoly outcome; that output rises and price falls as J increases; and that as $J \to \infty$ the price and aggregate output in the market approach their competitive levels.

In contrast with the Bertrand model, the Cournot model displays a gradual reduction in market power as the number of firms increases. Yet, the "farmer sending crops to market" scenario may not seem relevant to a wide class of situations. After all, most firms seem to choose their prices, not their quantities. For this reason, many economists have thought that the Cournot model gives the right answer for the wrong reason. Fortunately, the departure from the Bertrand model that we study next offers an alternative interpretation of the Cournot model. The basic idea is that we can think of the quantity choices in the Cournot model as long-run choices of *capacity*, with the determination of price from the inverse demand function being a proxy for the outcome of short-run price competition given these capacity choices.

Capacity Constraints and Decreasing Returns to Scale

In many settings, it is natural to suppose that firms operate under conditions of eventual decreasing returns to scale, at least in the short run when capital is fixed. One special case of decreasing returns occurs when a firm has a capacity constraint that prevents it from producing more than some maximal amount, say $\bar{q}$. Here we consider, somewhat informally, how the introduction of capacity constraints affects the prediction of the Bertrand model.

With capacity constraints (or, for that matter, costs that exhibit decreasing returns to scale in a smoother way), it is no longer sensible to assume that a price announcement represents a commitment to provide *any* demanded quantity, since the costs of an order larger than capacity are infinite. We therefore make a minimal adjustment to the rules of the Bertrand model by taking price announcements to be a commitment to supply demand only up to capacity. We also assume that capacities are commonly known among the firms.

To see how capacity constraints can affect the outcome of the duopoly pricing game, suppose that each of the two firms has a constant marginal cost of $c > 0$ and a capacity constraint of $\bar{q} = \frac{3}{4}x(c)$. As before, the market demand function $x(\cdot)$ is continuous, is strictly decreasing at all p such that $x(p) > 0$, and has $x(c) > 0$.

In this case, the Bertrand outcome $p_1^* = p_2^* = c$ is no longer an equilibrium. To see this, note that because firm 2 cannot supply all demand at price $p_2^* = c$, firm 1 can anticipate making a strictly positive level of sales if it raises p_1 slightly above c. As a result, it has an incentive to deviate from $p_1^* = c$.

In fact, whenever the capacity level $\bar{q}$ satisfies $\bar{q} < x(c)$, each firm can *assure* itself of a strictly positive level of sales at a strictly positive profit margin by setting its price below $p(\bar{q})$ but above c. This is illustrated in Figure 12.C.4. In the figure, we assume that the lower-priced firm 2 fills the highest-valuation demands. By charging

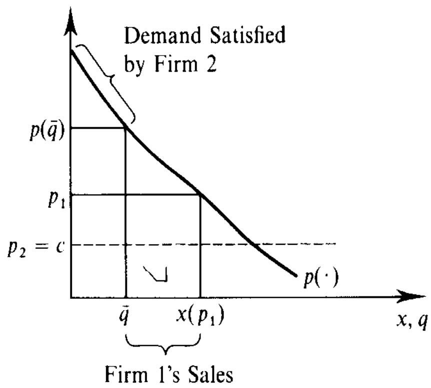


Figure 12.C.4
Calculation of demand in the presence of capacity constraints when the low-priced firm satisfies high-valuation demands first.

a price $p_1 \in (c, p(\bar{q}))$, firm 1 sells to the remaining demand at price p_1, making sales of $x(p_1) - \bar{q} > 0$. Hence, with capacity constraints, competition will not generally drive price down to cost, a point originally noted by Edgeworth (1897).

Determining the equilibrium outcome in situations in which capacity constraints are present can be tricky because knowledge of prices is no longer enough to determine each firm's sales. When the prices quoted are such that the low-priced firm cannot supply all demand at its quoted price, the demand for the higher-priced firm will generally depend on precisely *who* manages to buy from the low-priced firm. The high-priced firm will typically have greater sales if consumers with low valuations buy from the low-priced firm (in contrast with the assumption made in Figure 12.C.4) than if high-valuation consumers do. Thus, to determine demand functions for the firms, we now need to state a *rationing rule* specifying which consumers manage to buy from the low-priced firm when demand exceeds its capacity. In fact, the choice of a rationing rule can have important effects on equilibrium behavior. Exercise 12.C.11 asks you to explore some of the features of the equilibrium outcome when the highest valuation demands are served first, as in Figure 12.C.4. This is the rationing rule that tends to give the nicest results. Yet, it is neither more nor less plausible than other rules, such as a queue system or a random allocation of available units among possible buyers.

Up to this point in our discussion, we have taken a firm's capacity level as exogenous. Typically, however, we think of firms as *choosing* their capacity levels. This raises a natural question: What is the outcome in a model in which firms first choose their capacity levels and then compete in prices? Kreps and Scheinkman (1983) address this question and show that under certain conditions (among these is the assumption that high-valuation demands get served first when demand for a low-priced firm outstrips its capacity), the unique subgame perfect Nash equilibrium in this two-stage model is the *Cournot outcome*. This result is natural: the computation of price from the inverse demand curve in the Cournot model can be thought of as a proxy for this second-stage price competition. Indeed, for a wide range of capacity choices $(\bar{q}_1, \bar{q}_2)$, the unique equilibrium of the pricing subgame involves both firms setting their prices equal to $p(\bar{q}_1 + \bar{q}_2)$ (see Exercise 12.C.11). Thus, this two-stage model of capacity choice/price competition gives us the promised reinterpretation of the Cournot model: We can think of Cournot quantity competition as capturing long-run competition through capacity choice, with price competition occurring in the short run given the chosen levels of capacity.

Product Differentiation

In the Bertrand model, firms faced an infinitely elastic demand curve in equilibrium: With an arbitrarily small price differential, every consumer would prefer to buy from the lowest-priced firm. Often, however, consumers perceive differences among the products of different firms. When product differentiation exists, each firm will possess some market power as a result of the uniqueness of its product. Suppose, for example, that there are $J > 1$ firms. Each firm produces at a constant marginal cost of $c > 0$. The demand for firm j's product is given by the continuous function $x_j(p_j, p_{-j})$, where p_{-j} is a vector of prices of firm j's rivals.[13] In a setting of simultaneous price

13. Note the departure from the Bertrand model: In the Bertrand model, $x_j(p_j, p_{-j})$ is discontinuous at $p_j = \text{Min}_{k \neq j}\, p_k$.

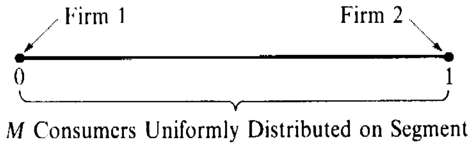


Figure 12.C.5
The linear city.

choices, each firm j takes its rivals' price choices $\bar{p}_{-j}$ as given and chooses p_j to solve

$$\underset{p_j}{\text{Max}} \quad (p_j - c)x_j(p_j, \bar{p}_{-j}).$$

Note that as long as $x_j(c, \bar{p}_{-j}) > 0$, firm j's best response necessarily involves a price in excess of its costs ($p_j > c$) because it can assure itself of strictly positive profits by setting its price slightly above c. Thus, in the presence of product differentiation, equilibrium prices will be above the competitive level. As with quantity competition and capacity constraints, the presence of product differentiation softens the strongly competitive result of the Bertrand model.

A number of models of product differentiation are popular in the applied literature. Example 12.C.2 describes one in some detail.

Example 12.C.2: *The Linear City Model of Product Differentiation.* Consider a city that can be represented as lying on a line segment of length 1, as shown in Figure 12.C.5. There is a continuum of consumers whose total number (or, more precisely, measure) is M and who are assumed to be located uniformly along this line segment. A consumer's location is indexed by $z \in [0, 1]$, the distance from the left end of the city. At each end of the city is located one supplier of widgets: Firm 1 is at the left end; firm 2, at the right. Widgets are produced at a constant unit cost of $c > 0$. Every consumer wants at most 1 widget and derives a gross benefit of v from its consumption. The total cost of buying from firm j for a consumer located a distance d from firm j is $p_j + td$, where $t/2 > 0$ can be thought of as the cost or disutility per unit of distance traveled by the consumer in going to and from firm j's location. The presence of travel costs introduces differentiation between the two firms' products because various consumers may now strictly prefer purchasing from one of the two firms even when the goods sell at the same price.

Figure 12.C.6(a) illustrates the purchase decisions of consumers located at various points in the city for a given pair of prices p_1 and p_2. Consumers at locations $[0, z_1)$

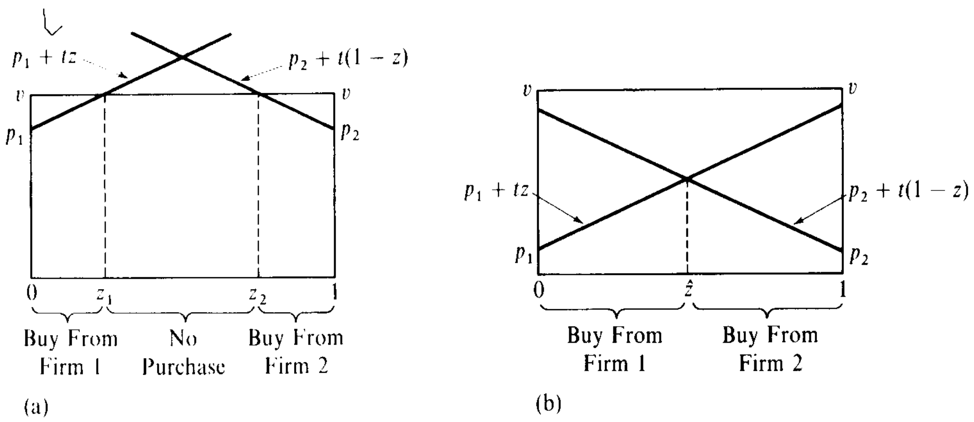


Figure 12.C.6
Consumer purchase decisions given p_1 and p_2. (a) Some consumers do not buy. (b) All consumers buy.

buy from firm 1. At these locations, $p_1 + tz < p_2 + t(1 - z)$ (purchasing from firm 1 is better than purchasing from firm 2), and $v - p_1 - tz > 0$ (purchasing from firm 1 is better than not purchasing at all). At location z_1, a consumer is indifferent between purchasing from firm 1 and not purchasing at all; that is, z_1 satisfies $v - p_1 - tz_1 = 0$. In Figure 12.C.6(a), consumers in the interval (z_1, z_2) do not purchase from either firm, while those in the interval $(z_2, 1]$ buy from firm 2.

Figure 12.C.6(b), by contrast, depicts a case in which, given prices p_1 and p_2, all consumers can obtain a strictly positive surplus by purchasing the good from one of the firms. The location of the consumer who is indifferent between the two firms is the point $\hat{z}$ such that

$$p_1 + t\hat{z} = p_2 + t(1 - \hat{z})$$

or

$$\hat{z} = \frac{t + p_2 - p_1}{2t}. \tag{12.C.7}$$

In general, the analysis of this model is complicated by the fact that depending on the parameters (v, c, t), the equilibria may involve market areas for the firms that do not touch [as in Figure 12.C.6(a)], or may have the firms battling for consumers in the middle of the market [as in Figure 12.C.6(b)]. To keep things as simple as possible here, we shall assume that consumers' value from a widget is large relative to production and travel costs, or more precisely, that $v > c + 3t$. In this case, it can be shown that a firm never wants to set its price at a level that causes some consumers not to purchase from either firm (see Exercise 12.C.13). In what follows, we shall therefore ignore the possibility of nonpurchase.

Given p_1 and p_2, let $\hat{z}$ be defined as in (12.C.7). Then firm 1's demand, given a pair of prices (p_1, p_2), equals $M\hat{z}$ when $\hat{z} \in [0, 1]$, M when $\hat{z} > 1$, and 0 when $\hat{z} < 0$.[14] Substituting for $\hat{z}$ from (12.C.7), we have

$$x_1(p_1, p_2) = \begin{cases} 0 & \text{if } p_1 > p_2 + t \\ (t + p_2 - p_1)M/2t & \text{if } p_1 \in [p_2 - t, p_2 + t] \\ M & \text{if } p_1 < p_2 - t. \end{cases} \tag{12.C.8}$$

By the symmetry of the two firms, the demand function of firm 2, $x_2(p_1, p_2)$, is

$$x_2(p_1, p_2) = \begin{cases} 0 & \text{if } p_2 > p_1 + t \\ (t + p_1 - p_2)M/2t & \text{if } p_2 \in [p_1 - t, p_1 + t] \\ M & \text{if } p_2 < p_1 - t. \end{cases} \tag{12.C.9}$$

Note from (12.C.8) and (12.C.9) that each firm j, in searching for its best response to any price choice $\bar{p}_{-j}$ by its rival, can restrict itself to prices in the interval $[\bar{p}_{-j} - t, \bar{p}_{-j} + t]$. Any price $p_j > \bar{p}_{-j} + t$ yields the same profits as setting $p_j = \bar{p}_{-j} + t$ (namely, zero), and any price $p_j < \bar{p}_{-j} - t$ yields lower profits than setting $p_j = \bar{p}_{-j} - t$ (all such prices result in sales of M units). Thus, firm j's best

14. Recall that the M consumers are uniformly distributed on the line segment, so $\hat{z}$ is the fraction who buy from firm 1.

response to $\bar{p}_{-j}$ solves

$$\underset{p_j}{\text{Max}} \quad (p_j - c)(t + \bar{p}_{-j} - p_j)\frac{M}{2t} \qquad (12.C.10)$$
$$\text{s.t. } p_j \in [\bar{p}_{-j} - t, \bar{p}_{-j} + t].$$

The necessary and sufficient (Kuhn–Tucker) first-order condition for this problem is

$$t + \bar{p}_{-j} + c - 2p_j \begin{cases} \leq 0 & \text{if } p_j = \bar{p}_{-j} - t \\ = 0 & \text{if } p_j \in (\bar{p}_{-j} - t, \bar{p}_{-j} + t) \\ \geq 0 & \text{if } p_j = \bar{p}_{-j} + t. \end{cases} \qquad (12.C.11)$$

Solving (12.C.11), we find that each firm j's best-response function is

$$b(\bar{p}_{-j}) = \begin{cases} \bar{p}_{-j} + t & \text{if } \bar{p}_{-j} \leq c - t \\ (t + \bar{p}_{-j} + c)/2 & \text{if } \bar{p}_{-j} \in (c - t, c + 3t) \\ \bar{p}_{-j} - t & \text{if } \bar{p}_{-j} \geq c + 3t. \end{cases} \qquad (12.C.12)$$

When $\bar{p}_{-j} < c - t$, firm j prices in a manner that leads its sales to equal zero (it cannot make profits because it cannot make sales at any price above c). When $\bar{p}_{-j} > c + 3t$, firm j prices in a manner that captures the entire market. In the intermediate case, firm j's best response to $\bar{p}_{-j}$ leaves both firms with strictly positive sales levels.

Given the symmetry of the model, we look for a symmetric equilibrium, that is, an equilibrium in which $p_1^* = p_2^* = p^*$. In any symmetric equilibrium, $p^* = b(p^*)$. Examining (12.C.12), we see that this condition can be satisfied only in the middle case (note also that this is the only case in which both firms can have strictly positive sales, as they must in any symmetric equilibrium). Thus, p^* must satisfy

$$p^* = \tfrac{1}{2}(t + p^* + c),$$

and so

$$p^* = c + t.$$

In this Nash equilibrium, each firm has sales of $M/2$ and a profit of $tM/2$. Note that as t approaches zero, the firms' products become completely undifferentiated and the equilibrium prices approach c, as in the Bertrand model. In the other direction, as the travel cost t becomes greater, thereby increasing the differentiation between the firms' products, equilibrium prices and profits increase.

Figure 12.C.7 depicts the best-response functions for the two firms (for prices greater than or equal to c) and the Nash equilibrium. As usual, the Nash equilibrium lies at the intersection of the graphs of these best-response functions. Note that there are no asymmetric equilibria here. ■

Matters become more complicated when $v < c + 3t$ because firms may wish to set prices at which some consumers do not want to purchase from either firm. One can show, however, that the equilibrium just derived remains valid as long as $v \geq c + \frac{3}{2}t$. In contrast, when $v < c + t$, in equilibrium the firms' market areas do not touch (the firms are like "local monopolists"). In the intermediate case where $v \in [c + t, c + \frac{3}{2}t]$, firms are at a "kink" in their demand functions and the consumer at the indifferent location $\hat{z}$ receives no surplus from his purchase in the equilibrium. Exercise 12.C.14 asks you to investigate these cases.

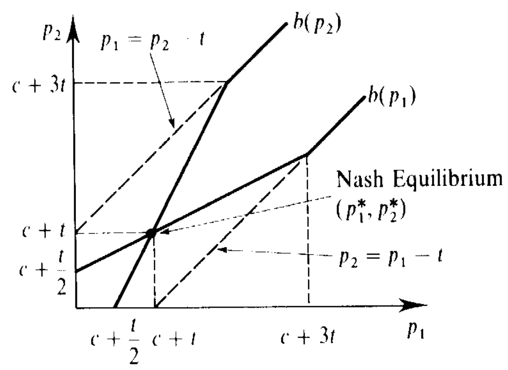


Figure 12.C.7 (left)
Best-response functions and Nash equilibrium in the linear city model when $v > c + 3t$.

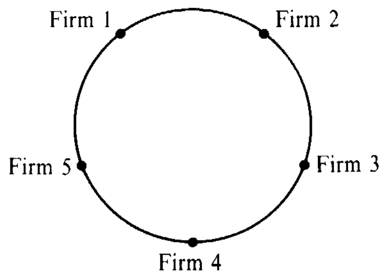


Figure 12.C.8 (right)
The circular city model when $J = 5$.

The essential features of the linear city model can be extended to the case in which $J > 2$. In doing so, it is often most convenient for analytical purposes to consider instead a model of a *circular city*, so that firms can be kept in symmetric positions.[15] In this model, which is due to Salop (1979), consumers are uniformly distributed along a circle of circumference 1, and the firms are positioned at equal intervals from one another. Figure 12.C.8 depicts a case where $J = 5$.

Models like the linear and circular city models are known as *spatial* models of product differentiation because each firm is identified with an "address" in product space. More generally, we can imagine firms' products located in some N-dimensional characteristics space, with consumers' "addresses" (their ideal points of consumption) distributed over this space.

Spatial models share the characteristic that each firm competes for customers only locally, that is, solely with the firms offering similar products. A commonly used alternative to spatial formulations, in which each product competes instead for sales with *all* other products, is the *representative consumer model* introduced by Spence (1976) and Dixit and Stiglitz (1977). In this model, a representative consumer is postulated whose preferences over consumption of the J products $(x_1, \ldots, x_J)$ and a numeraire good m take the quasilinear form

$$u(m, x_1, \ldots, x_J) = G\left(\sum_{j=1}^{J} f(x_j)\right) + m,$$

where both $G(\cdot)$ and $f(\cdot)$ are concave.[16] Normalizing the price of the numeraire to be 1, the first-order conditions for the representative consumer's maximization problem are

$$G'\left(\sum_{j=1}^{J} f(x_j)\right) f'(x_j) = p_j \qquad \text{for } j = 1, \ldots, J. \qquad (12.C.13)$$

These first-order conditions can be inverted to yield demand functions $x_j(p_1, \ldots, p_J)$ for $j = 1, \ldots, J$, which can then be used to specify a game of simultaneous price choices.[17]

An important variant of this representative consumer model arises in the limiting case where we have many products, each of which constitutes a small fraction of the sales in the overall market. In the limit, we can write the representative consumer's utility function as $G(\int f(x_j)\, dj) + m$, where x_j is now viewed as a function of the continuous index variable j.

15. In the segment [0, 1], only with two firms can we have symmetric positioning. With more than two firms, the two firms closest to the endpoints of the segment would have only one nearest neighbor but the firms in the interior would have two.

16. Dixit and Stiglitz (1977) actually consider more general utility functions of the form $u(G(\sum_j f(x_j)), m)$.

17. It is also common in the literature to study games of simultaneous quantity choices, using the expression in (12.C.13) directly as the inverse demand functions for the firms.

This leads to a considerable simplification because each firm j, in deciding on its price choice, can take the value of $\bar{x} = \int f(x_j)\,dj$, called the *index of aggregate output*, as given; its own production has no effect on the value of this index. Given the value of $\bar{x}$, firm j faces the demand function

$$x_j(p_j, \bar{x}) = \psi\left(\frac{p_j}{G'(\bar{x})}\right),$$

where $\psi(\cdot) = f'^{-1}(\cdot)$. Its optimal choice can then be viewed as a function $p_j^*(\bar{x})$ of the index $\bar{x}$. Thus, the equilibrium value of the aggregate output index, say $\bar{x}^*$, satisfies $\bar{x}^* = \int f(x_j(p_j^*(\bar{x}^*), \bar{x}^*))\,dj$.

This limiting case is known as the *monopolistic competition model*. It originates in Chamberlin (1933); see Hart (1985) for a modern treatment. In markets characterized by monopolistic competition, market power is accompanied by a low level of strategic interaction, in that the strategies of any particular firm do not affect the payoff of any other firm.[18]

12.D Repeated Interaction

One unrealistic assumption in the models presented in Section 12.C was their static, one-shot nature. In these models, a firm never had to consider the reaction of its competitors to its price or quantity choice. In the Bertrand model, for instance, a firm could undercut its rival's price by a penny and steal all the rival's customers. In practice, however, a firm in this circumstance may well worry that if it does undercut its rival in this manner, the rival will respond by cutting its own price, ultimately leading to only a short-run gain in sales but a long-run reduction in the price level in the market.

In this section, we consider the simplest type of dynamic model in which these concerns arise. Two identical firms complete for sales repeatedly, with competition in each period t described by the Bertrand model. When they do so, the two firms know all the prices that have been chosen (by *both* firms) previously. There is a discount factor $\delta < 1$, and each firm j attempts to maximize the discounted value of profits, $\sum_{t=1}^{\infty} \delta^{t-1} \pi_{jt}$, where π_{jt} is firm j's profit in period t. The game that this situation gives rise to is a dynamic game (see Chapter 9) of a special kind: it is obtained by repeated play of the same static simultaneous-move game and is known as a *repeated game*.

In this repeated Bertrand game, firm j's strategy specifies what price p_{jt} it will charge in each period t as a function of the history of all past price choices by the two firms, $H_{t-1} = \{p_{1\tau}, p_{2\tau}\}_{\tau=1}^{t-1}$. Strategies of this form allow for a range of interesting behaviors. For example, a firm's strategy could call for retaliation if the firm's rival ever lowers its price below some "threshold price." This retaliation could be brief, calling for the firm to lower its price for only a few periods after the rival "crosses the line," or it could be unrelenting. The retaliation could be tailored to the amount by which the firm's rival undercut it, or it could be severe no matter how minor the rival's transgression. The firm could also respond with increasingly cooperative

18. In contrast, in spatial models, even in the limit of a continuum of firms, strategic interaction remains. In that case, firms interact locally, and neighbors count, no matter how large the economy is.

behavior in return for its rival acting cooperatively in the past. And, of course, the firm's strategy could also make the firm's behavior in any period t independent of past history (a strategy involving no retaliation or rewards).

Of particular interest to us is the possibility that these types of behavioral responses could allow firms, in settings of repeated interaction, to sustain behavior more cooperative than the outcome predicted by the simple one-shot Bertrand model. We explore this possibility in the remainder of this section.

We begin by considering the case in which the firms compete only a finite number of times T (this is known as a *finitely repeated game*). Can the rich set of possible behaviors just described actually arise in a subgame perfect Nash equilibrium of this model? Recalling Proposition 9.B.4, we see the answer is "no." The unique subgame perfect Nash equilibrium of the finitely repeated Bertrand game simply involves T repetitions of the static Bertrand equilibrium in which prices equal cost. This is a simple consequence of backward induction: In the last period, T, we must be at the Bertrand solution, and therefore profits are zero in that period *regardless of what has happened earlier*. But, then, in period $T - 1$ we are, strategically speaking, at the last period, and the Bertrand solution must arise again. And so on, until we get to the first period. In summary, backward induction rules out the possibility of more cooperative behavior in the finitely repeated Bertrand game.

Things can change dramatically, however, when the horizon is extended to an infinite number of periods (this is known as an *infinitely repeated game*). To see this, consider the following strategies for firms $j = 1, 2$:

$$p_{jt}(H_{t-1}) = \begin{cases} p^m & \text{if all elements of } H_{t-1} \text{ equal } (p^m, p^m) \text{ or } t = 1 \\ c & \text{otherwise.} \end{cases} \qquad (12.D.1)$$

In words, firm j's strategy calls for it to initially play the monopoly price p^m in period 1. Then, in each period $t > 1$, firm j plays p^m if in every previous period both firms have charged price p^m and otherwise charges a price equal to cost. This type of strategy is called a *Nash reversion strategy*: Firms cooperate until someone deviates, and any deviation triggers a permanent retaliation in which both firms thereafter set their prices equal to cost, the one-period Nash strategy. Note that if both firms follow the strategies in (12.D.1), then both firms will end up charging the monopoly price in every period. They start by charging p^m, and therefore no deviation from p^m will ever be triggered.

For the strategies in (12.D.1), we have the result presented in Proposition 12.D.1.

Proposition 12.D.1: The strategies described in (12.D.1) constitute a subgame perfect Nash equilibrium (SPNE) of the infinitely repeated Bertrand duopoly game if and only if $\delta \geq \frac{1}{2}$.

Proof: Recall that a set of strategies is an SPNE of an infinite horizon game if and only if it specifies Nash equilibrium play in every subgame (see Section 9.B). To start, note that although each subgame of this repeated game has a distinct history of play leading to it, all of these subgames have an identical structure: Each is an infinitely repeated Bertrand duopoly game exactly like the game as a whole. Thus, to establish that the strategies in (12.D.1) constitute an SPNE, we need to show that after any previous history of play, the strategies specified for the remainder of the game constitute a Nash equilibrium of an infinitely repeated Bertrand game.

In fact, given the form of the strategies in (12.D.1), we need to be concerned with only two types of previous histories: those in which there has been a previous deviation (a price not equal to p^m) and those in which there has been no deviation.

Consider, first, a subgame arising after a deviation has occurred. The strategies call for each firm to set its price equal to c in every future period regardless of its rival's behavior. This pair of strategies is a Nash equilibrium of an infinitely repeated Bertrand game because each firm j can earn at most zero when its opponent always sets its price equal to c, and it earns exactly this amount by itself setting its price equal to c in every remaining period.

Now consider a subgame starting in, say, period t after no previous deviation has occurred. Each firm j knows that its rival's strategy calls for it to charge p^m until it encounters a deviation from p^m and to charge c thereafter. Is it in firm j's interest to use this strategy itself given that its rival does? That is, do these strategies constitute a Nash equilibrium in this subgame?

Suppose that firm j contemplates deviating from price p^m in period $\tau \geq t$ of the subgame if no deviation has occurred prior to period τ.[19] From period t through period $\tau - 1$, firm j will earn $\frac{1}{2}(p^m - c)x(p^m)$ in each period, exactly as it does if it never deviates. Starting in period τ, however, its payoffs will differ from those that would arise if it does not deviate. In periods after it deviates (periods $\tau + 1, \tau + 2, \ldots$), firm j's rival charges a price of c regardless of the form of firm j's deviation in period τ, and so firm j can earn at most zero in each of these periods. In period τ, firm j optimally deviates in a manner that maximizes its payoff in that period (note that the payoffs firm j receives in later periods are the same for any deviation from p^m that it makes). It will therefore charge $p^m - \varepsilon$ for some arbitrarily small $\varepsilon > 0$, make all sales in the market, and earn a one-period payoff of $(p^m - c - \varepsilon)x(p^m)$. Thus, its overall discounted payoff from period τ onward as a result of following this deviation strategy, discounted to period τ, can be made arbitrarily close to $(p^m - c)x(p^m)$.

On the other hand, if firm j never deviates, it earns a discounted payoff from period τ onward, discounted to period τ, of $[\frac{1}{2}(p^m - c)x(p^m)]/(1 - \delta)$. Hence, for any t and $\tau \geq t$, firm j will prefer no deviation to deviation in period τ if and only if

$$\frac{1}{1-\delta}[\tfrac{1}{2}(p^m - c)x(p^m)] \geq (p^m - c)x(p^m),$$

or

$$\delta \geq \tfrac{1}{2}. \tag{12.D.2}$$

Thus, the strategies in (12.D.1) constitute an SPNE if and only if $\delta \geq \frac{1}{2}$. ■

The implication of Proposition 12.D.1 is that the perfectly competitive outcome of the static Bertrand game may be avoided if the firms foresee infinitely repeated interaction. The reason is that, in contemplating a deviation, each firm takes into account not only the one-period gain it earns from undercutting its rival but also the profits forgone by triggering retaliation. The size of the discount factor δ is

19. From our previous argument, we know that once a deviation has occurred within this subgame, firm j can do no better than to play c in every period given that its rival will do so. Hence, to check whether these strategies form a Nash equilibrium in this subgame, we need only check whether firm j will wish to deviate from p^m if no such deviation has yet occurred.

important here because it affects the relative weights put on the future losses versus the present gain from a deviation. The monopoly price is sustainable if and only if the present value of these future losses is large enough relative to the possible current gain from deviation to keep the firms from going for short-run profits.

The discount factor need not be interpreted literally. For example, in a model in which market demand is growing at rate γ [i.e., $x_t(p) = \gamma^t x(p)$], larger values of γ make the model behave as if there is a larger discount factor because demand growth increases the size of any future losses caused by a current deviation. Alternatively, we can imagine that in each period there is a probability γ that the firms' interaction might end. The larger γ is, the more firms will effectively discount the future. (This interpretation makes clear that the infinitely repeated game framework can be relevant even when the firms may cease their interaction within some finite amount of time; what is needed to fit the analysis into the framework above is a strictly positive probability of continuing upon having reached any period.) Finally, the value of δ can reflect how long it takes to detect a deviation. These interpretations are developed in Exercise 12.D.1.

Although the strategies in (12.D.1) constitute an SPNE when $\delta \geq \frac{1}{2}$, they are *not* the only SPNE of the repeated Bertrand model. In particular, we can obtain the result presented in Proposition 12.D.2.

Proposition 12.D.2: In the infinitely repeated Bertand duopoly game, when $\delta \geq \frac{1}{2}$ repeated choice of any price $p \in [c, p^m]$ can be supported as a subgame perfect Nash equilibrium outcome path using Nash reversion strategies. By contrast, when $\delta < \frac{1}{2}$, any subgame perfect Nash equilibrium outcome path must have all sales occurring at a price equal to c in every period.

Proof: For the first part of the result, we have already shown in Proposition 12.D.1 that repeated choice of price p^m can be sustained as an SPNE outcome when $\delta \geq \frac{1}{2}$. The proof for any price $p \in [c, p^m)$ follows exactly the same lines; simply change price p^m in the strategies of (12.D.1) to $p \in [c, p^m)$.

The proof of the second part of the result is presented in small type.

We now show that all sales must occur at a price equal to c when $\delta < \frac{1}{2}$. To begin, let $v_{jt} = \sum_{\tau \geq t} \delta^{\tau - t} \pi_{jt}$ denote firm j's profits, discounted to period t, when the equilibrium strategies are played from period t onward. Also define $\pi_t = \pi_{1t} + \pi_{2t}$.

Observe that, because every firm j finds it optimal to conform to the equilibrium strategies in every period t, it must be that

$$\pi_t \leq v_{jt} \quad \text{for } j = 1, 2 \text{ and every } t, \tag{12.D.3}$$

since each firm j can obtain a payoff arbitrarily close to π_t in period t by deviating and undercutting the lowest price in the market by an arbitrarily small amount and can assure itself a nonnegative payoff in any period thereafter.

Suppose that there exists at least one period t in which $\pi_t > 0$. We will derive a contradiction. There are two cases to consider:

(i) Suppose, first, that there is a period τ with $\pi_\tau > 0$ such that $\pi_\tau \geq \pi_t$ for all t. If so, then adding (12.D.3) for $t = \tau$ over $j = 1, 2$, we have

$$2\pi_\tau \leq (v_{1\tau} + v_{2\tau}).$$

But $(v_{1\tau} + v_{2\tau}) \leq [1/(1 - \delta)]\pi_\tau$, and so this is impossible if $\delta < \frac{1}{2}$.

(ii) Suppose, instead, that no such period exists; that is, for any period t, there is a period $\tau > t$ such that $\pi_\tau > \pi_t$. Define $\tau(t)$ for $t \geq 1$ recursively as follows: Let $\tau(1) = 1$ and for $t \geq 2$ define $\tau(t) = \text{Min}\,\{\tau > \tau(t-1): \pi_\tau > \pi_{\tau(t-1)}\}$. Note that, for all t, π_t is bounded above by the monopoly profit level $\pi^m = (p^m - c)x(p^m)$ and that the sequence $\{\pi_{\tau(t)}\}_{t=1}^{\infty}$ is monotonically increasing. Hence, as $t \to \infty$, $\pi_{\tau(t)}$ must converge to some $\bar{\pi} \in (0, \pi^m]$ such that $\pi_t < \bar{\pi}$ for all t. Now, adding (12.D.3) over $j = 1, 2$, we see that we must have

$$2\pi_{\tau(t)} \leq v_{1\tau(t)} + v_{2\tau(t)} \tag{12.D.4}$$

for all t. Moreover, $v_{1\tau(t)} + v_{2\tau(t)} \leq [1/(1-\delta)]\bar{\pi}$ for all t, and so we must have

$$2\pi_{\tau(t)} \leq \frac{1}{1-\delta}\bar{\pi} \tag{12.D.5}$$

for all t. But when $\delta < \frac{1}{2}$, condition (12.D.5) must be violated for t sufficiently large.

This completes the proof of the proposition. ■

The presence of multiple equilibria identified in Proposition 12.D.2 for $\delta \geq \frac{1}{2}$ is common in infinitely repeated games. Typically, a range of cooperative equilibria is possible for a given level of δ, as is a complete lack of cooperation in the form of the static Nash equilibrium outcome repeated forever.

Proposition 12.D.2 also tells us that the set of SPNE of the repeated Bertrand game grows as δ gets larger.[20] The discontinuous behavior as a function of δ of the set of SPNE displayed in Proposition 12.D.2 is, however, a special feature of the repeated Bertrand model. The repeated Cournot model and models of repeated price competition with differentiated products generally display a smoother increase in the maximal level of joint profits that can be sustained as δ increases (see Exercise 12.D.3).

In fact, a general result in the theory of repeated games, known as the *folk theorem*, tells us the following: In an infinitely repeated game, *any feasible discounted payoffs that give each player, on a per-period basis, more than the lowest payoff that he could guarantee himself in a single play of the simultaneous-move component game can be sustained as the payoffs of an SPNE if players discount the future to a sufficiently small degree.* In Appendix A, we provide a more precise statement and extended discussion of the folk theorem for general repeated games. Its message is clear: Although infinitely repeated games allow for cooperative behavior, they also allow 1
for an *extremely wide range* of possible behavior.

The wide range of equilibria in repeated game models of oligopoly is somewhat disconcerting. From a practical point of view, how do we know which equilibrium behavior will arise? Can "anything happen" in oligopolistic markets? To get around this problem, researchers often assume that symmetrically placed firms will find the symmetric profit-maximizing equilibrium focal (see Section 8.D). However, even restricting attention to the case of symmetric firms, the validity of this assumption is likely to depend on the setting. For example, the history of an industry could make other equilibria focal: An industry that has historically been very noncooperative (maybe because δ has always been low) may find noncooperative outcomes more focal. The assumption that the symmetric profit-maximizing equilibrium arises is

20. Strictly speaking, Proposition 12.D.2 shows this only for the class of stationary, symmetric equilibria (i.e., equilibria in which the firms adopt identical strategies and in which, on the equilibrium path, the actions taken are the same in every period).

more natural when the self-enforcing agreement interpretation of these equilibria is relevant, as when oligopolists secretly meet to discuss their pricing plans. Because antitrust laws preclude oligopolists from writing a formal contract specifying their behavior, any secret collusive agreement among them must be self-enforcing and so must constitute an SPNE. It seems reasonable to think that, in such circumstances, identical firms will therefore agree to the most profitable symmetric SPNE. (If the firms are not identical, similar logic suggests that the firms would agree to an SPNE corresponding to a point on the frontier of their set of SPNE payoffs.)

Finally, just as with the static models discussed in Section 12.C, it is of interest to investigate how the number of firms in a market affects its competitiveness. You are asked to do so in Exercise 12.D.2.

Exercise 12.D.2: Show that with J firms, repeated choice of any price $p \in (c, p^m]$ can be sustained as a stationary SPNE outcome path of the infinitely repeated Bertrand game using Nash reversion strategies if and only if $\delta \geq (J-1)/J$. What does this say about the effect of having more firms in a market on the difficulty of sustaining collusion?

In practice, an important feature of many settings of oligopolistic collusion (as well as other settings of cooperation) is that firms are likely to be able to observe their rivals' behavior only imperfectly. For example, as emphasized by Stigler (1960), an oligopolist's rivals may make secret price cuts to consumers. If the market demand is stochastic, a firm will be unable to tell with certainty whether there have been any deviations from collusive pricing simply from observation of its own demand. This possibility leads formally to study of *repeated games with imperfect observability*; see, for example, Green and Porter (1984) and Abreu, Pearce, and Stachetti (1990). A feature of this class of models is that they are able to explain observed breakdowns of cooperation as being an inevitable result of attempts to cooperate in environments characterized by imperfect observability. This is so because equilibrium strategies must be such that some negative realizations of demand result in a breakdown of cooperation if firms are to be prevented from secretly deviating from a collusive scheme.

12.E Entry

In Sections 12.B to 12.D, we analyzed monopolistic and oligopolistic market outcomes, keeping the number of active firms exogenously fixed. In most cases, however, we wish to view the number of firms that will be operating in an industry as an endogeneous variable. Doing so also raises a new question regarding the welfare properties of situations in which market power is present: Is the equilibrium number of firms that enter the market socially efficient? In Section 10.F, we saw that the answer to this question is "yes" in the case of competitive markets as long as an equilibrium exists. In this section, however, we shall see that this is no longer true when market power is present.

We now take the view that there is an infinite (or finite but very large) number of potential firms, each of which could enter and produce the good under consideration if it were profitable to do so. As in Section 10.F, we focus on the case in which all potential firms are identical. (See Exercise 12.E.1 for a case in which they are not.)

A natural way to conceptualize entry in oligopolistic settings is as a two-stage process in which a firm first incurs some setup cost $K > 0$ in entering the industry and then, once this cost is sunk, competes for business. The simplest sort of model that captures this idea has the following structure:

Stage 1: All potential firms simultaneously decide "in" or "out." If a firm decides "in," it pays a setup cost $K > 0$.

Stage 2: All firms that have entered play some oligopolistic game.

The oligopoly game in stage 2 could be any of those considered in Sections 12.C and 12.D.

Formally, this two-stage entry model defines a dynamic game (see Chapter 9). Note that its stage 2 subgames are exactly like the games we have analyzed in the previous sections because, at that stage, the number of firms is fixed. Throughout our discussion we shall assume that for each possible number of active firms, there is a unique, symmetric (across firms) equilibrium in stage 2, and we let π_J denote the profits of a firm in this stage 2 equilibrium when J firms have entered (π_J does not include the entry cost K).

This two-stage entry model provides a very simple representation of the entry process. There is very little dynamic structure, and no firm has any "first-mover" advantage that enables it to deter entry or lessen competition from other firms (see Section 12.G and Appendix B for a discussion of these possibilities).

Consider now the (pure strategy) subgame perfect Nash equilibria (SPNEs) of this model. In any SPNE of this game, no firm must want to change its entry decision given the entry decisions of the other firms. For expositional purposes, we shall also adopt the convention that a firm chooses to enter the market when it is indifferent. With this assumption, there is an equilibrium with J^* firms choosing to enter the market if and only if

$$\pi_{J^*} \geq K \tag{12.E.1}$$

and

$$\pi_{J^*+1} < K. \tag{12.E.2}$$

Condition (12.E.1) says that a firm that has chosen to enter does at least as well by doing so as it would do if it were to change its decision to "out," given the anticipated result of competition with J^* firms. Condition (12.E.2) says that a firm that has decided to remain out of the market does strictly worse by changing its decision to "in," given the anticipated result of competition with $J^* + 1$ firms.

Typically, we expect that π_J is decreasing in J and that $\pi_J \to 0$ and $J \to \infty$. In this case, there is a unique integer $\hat{J}$ such that $\pi_J \geq K$ for all $J \leq \hat{J}$ and $\pi_J < K$ for all $J > \hat{J}$, and so $J^* = \hat{J}$ is the unique equilibrium number of firms.[21,22]

21. Note, however, that although there is a unique number of entrants, there are many equilibria, in each of which the particular firms choosing to enter differ.

22. Without the assumption that firms enter when indifferent, condition (12.E.2) would be a weak inequality. This change in (12.E.2) matters for the identification of the equilibrium number of firms only in the case in which there is an integer number of firms $\tilde{J}$ such that $\pi_{\tilde{J}} = K$ (so that with $\tilde{J}$ firms in the market each firm earns exactly zero net of its entry cost K). When this is so, this change allows both $\tilde{J}$ and $\tilde{J} - 1$ to be equilibria. With minor adaptations but some loss of expositional simplicity, all the points made in this section can be extended to cover this case.

We illustrate the determination of the equilibrium number of firms with two examples in which the stage 2 oligopoly games correspond, respectively, to the Cournot and Bertrand models discussed in Section 12.C.

Example 12.E.1: *Equilibrium Entry with Cournot Competition.* Suppose that competition in stage 2 of the two-stage entry game corresponds to the Cournot model studied in Section 12.C, with $c(q) = cq$, $p(q) = a - bq$, $a > c \geq 0$, and $b > 0$. The stage 2 output per firm, q_J, and profit per firm, π_J, are given (see Exercise 12.C.7) by

$$q_J = \left(\frac{a-c}{b}\right)\left(\frac{1}{J+1}\right), \tag{12.E.3}$$

$$\pi_J = \left(\frac{a-c}{J+1}\right)^2\left(\frac{1}{b}\right). \tag{12.E.4}$$

Note that π_J is strictly decreasing in J and that $\pi_J \to 0$ as $J \to \infty$. Also, $Jq_J \to (a-c)/b$ as $J \to \infty$, so that aggregate quantity approaches the competitive level. Solving for the real number $\tilde{J} \in \mathbb{R}$ at which $\pi_{\tilde{J}} = K$ gives

$$(\tilde{J}+1)^2 = \frac{(a-c)^2}{bK}$$

or

$$\tilde{J} = \frac{(a-c)}{\sqrt{bK}} - 1.$$

The equilibrium number of entrants J^* is the largest integer that is less than or equal to $\tilde{J}$. Note that as K decreases, the number of firms active in the market (weakly) increases, and that as more firms become active, aggregate output increases and price decreases. Indeed, $J^* \to \infty$ as $K \to 0$, and output and price approach their competitive levels. Note also that a proportional increase in demand at every price, captured by a reduction in b, changes the equilibrium number of firms and price in a manner that is identical to a decrease in K. ■

Example 12.E.2: *Equilibrium Entry with Bertrand Competition.* Suppose now that competition in stage 2 of the two-stage entry game takes the form of the Bertrand model studied in Section 10.C. Once again, $c(q) = cq$, $p(q) = a - bq$, $a > c \geq 0$, and $b > 0$. Now $\pi_1 = \pi^m$, the monopoly profit level, and $\pi_J = 0$ for all $J \geq 2$. Thus, assuming that $\pi^m > K$, the SPNE must have $J^* = 1$ and result in the monopoly price and quantity levels. Comparing this result with the result in Example 12.E.1 for the Cournot model, we see that the presence of more intense stage 2 competition here actually *lowers* the ultimate level of competition in the market! ■

Entry and Welfare

Consider now how the number of firms entering an oligopolistic market compares with the number that would maximize social welfare given the presence of oligopolistic competition in the market. We begin by considering this issue for the case of a homogeneous-good industry.

Let q_J be the symmetric equilibrium output per firm when there are J firms in the market. As usual, the inverse demand function is denoted by $p(\cdot)$. Thus, $p(Jq_J)$ is the price when there are J active firms; and so $\pi_J = p(Jq_J)q_J - c(q_J)$, where $c(\cdot)$ is the cost function of a firm after entry. We assume that $c(0) = 0$.

We measure welfare here by means of Marshallian aggregate surplus (see Section 10.E). In this case, social welfare when there are J active firms is given by

$$W(J) = \int_0^{Jq_J} p(s)\,ds - Jc(q_J) - JK. \tag{12.E.5}$$

The socially optimal number of active firms in this oligopolistic industry, which we denote by J^o, is any integer number that solves $\text{Max}_J\, W(J)$. Example 12.E.3 illustrates that in contrast with the conclusion arising in the case of a competitive market, the equilibrium number of firms here need not be socially optimal.

Example 12.E.3: Consider the Cournot model of Example 12.E.1. For the moment, ignore the requirement that the number of firms is an integer, and solve for the number of firms $\bar{J}$ at which $W'(\bar{J}) = 0$. This gives

$$(\bar{J} + 1)^3 = \frac{(a - c)^2}{bK}. \tag{12.E.6}$$

If $\bar{J}$ turns out to be an integer, then the socially optimal number of firms is $J^o = \bar{J}$. Otherwise, J^o is one of the two integers on either side of $\bar{J}$ [recall that $W(\cdot)$ is concave]. Now, recall from (12.E.4) that $\pi_J = (1/b)[(a - c)/(J + 1)]^2$. As noted in Example 12.E.1, if we let $\tilde{J}$ be the real number such that

$$(\tilde{J} + 1)^2 = \frac{(a - c)^2}{bK}, \tag{12.E.7}$$

the equilibrium number of firms is the largest integer less than or equal to $\tilde{J}$. From (12.E.6) and (12.E.7), we see that

$$(\tilde{J} + 1) = (\bar{J} + 1)^{3/2}.$$

Thus, when the demand and cost parameters are such that the optimal number of firms is exactly two ($J^o = \bar{J} = 2$), four firms actually enter this market ($J^* = 4$, since $\tilde{J} \simeq 4.2$); when the social optimum is for exactly three firms to enter ($J^o = \bar{J} = 3$), seven firms actually do ($J^* = 7$, since $\tilde{J} = 7$); when the social optimum is for exactly eight firms to enter ($J^o = \bar{J} = 8$), 26 actually enter ($J^* = 26$, since $\tilde{J} = 26$). ■

Can we say anything general about the nature of the entry bias? It turns out that we can as long as stage 2 competition satisfies three weak conditions [we follow Mankiw and Whinston (1986) here]:

(A1) $Jq_J \geq J'q_{J'}$ whenever $J > J'$;
(A2) $q_J \leq q_{J'}$ whenever $J > J'$;
(A3) $p(Jq_J) - c'(q_J) \geq 0$ for all J.

Conditions (A1) and (A3) are straightforward: (A1) requires that aggregate output increases (price falls) when more firms enter the industry, and (A3) says that price is not below marginal cost regardless of the number of firms entering the industry. Condition (A2) is more interesting. It is the assumption of *business stealing*. It says that when an additional firm enters the market, the sales of existing firms fall (weakly). Hence, part of the new firm's sales come at the expense of existing firms. These conditions are satisfied by most, although not all, oligopoly models. [In the Bertrand model, for example, condition (A3) does not hold.]

For markets satisfying these three conditions we have the result shown in Proposition 12.E.1.

Proposition 12.E.1: Suppose that conditions (A1) to (A3) are satisfied by the post-entry oligopoly game, that $p'(\cdot) < 0$, and that $c''(\cdot) \geq 0$. Then the equilibrium number of entrants, J^*, is at least $J^\circ - 1$, where J° is the socially optimal number of entrants.[23]

Proof: The result is trivial for $J^\circ = 1$, so suppose that $J^\circ > 1$. Under the assumptions of the proposition, π_J is decreasing in J (Exercise 12.E.2 asks you to show this). To establish the result, we therefore need only show that $\pi_{J^\circ - 1} \geq K$.

To prove this, note first that by the definition of J° we must have $W(J^\circ) - W(J^\circ - 1) \geq 0$, or

$$\int_{Q_{J^\circ - 1}}^{Q_{J^\circ}} p(s)\,ds - J^\circ c(q_{J^\circ}) + (J^\circ - 1)c(q_{J^\circ - 1}) \geq K,$$

where we let $Q_J = Jq_J$. We can rearrange this expression to yield

$$\pi_{J^\circ - 1} - K \geq p(Q_{J^\circ - 1})q_{J^\circ - 1} - \int_{Q_{J^\circ - 1}}^{Q_{J^\circ}} p(s)\,ds + J^\circ[c(q_{J^\circ}) - c(q_{J^\circ - 1})].$$

Given $p'(\cdot) < 0$ and condition (A1), this implies that

$$\pi_{J^\circ - 1} - K \geq p(Q_{J^\circ - 1})[q_{J^\circ - 1} + Q_{J^\circ - 1} - Q_{J^\circ}] + J^\circ[c(q_{J^\circ}) - c(q_{J^\circ - 1})]. \quad (12.E.8)$$

But since $c''(\cdot) \geq 0$, we know that $c'(q_{J^\circ - 1})[q_{J^\circ} - q_{J^\circ - 1}] \leq c(q_{J^\circ}) - c(q_{J^\circ - 1})$. Using this inequality with (12.E.8) and the fact that $q_{J^\circ - 1} + Q_{J^\circ - 1} - Q_{J^\circ} = J^\circ(q_{J^\circ - 1} - q_{J^\circ})$ yields

$$\pi_{J^\circ - 1} - K \geq [p(Q_{J^\circ - 1}) - c'(q_{J^\circ - 1})]J^\circ(q_{J^\circ - 1} - q_{J^\circ}).$$

Conditions (A2) and (A3) then imply that $\pi_{J^\circ - 1} \geq K$.[24] ∎

The idea behind the proof of Proposition 12.E.1 is illustrated in Figure 12.E.1 for the case where $c(q) = 0$ for all q. In the figure, the incremental welfare benefit of the J°th firm, before taking its entry cost into account, is represented by the shaded area $(abcd)$. Since entry of this firm is socially efficient, this area must be at least K. But area $(abcd)$ is less than area $(abce)$, which equals $p(Q_{J^\circ - 1})(Q_{J^\circ} - Q_{J^\circ - 1})$. Moreover, business stealing implies that $(Q_{J^\circ} - Q_{J^\circ - 1}) = J^\circ q_{J^\circ} - (J^\circ - 1)q_{J^\circ - 1} \leq q_{J^\circ - 1}$, and so we see that area $(abce) \leq p(Q_{J^\circ - 1})q_{J^\circ - 1} = \pi_{J^\circ - 1}$ [the value of $\pi_{J^\circ - 1}$ is represented in Figure 12.E.1 by area $(abfg)$]. Hence $\pi_{J^\circ - 1} \geq K$.

The tendency for excess entry in the presence of market power is fundamentally driven by the business-stealing effect. When business stealing accompanies new entry and price exceeds marginal cost, part of a new entrant's profit comes at the expense of existing firms, creating an excess incentive for the new firm to enter.

Of course, as Proposition 12.E.1 indicates, we may also see too few firms in an industry. The classic example concerns a situation in which the socially optimal number of firms is one. A single firm deciding whether to enter a market as a

23. If there is more than one maximizer of $W(J)$, say $\{J_1^\circ, \ldots, J_N^\circ\}$, then $J^* \geq \text{Max}\{J_1^\circ, \ldots, J_N^\circ\} - 1$.

24. Note that if (A1) holds with strict inequality, then this conclusion can be strengthened to $\pi_{J^\circ - 1} > K$ [a strict inequality appears in (12.E.8)]. In this case, $J^* \geq J^\circ - 1$ even if firms do not enter when indifferent.

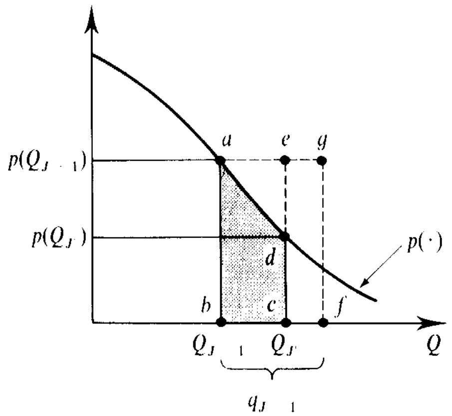


Figure 12.E.1 (left) Diagrammatic explanation of Proposition 12.E.1.

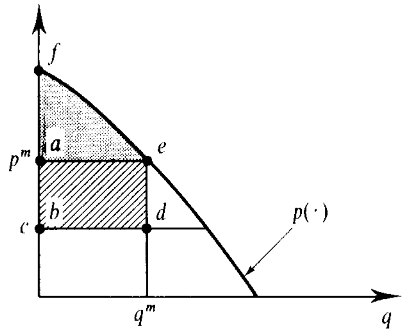


Figure 12.E.2 (right) An insufficient entry incentive.

monopolist compares its monopoly profit—the hatched area ($abde$) in Figure 12.E.2—with the entry cost K. However, the firm fails to capture, and therefore ignores, the increase in consumer surplus that its entry generates—the shaded area (fae). As a result, the firm may find entry unprofitable even though it is socially desirable. Proposition 12.E.1 tells us, however, that if we have too little entry in a homogeneous-good market, this can be at most by a single firm.

What happens when product differentiation is present? It turns out that we can then say very little of a general nature. The reason is that the sort of problem illustrated in Figure 12.E.2 can now happen for many products, leading to many "too few by one" conclusions. An additional issue is that, with product differentiation, the number of firms is not all that matters. We may also fail to have the right selection of products.[25]

An alternative approach to the two-stage entry game models the actions of entry and quantity/price choice as simultaneous. In this *one-stage entry game*, a firm incurs its setup cost only if it sells a positive amount. For example, the one-stage versions of Examples 12.E.1 and 12.E.2 are Cournot and Bertrand games, respectively, with cost functions

$$C(q) = \begin{cases} K + c(q) & \text{if } q > 0 \\ 0 & \text{if } q = 0 \end{cases}$$

and an infinite (or very large) number of firms. For models of price competition, this change can have dramatic consequences. Consider the effect on the result of Example 12.E.2 that is illustrated in Example 12.E.4.

Example 12.E.4: *The One-Stage Entry Model with Bertrand Competition.* Suppose that $p > [K + cx(p)]/x(p)$ for some p (the parameter $c > 0$ is the cost per unit); that is, suppose there is some price level at which a monopolist can earn strictly positive profits after paying its set up cost K. Assume that many firms simultaneously name prices and that a firm incurs the setup cost K only if it actually makes sales. Any equilibrium of this game has all sales occurring at price $p^* = \text{Min}\{p\colon p \geq [K + cx(p)]/x(p)\}$ (if price is above p^*, some firm could gain by setting a price $p^* - \varepsilon$; if price is below p^*, some firm must be making strictly negative profits), and one firm satisfying all demand at this price (if the demand were split among

25. See Spence (1976), Dixit and Stiglitz (1977), Salop (1979), and Mankiw and Whinston (1986) for more on the case of product differentiation.

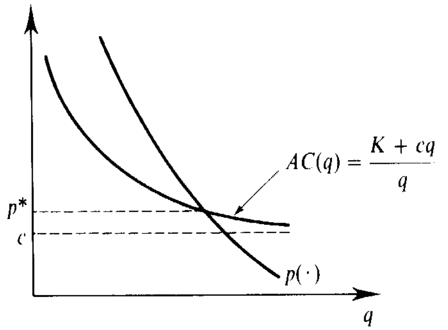


Figure 12.E.3
Equilibrium in the one-stage entry game discussed in Example 12.E.4.

several firms at price p^*, none of them could cover their cost).[26] In this equilibrium, all firms make zero profits. The equilibrium outcome is depicted in Figure 12.E.3. Observe that it is strictly superior in welfare terms to the outcome that arises in the two-stage entry process considered in Example 12.E.2, where there is also a single firm active but it quotes a monopoly price.[27] ■

What is the critical difference between the one-stage and two-stage entry processes? In the two-stage model an entrant must sink its fixed costs prior to competing, whereas in the one-stage model it can compete for sales while retaining the option not to sink these costs if it does not make any sales. We can think of the two-stage case as a model of a firm incurring a once-and-for-all sunk entry cost that allows for many later periods of competitive interaction, whereas the one-stage case captures a setting in which "hit-and-run" entry is possible (i.e., entry for just one period while paying only the one-period rental price of capital). When a firm must incur a sunk cost in entering it must consider the reaction of other firms to its entry. In the Bertrand model with constant costs this reaction is severe: price falls to cost and the firm loses money by entering. In contrast, in the one-stage game the firm can enter and undercut active firms' prices without fearing their reactions. This makes entry more aggressive and leads to a lower equilibrium price. This one-stage entry model with price competition provides one formalization of what Baumol, Panzar, and Willig (1982) call a *contestable market*.

12.F The Competitive Limit

In Chapter 10, we introduced the idea that a competitive market might usefully be thought of as a limiting case of an oligopolistic market in which firms' market power grows increasingly small (see Section 10.B). We also noted that this view could provide a framework for reconciling cases in which competitive equilibria fail to exist in the presence of free entry and average costs that exhibit a strictly positive efficient

26. Note that we now allow consumer demand to be given entirely to one firm when several firms name the same price (before, we had taken the division of demand in this case to be exogenously given). This is the only division of demand that is compatible with equilibrium in this example. It can be formally justified as the limit of the equilibria that arise when prices must be quoted in discrete units as the size of these units grows small.

27. In fact, this equilibrium outcome is the solution to the problem faced by a welfare-maximizing planner who can control the outputs q_j of the firms but must guarantee a nonnegative profit to all active firms, that is, who faces the constraint that $p(\sum_k q_k)q_j \geq cq_j + K$ for every j with $q_j > 0$.

scale (see Section 10.F). In this situation, we argued, as long as many firms could fit into the market, the market outcome ought to be close to the competitive outcome that would arise if industry average costs were actually constant at the level of minimum average cost. In this section, we elaborate on these points and develop, in a setting of free entry, the theme that if the size of individual firms is small relative to the size of the market, then the equilibrium will be nearly competitive.

We have already seen one example of this phenomenon in Example 12.E.1. Here we establish the point in a more general way. We now let market demand be $x_\alpha(p) = \alpha x(p)$, where $x(p)$ is differentiable and $x'(\cdot) < 0$. Increases in α correspond to proportional increases in demand at all prices. Letting $p(q)$ be the inverse demand function associated with $x(p)$, the inverse demand function associated with $x_\alpha(p)$ is then $p_\alpha(q) = p(q/\alpha)$. All potential firms have a strictly convex cost function $c(q)$ and entry cost $K > 0$. We denote the level of minimum average cost for a firm by $\bar{c} = \text{Min}_{q>0}\,[K + c(q)]/q$, and we let $\bar{q} > 0$ denote a firm's (unique) efficient scale.

As in Example 12.E.1, we focus on the case of a two-stage entry model with Cournot competition in the second stage, in which the cost K is incurred only if the firm decides to enter in stage 1. We let $b(Q_{-j})$ denote active firm j's optimal output level for any given level of aggregate output by its rivals, Q_{-j}, and we assume that this best response is unique for all Q_{-j}.

Finally, we let p_α and Q_α denote the price and aggregate output in a subgame perfect Nash equilibrium (SPNE) of the two-stage Cournot entry model when the market size is α. We denote by P_α the set of all SPNE prices for market size α.

Proposition 12.F.1: As the market size grows, the price in any subgame perfect Nash equilibrium of the two-stage Cournot entry model converges to the level of minimum average cost (the "competitive" price). Formally,

$$\underset{p_\alpha \in P_\alpha}{\text{Max}}\,|p_\alpha - \bar{c}| \to 0 \quad \text{as} \quad \alpha \to \infty.$$

Proof: The argument consists of three steps:

(i) First, you are asked in Exercise 12.F.1 to show that for large enough α, an active firm's best-response function $b(Q_{-j})$ is (weakly) decreasing in Q_{-j}.

(ii) Second, we argue that if $b(Q_{-j})$ is decreasing, then we must have $Q_\alpha \geq \alpha x(\bar{c}) - \bar{q}$ in any SPNE of the two-stage entry game with market size α. To see why this is so, suppose that with market size α we had an SPNE with J_α firms entering and an aggregate output level $Q_\alpha < \alpha x(\bar{c}) - \bar{q}$. Consider any firm j whose equilibrium entry choice is "out" in this equilibrium, and suppose that firm j instead decided to enter and produce quantity $\bar{q}$. Because $b(\cdot)$ is decreasing, it is intuitively plausible that the aggregate output level of the original J_α active firms cannot increase when firm j enters in this way (see the small-type paragraph that follows for the formal argument behind this claim). As a result, aggregate output in the market following firm j's entry is no more than $(Q_\alpha + \bar{q})$; and since $(Q_\alpha + \bar{q}) < \alpha x(\bar{c})$, the resulting (post-entry) price is above $\bar{c}$. Hence, firm j would earn strictly positive profits by entering in this fashion, contradicting the hypothesis that we were at an SPNE to start with.

The argument that the output of the existing J_α firms cannot increase following entry of firm j is as follows: Let Q_{-j} be the initial equilibrium level of these firms' aggregate output, and let $\tilde{Q}_{-j}$ be their post-entry aggregate output. Suppose that $\tilde{Q}_{-j} > Q_{-j}$. Then at least one of these firms, say firm k, must have increased its output level in response to firm j's entry, say from q_k to $\tilde{q}_k > q_k$. Because $b(\cdot)$ is decreasing, it must be that $\tilde{Q}_{-k} < Q_{-k}$; that is, the post-entry output $\tilde{Q}_{-k}$ of active firms other than k (which includes firm j) must be less than their pre-entry output, Q_{-k}. By part (c) of Exercise 12.C.8, this implies that $q_k + Q_{-k} \geq \tilde{q}_k + \tilde{Q}_{-k}$. But $Q_{-j} = q_k + Q_{-k}$ (since firm j initially produces nothing), and $\tilde{q}_k + \tilde{Q}_{-k} \geq \tilde{Q}_{-j}$ (because firm j's post-entry output is nonnegative). Hence, $Q_{-j} \geq \tilde{Q}_{-j}$, which is a contradiction.

(iii) Finally, we argue that the conclusion of (ii) implies the result. To see this, consider how much above $\bar{c}$ the price can be if aggregate output is no more than $\bar{q}$ below $\alpha x(\bar{c})$. This is given by

$$\begin{aligned}\Delta p_\alpha &= p_\alpha(\alpha x(\bar{c}) - \bar{q}) - p_\alpha(\alpha x(\bar{c})) \\ &= p\left(\frac{\alpha x(\bar{c}) - \bar{q}}{\alpha}\right) - p(x(\bar{c})).\end{aligned}$$

But as $\alpha \to \infty$, $[\alpha x(\bar{c}) - \bar{q}]/\alpha \to x(\bar{c})$, so that $\Delta p_\alpha \to 0$. ■

There are two forces driving Proposition 12.F.1. First, the entry process ensures that firms will enter if there is too much "room" left in the market. Second, in a market that is very large relative to the minimum efficient scale, a reduction of output equal to the level of minimum efficient scale has very little effect on price. The consequence of these two facts is that as the market size grows large, firms' market power is dissipated and price approaches the level of minimum average cost (the competitive level). In this limiting outcome, welfare approaches its optimal level.[28]

In Example 12.E.2, we saw that in a two-stage Bertrand market, no such limiting result holds.[29] Because price drops to marginal cost if even two firms enter, the market is always monopolized, no matter what its size. However, the two-stage Bertrand model's limiting properties are quite special. As long as, for any market size, price is above marginal cost for any finite number of firms that enter the market, and approaches marginal cost as the number of firms grows large, a limiting result like that in Proposition 12.F.1 holds.

Finally, Proposition 12.F.1 applies only for the case of homogeneous-good markets. With product differentiation, we must be careful. Firms may be small relative to the size of the entire set of interrelated markets, but they may still be large relative to their own particular niche. In this case, each firm may maintain substantial market power even in the limit, and the limiting equilibrium can be far from efficient (see Exercise 12.F.4).

28. The sense of approximation is relative to the size parameter of the market α. Assuming that α is a proxy for the number of consumers, this means that the welfare loss per consumer relative to the social optimum goes to zero.

29. Strictly speaking, firms' cost functions in Example 12.E.2 differ from the cost functions assumed in Proposition 12.F.1 (average costs including K are declining everywhere in Example 12.E.2). Nevertheless, for the two-stage Cournot model, Proposition 12.F.1 can be shown to be valid for the cost function of Example 12.E.2 (letting $\bar{c}$ in the statement of the proposition now be the limiting value of average cost as a firm's output grows large).

12.G Strategic Precommitments to Affect Future Competition

An important feature of many oligopolistic settings is that firms attempt to make strategic precommitments in order to alter the conditions of future competition in a manner that is favorable to them. Examples of strategic precommitments abound. For example, investments in cost reduction, capacity, and new-product development all lead to long-lasting changes that can affect the nature of future competition. In practice, these types of decisions can be among the most important competitive decisions that firms make.

Some general features of these types of strategic precommitments can be usefully illuminated through examination of the following simple two-stage duopoly model:

Stage 1: Firm 1 has the option to make a strategic investment, whose level we denote by $k \in \mathbb{R}$. This choice is observable.

Stage 2: Firms 1 and 2 play some oligopoly game, choosing strategies $s_1 \in S_1 \subset \mathbb{R}$ and $s_2 \in S_2 \subset \mathbb{R}$, respectively. Given investment level k and strategy choices (s_1, s_2), profits for firms 1 and 2 are given by $\pi_1(s_1, s_2, k)$ and $\pi_2(s_1, s_2)$, respectively.

For example, k might be an investment that reduces firm 1's marginal cost of production with the stage 2 game being Cournot competition (so $s_j = q_j$, firm j's quantity choice). Alternatively, stage 2 competition could be differentiated products price competition.

We suppose that there is a unique Nash equilibrium in stage 2 given any choice of k, $(s_1^*(k), s_2^*(k))$, and we assume for convenience that it is differentiable in k. We also assume for purposes of our discussion that $\partial \pi_1(s_1, s_2, k)/\partial s_2 < 0$ and $\partial \pi_2(s_1, s_2)/\partial s_1 < 0$, that is, that stage 2 actions are "aggressive" in the sense that a higher level of s_{-j} by firm j's rival lowers firm j's profit. Hence, firm 1 would be better off, all else being equal, if it could induce firm 2 to lower its choice of s_2.

When can investment by firm 1 cause firm 2 to lower s_2? Letting $b_1(s_2, k)$ and $b_2(s_1)$ denote firm 1's and firm 2's stage 2 best-response functions (note that firm 1's best response depends on k), we can differentiate the equilibrium condition $s_2^* = b_2(b_1(s_2^*, k))$ to get

$$\frac{ds_2^*(k)}{dk} = \frac{db_2(s_1^*(k))}{ds_1}\left(\frac{\partial b_1(s_2^*(k), k)/\partial k}{1 - [\partial b_1(s_2^*(k), k)/\partial s_2][db_2(s_1^*(k))/ds_1]}\right). \qquad (12.G.1)$$

The denominator of the second term on the right-hand side of (12.G.1) being nonegative is often called the *stability condition.* It implies that the simple dynamic adjustment process in which the firms take turns myopically playing a best response to each others' current strategies converges to the Nash equilibrium from any strategy pair in a neighborhood of the equilibrium. We shall maintain this assumption for the remainder of our discussion. Thus, the effect of k on s_2 can be seen to depend on two factors: (i) Does k make firm 1 more or less "aggressive" in stage 2 competition [i.e., what is the sign of $\partial b_1(s_2^*(k), k)/\partial k$]? and (ii) Does firm 2 respond to the anticipation of more aggressive play by firm 1 with more aggression itself or with less [i.e., what is the sign of $db_2(s_1^*(k))/ds_1$]?

	Strategic Substitutes: $\dfrac{db_2(\cdot)}{ds_1} < 0$	Strategic Complements $\dfrac{db_2(\cdot)}{ds_1} > 0$
$\dfrac{\partial b_1(\cdot)}{\partial k} > 0$	$\dfrac{ds_2^*(k)}{dk} < 0$	$\dfrac{ds_2^*(k)}{dk} > 0$
$\dfrac{\partial b_1(\cdot)}{\partial k} < 0$	$\dfrac{ds_2^*(k)}{dk} > 0$	$\dfrac{ds_2^*(k)}{dk} < 0$

Figure 12.G.1
Determinants of the sign of $ds_2^*(k)/dk$.

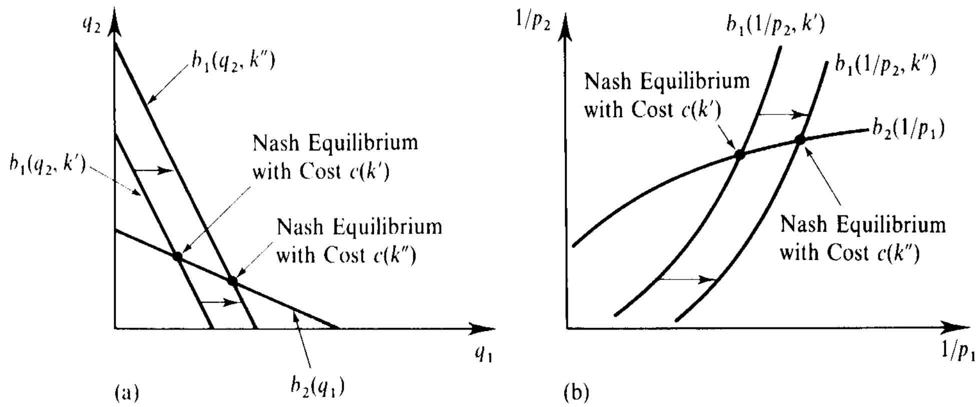


Figure 12.G.2
Strategic effects of a reduction in marginal cost from $c(k')$ to $c(k'') < c(k')$.
(a) Quantity model.
(b) Price model.

When firm 2 responds in kind to more aggressive choices of s_1 by firm 1 [i.e., when $db_2(s_1^*(k))/ds_1 > 0$], we say that s_2 is a *strategic complement* of s_1; and if firm 2 becomes less aggressive in the face of more aggressive play by firm 1 [i.e., if $db_2(s_1^*(k))/ds_1 < 0$], s_2 is a *strategic substitute* of s_1. [This terminology is derived from Bulow, Geanakoplos, and Klemperer (1985); see also Fudenberg and Tirole (1984) for a related taxonomy.]

Figure 12.G.1 summarizes these two determinants of firm 2's response, $ds_2^*(k)/dk$.

Example 12.G.1: *The Strategic Effects from Investment in Marginal Cost Reduction.* The importance for strategic behavior of the distinction between cases of strategic complements and strategic substitutes is nicely illustrated by examining the strategic effects of investments in marginal cost reduction for models of quantity versus price competition.

Suppose that if firm 1 invests k then its (constant) per-unit production costs are $c(k)$, where $c'(k) < 0$. Consider, first, the case in which stage 2 competition takes the form of the Cournot model of Example 12.C.1, so that the stage 2 strategic variable is $s_j = q_j$, firm j's quantity choice. In this model, we have a situation of strategic substitutes because firm 2's best-response function in stage 2 is downward sloping $[db_2(q_1)/dq_1 < 0$ at all q_1 such that $b_2(q_1) > 0]$. As shown in Figure 12.G.2(a), the lowering of firm 1's marginal cost because of an increase in k from, say, k' to $k'' > k'$, shifts firm 1's best-response function outward from $b_1(q_2, k')$ to $b_1(q_2, k'')$; with lower marginal costs, firm 1 will wish to produce more for any quantity choice of its rival [and so, in terms of our earlier analysis, $\partial b_1(q_2^*(k), k)/\partial k > 0$]. Thus, in this model, investment in cost reduction leads to a reduction in firm 2's output level, an effect that is beneficial for firm 1 [see Figure 12.G.2(a)].

In contrast, suppose that stage 2 competition takes the form of the differentiated price competition model of Example 12.C.2. Here we take $s_j = (1/p_j)$ to conform with the interpretation of s_j as an "aggressive" variable [i.e., $\partial \pi_1(s_1, s_2, k)/\partial s_2 < 0$]. In this model, we have a situation of strategic complements: an anticipated reduction in firm 1's price causes firm 2 to reduce its price also [i.e., $db_2(1/p_1)/d(1/p_1) > 0$]. As depicted in Figure 12.G.2(b), a reduction in firm 1's marginal cost because of an increase in k from k' to $k'' > k'$ once again makes firm 1 more aggressive, leading it to choose a lower price given any price choice of its rival; its best-response function shifts to the right from $b_1(1/p_2, k')$ to $b_1(1/p_2, k'')$ [hence, in terms of our earlier analysis, $\partial b_1(1/p_2^*(k), k)/\partial k > 0$]. With strategic complements, the result of the reduction in firm 1's marginal cost is therefore to lower firm 2's equilibrium price, an effect that is undesirable for firm 1.

Thus, the strategic effects of a reduction in firm 1's marginal cost differ between the two models, being beneficial to firm 1 in the quantity model and detrimental in the price model.[30] Which model more accurately captures the nature of competitive interaction depends on the particulars of an industry's situation. For example, if firms in a mature industry have excess capacity, the price model is likely to be more descriptive, and the strategic effect will be detrimental. On the other hand, in a new market where firms are investing in capacity, the strategic effect is likely to be better captured by the quantity model (recall our interpretation of the Cournot model in terms of capacity choices in Section 12.C). ■

In deciding on its level of investment, firm 1 must therefore consider not only the direct effects of its investment (say, the direct benefit of lower costs), but also the strategic effects that arise through induced changes in its its rival's behavior. Formally, the derivative of firm 1's profits with respect to a change in k can be written as

$$\frac{d\pi_1(s_1^*(k), s_2^*(k), k)}{dk} = \frac{\partial \pi_1(s_1^*(k), s_2^*(k), k)}{\partial k} + \frac{\partial \pi_1(s_1^*(k), s_2^*(k), k)}{\partial s_1}\frac{ds_1^*(k)}{dk} + \frac{\partial \pi_1(s_1^*(k), s_2^*(k), k)}{\partial s_2}\frac{ds_2^*(k)}{dk}.$$

Since at a Nash equilibrium in stage 2 given investment level k we have $\partial \pi_1(s_1^*(k), s_2^*(k), k)/\partial s_1 = 0$, this simplifies to

$$\frac{d\pi_1(s_1^*(k), s_2^*(k), k)}{dk} = \frac{\partial \pi_1(s_1^*(k), s_2^*(k), k)}{\partial k} + \frac{\partial \pi_1(s_1^*(k), s_2^*(k), k)}{\partial s_2}\frac{ds_2^*(k)}{dk}. \quad (12.G.2)$$

The first term on the right-hand side of (12.G.2) is the *direct effect* on firm 1's profits from changing k; the second term is the *strategic effect* that arises because of firm 2's equilibrium response to the change in k. Since $\partial \pi_1(s_1^*(k), s_2^*(k), k)/\partial s_2 < 0$, the strategic effect on firm 1's profits is positive if $ds_2^*(k)/dk < 0$, that is, if firm 2's response to increases in firm 1's investment is to lower its choice of s_2.

30. Best-response functions need not always slope this way in the price and quantity models, but the particular examples considered here represent the "normal" cases; see Exercise 12.C.12.

In the above discussion, we have considered situations in which a firm makes a strategic precommitment to affect future competition with another firm who is (or will be) in the market. A particularly striking example of strategic precommitment to affect future market conditions, however, arises when one firm is the first into an industry and seeks to use its first-mover advantage to deter further entry into its market. We can analyze this case formally by introducing a stage between stages 1 and 2, say stage 1.5, at which firm 2 decides whether to be in the market and by supposing that if firm 2 chooses "in" then it must pay a set-up cost $F > 0$. Firm 2 will therefore choose "out" given firm 1's stage 1 choice of k if its anticipated profit in stage 3, $\pi_2(s_1^*(k), s_2^*(k))$, is less than F. Given this fact, the incumbent would, of course, like simply to announce that in response to any entry it will engage in predatory pricing (i.e., it will choose a very high level of s_1 in stage 3). The problem, however, is that this threat must be *credible* (recall the discussion in Chapter 9). Thus, what the incumbent needs to do to deter entry is choose a level of k that precommits it to sufficiently aggressive behavior that firm 2 chooses not to enter. In any particular problem, this may or may not be possible, and it may or may not be profitable. As a general matter, there are many potential mechanisms (i.e., many types of variables k) by which such precommitments can be made. In Appendix B, we examine in some detail the classic mechanism of entry deterrence through capacity expansion first studied by Spence (1977) and Dixit (1980).

APPENDIX A: INFINITELY REPEATED GAMES AND THE FOLK THEOREM

In this appendix, we extend the discussion in Section 12.D of infinitely repeated games to a more general setting. Our primary aim is to develop a formal statement of a version of the *folk theorem* of infinitely repeated games. Infinitely repeated games have a very rich theoretical structure and we shall only touch on a limited number of their properties. Fudenberg and Tirole (1992) and Osborne and Rubinstein (1994) provide more extended discussions.

The Model

An infinitely repeated game consists of an infinite sequence of repetitions of a one-period simultaneous-move game, known as the *stage game*. For expositional simplicity, we focus here on the case in which there are two players.

In the one-period stage game, each player i has a compact strategy set S_i; $q_i \in S_i$ is a particular feasible action for player i. Denote $q = (q_1, q_2)$ and $S = S_1 \times S_2$. Player i's payoff function is $\pi_i(q_i, q_j)$. We restrict our attention throughout to pure strategies. It will be convenient to define player i's one-period best-response payoff given that his rival plays q_j by $\hat{\pi}_i(q_j) = \text{Max}_{q \in S_i}\, \pi_i(q, q_j)$.[31] We assume that the stage game has a unique pure strategy Nash equilibrium $q^* = (q_1^*, q_2^*)$ (the assumption of uniqueness is for expositional simplicity only).

In the infinitely repeated game, actions are taken and payoffs are earned at the beginning of each period. The players discount payoffs with discount factor $\delta < 1$.

31. We assume that conditions on the sets S_i and functions $\pi_i(q_i, q_j)$ hold such that this function exists (i.e., such that each player's best response is always well defined).

Players observe each other's action choices in each period and have perfect recall. A pure strategy in this game for player i, s_i, is a sequence of functions $\{s_{it}(\cdot)\}_{t=1}^{\infty}$ mapping from the history of previous action choices (denoted H_{t-1}) to his action choice in period t, $s_{it}(H_{t-1}) \in S_i$. The set of all such pure strategies for player i is denoted by Σ_i, and $s = (s_1, s_2) \in \Sigma_1 \times \Sigma_2$ is a profile of pure strategies for the two players.

Any pure strategy profile $s = (s_1, s_2)$ induces an *outcome path* $Q(s)$, an infinite sequence of actions $\{q_t = (q_{1t}, q_{2t})\}_{t=1}^{\infty}$ that will actually be played when the players follow strategies s_1 and s_2. Player i's discounted payoff from outcome path Q is given by $v_i(Q) = \sum_{\tau=0}^{\infty} \delta^\tau \pi_i(q_{1+\tau})$. We also define player i's *average payoff* from outcome path Q to be $(1-\delta)v_i(Q)$; this is the per-period payoff that, if infinitely repeated, would give player i a discounted payoff of $v_i(Q)$. Finally, it is also useful to define the discounted continuation payoff from outcome path Q from some period t onward (discounted to period t) by $v_i(Q, t) = \sum_{\tau=0}^{\infty} \delta^\tau \pi_i(q_{t+\tau})$.

We can note immediately the following fact: The strategies that call for each player i to play his stage game Nash equilibrium action q_i^* in every period, regardless of the prior history of play, constitute an SPNE for *any* value of $\delta < 1$. In the discussion that follows, we are interested in determining to what extent repetition allows other outcomes to emerge as SPNEs.

Nash Reversion and the Nash Reversion Folk Theorem

We begin by considering strategies with the Nash reversion form that we considered for the Bertrand pricing game in Section 12.D.

Definition 12.AA.1: A strategy profile $s = (s_1, s_2)$ in an infinitely repeated game is one of *Nash reversion* if each player's strategy calls for playing some outcome path Q until someone defects and playing the stage game Nash equilibrium $q^* = (q_1^*, q_2^*)$ thereafter.

What outcome paths Q can be supported as outcome paths of an SPNE using Nash reversion strategies? Following logic similar to that discussed in Section 12.D, we can derive the test in Lemma 12.AA.1.

Lemma 12.AA.1: A Nash reversion strategy profile that calls for playing outcome path $Q = \{q_{1t}, q_{2t}\}_{t=1}^{\infty}$ prior to any deviation is an SPNE if and only if

$$\hat{\pi}_i(q_{jt}) + \frac{\delta}{1-\delta}\pi_i(q_1^*, q_2^*) \leq v_i(Q, t) \tag{12.AA.1}$$

(where $j \neq i$) for all t and $i = 1, 2$.

Proof: As discussed in Section 12.D, the prescribed play after any deviation is a Nash equilibrium in the continuation subgame; so we need only check whether these strategies induce a Nash equilibrium in the subgame starting in any period t when there has been no previous deviation. Note first that if for some i and t condition (12.AA.1) did not hold, then we could not have an SPNE. That is, if no deviation had occurred prior to period t, then in the continuation subgame, player i would not find following path Q to be his best response to player j's doing so (in particular, a deviation by player i in period t that maximizes his payoff in that period, followed by his playing q_i^* thereafter, would be superior for him).

In the other direction, suppose that condition (12.AA.1) is satisfied for all i and t but that we do not have an SPNE. Then there must be some period t in which some player i finds it worthwhile to deviate from outcome path Q if no previous deviation has occurred. Now, when his opponent follows a Nash revision strategy, player i's optimal deviation will involve deviating in a manner that maximizes his payoff in period t and then playing q_i^* thereafter. But his payoff from this deviation is exactly that on the left side of condition (12.AA.1), and so this deviation cannot raise his payoff. ■

Condition (12.AA.1) can be written to emphasize the trade-off between one-period gains and future losses as follows:

$$\hat{\pi}_i(q_{jt}) - \pi_i(q_{1t}, q_{2t}) \leq \delta\left(v_i(Q, t+1) - \frac{\pi_i(q_1^*, q_2^*)}{1-\delta}\right) \qquad (12.AA.2)$$

for all t and $i = 1, 2$. The left-hand side of condition (12.AA.2) gives player i's one-period gain from deviating in period t, and the right-hand side gives player i's discounted future losses from reversion to the Nash equilibrium starting in period $t + 1$.

For stationary outcome paths of the sort considered in Section 12.D [where each player i takes the same action q_i in every period, so that $Q = (q_1, q_2), (q_1, q_2), \ldots$], the infinite set of inequalities that must be checked in condition (12.AA.2) reduce to just two: infinite repetition of (q_1, q_2) is an outcome path of an SPNE that uses Nash reversion if and only if, for $i = 1$ and 2,

$$\hat{\pi}_i(q_j) - \pi_i(q_1, q_2) \leq \frac{\delta}{1-\delta}[\pi_i(q_1, q_2) - \pi_i(q_1^*, q_2^*)]. \qquad (12.AA.3)$$

How much better than the static Nash equilibrium outcome $q^* = (q_1^*, q_2^*)$ can the players do using Nash reversion? First, under relatively mild conditions (which the Bertrand game considered in Section 12.D does not satisfy), the players can sustain a stationary outcome path that has strictly higher discounted payoffs than does infinite repetition of $q^* = (q_1^*, q_2^*)$ as long as $\delta > 0$. This fact is developed formally in Proposition 12.AA.1.

Proposition 12.AA.1: Consider an infinitely repeated game with $\delta > 0$ and $S_i \subset \mathbb{R}$ for $i = 1, 2$. Suppose also that $\pi_i(q)$ is differentiable at $q^* = (q_1^*, q_2^*)$, with $\partial\pi_i(q_i^*, q_j^*)/\partial q_j \neq 0$ for $j \neq i$ and $i = 1, 2$. Then there is some $q' = (q_1', q_2')$, with $[\pi_1(q'), \pi_2(q')] \gg [\pi_1(q^*), \pi_2(q^*)]$ whose infinite repetition is the outcome path of an SPNE that uses Nash reversion.

Proof: At $q = (q_1^*, q_2^*)$, condition (12.AA.3) holds with equality. Consider a differential change in q, (dq_1, dq_2), such that $[\partial\pi_i(q_i^*, q_j^*)/\partial q_j]\, dq_j > 0$ for $i = 1, 2$. The differential change in firm i's profits from this change is

$$\begin{aligned} d\pi_i(q_i^*, q_j^*) &= \frac{\partial\pi_i(q_i^*, q_j^*)}{\partial q_i} dq_i + \frac{\partial\pi_i(q_i^*, q_j^*)}{\partial q_j} dq_j \\ &= \frac{\partial\pi_i(q_i^*, q_j^*)}{\partial q_j} dq_j, \end{aligned} \qquad (12.AA.4)$$

since q_i^* is a best response to q_j^*. Thus,

$$d\pi_i(q_i^*, q_j^*) > 0. \qquad (12.AA.5)$$

On the other hand, the envelope theorem (see Section M.L of the Mathematical Appendix) tells us that at any q_j

$$d\hat{\pi}_i(q_j) = \frac{\partial \pi_i(b_i(q_j), q_j)}{\partial q_j} dq_j,$$

where $b_i(\cdot)$ is player i's best response to q_j in the stage game. Hence,

$$d\hat{\pi}_i(q_j^*) = \frac{\partial \pi_i(q_i^*, q_j^*)}{\partial q_j} dq_j. \quad (12.AA.6)$$

Together, (12.AA.4) and (12.AA.6) imply that, to first order, the value of the left-hand side of condition (12.AA.3) is unaffected by this change. However, (12.AA.5) implies that the right-hand side of (12.AA.3), to first order, increases. Hence, for a small enough change $(\Delta q_1, \Delta q_2)$ in direction (dq_1, dq_2), infinite repetition of $(q_1 + \Delta q_1, q_2 + \Delta q_2)$ is sustainable as the outcome path of an SPNE using Nash reversion strategies and, by (12.AA.5), yields strictly higher discounted payoffs to the two players than does infinite repetition of $q^* = (q_1^*, q_2^*)$. ■

Proposition 12.AA.1 tells us that with continuous strategy sets and differentiable payoff functions, as long as there is some possibility for a joint improvement in payoffs around the stage game Nash equilibrium, some cooperation can be sustained.

Going further, examination of condition (12.AA.2) tells us that cooperation becomes easier as δ grows.

Proposition 12.AA.2: Suppose that outcome path Q can be sustained as an SPNE outcome path using Nash reversion when the discount rate is δ. Then it can be so sustained for any $\delta' \geq \delta$.

In fact, as δ gets very large, a great number of outcomes become sustainable. The result presented in Proposition 12.AA.3, a version of the *Nash reversion folk theorem* [originally due to Friedman (1971)], shows that *any* stationary outcome path that gives each player a discounted payoff that exceeds that arising from infinite repetition of the stage game Nash equilibrium $q^* = (q_1^*, q_2^*)$ can be sustained as an SPNE if δ is sufficiently close to 1.

Proposition 12.AA.3: For any pair of actions $q = (q_1, q_2)$ such that $\pi_i(q_1, q_2) > \pi_i(q_1^*, q_2^*)$ for $i = 1, 2$, there exists a $\underline{\delta} < 1$ such that, for all $\delta > \underline{\delta}$, infinite repetition of $q = (q_1, q_2)$ is the outcome path of an SPNE using Nash reversion strategies.

The proof of Proposition 12.AA.3 follows immediately from condition (12.AA.3) letting $\delta \to 1$. In fact, with a more sophisticated argument, the logic of Proposition 12.AA.3 can be extended to nonstationary outcome paths. By doing so, it is possible to convexify the set of possible payoffs identified in Proposition 12.AA.3 by alternating between various action pairs (q_1, q_2). In this way, we can support any payoffs in the shaded region of Figure 12.AA.1 as the average payoffs of an SPNE.[32]

Exercise 12.AA.1: Argue that no pair of actions q such that $\pi_i(q_1, q_2) < \pi_i(q_1^*, q_2^*)$ for some i can be sustained as a stationary SPNE outcome path using Nash reversion.

More Severe Punishments and the Folk Theorem

It is intuitively clear that, for a given level of $\delta < 1$, the more severe the punishments that can be credibly threatened in response to a deviation, the easier it is to prevent

32. See Fudenberg and Maskin (1991) for details.

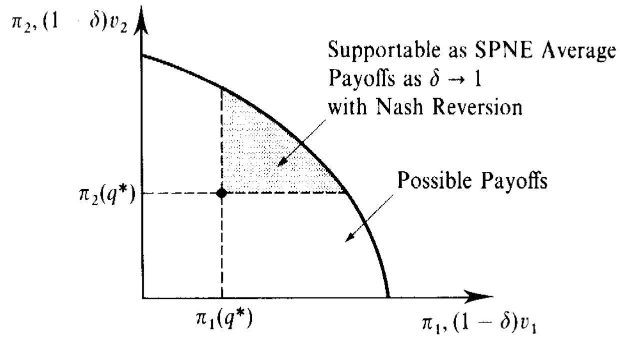


Figure 12.AA.1
The Nash reversion folk theorem.

players from deviating from any given outcome path. In general, Nash reversion is not the most severe credible punishment that is possible. Just as players can be induced to cooperate through the use of threatened punishments, they can also be induced to punish each other.

To consider this issue, it is useful to let $\underline{\pi}_i = \text{Min}_{q_j}[\text{Max}_{q_i} \pi_i(q_i, q_j)]$ denote player i's *minimax payoff*.[33] Payoff $\underline{\pi}_i$ is the lowest payoff that player i's rival can hold him to in the stage game if player i anticipates the action that his rival will play. Note, first, that player i's payoff in the stage game Nash equilibrium $q^* = (q_1^*, q_2^*)$ cannot be below $\underline{\pi}_i$. More importantly, regardless of the strategies played by his rival, player i's average payoff in the infinitely repeated game or in any subgame within it cannot be below $\underline{\pi}_i$. Thus, no punishment following a deviation can give player i an average payoff below $\underline{\pi}_i$. Payoffs that strictly exceed $\underline{\pi}_i$ for each player i are known as *individually rational payoffs*.

Note that for a punishment to be credible we must be sure that after an initial deviation occurs and the punishment is called for, no player wants to deviate from the prescribed punishment path. This means that a punishment is credible if and only if it itself constitutes an SPNE outcome path. Proposition 12.AA.4 tells us that as long as $\delta > 0$ and conditions similar to those in Proposition 12.AA.1 hold, SPNEs that yield more severe punishments than Nash reversion can be constructed whenever each player i's stage game Nash equilibrium payoff strictly exceeds $\underline{\pi}_i$. (You are asked to prove this result in Exercise 12.AA.2.)

Proposition 12.AA.4: Consider an infinitely repeated game with $\delta > 0$ and $S_i \subset \mathbb{R}$ for $i = 1, 2$. Suppose also that $\pi_i(q)$ is differentiable at $q^* = (q_1^*, q_2^*)$, with $\partial \pi_i(q_i^*, q_j^*)/\partial q_j \neq 0$ for $j \neq i$ and $i = 1, 2$, and that $\pi_i(q_1^*, q_2^*) > \underline{\pi}_i$ for $i = 1, 2$. Then there is some SPNE with discounted payoffs to the two players of (v_1', v_2') such that $(1 - \delta)v_i' < \pi_i(q_1^*, q_2^*)$ for $i = 1, 2$.

Under the conditions of Proposition 12.AA.4, for any $\delta \in (0, 1)$, more severe punishments than Nash reversion can credibly be threatened. We should therefore expect that more cooperative outcomes can be sustained than those sustainable through the threat of Nash reversion whenever a fully cooperative outcome is not already achievable using Nash reversion strategies.

33. In general, a player's minimax payoff will be lower if mixed strategies are allowed. In this case, the statement of the folk theorem given in Proposition 12.AA.5 remains unchanged, but with these (potentially) lower levels of $\underline{\pi}_i$.

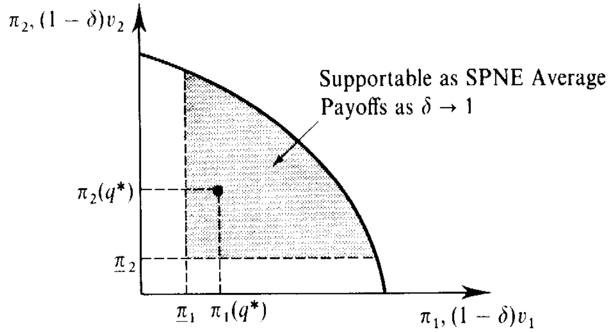


Figure 12.AA.2
The folk theorem.

For arbitrary $\delta < 1$, constructing the full set of SPNEs is a delicate process. Each SPNE, whether collusive or punishing, uses other SPNEs as threatened punishments. For details on how this is done, see the original contributions by Abreu (1986) and (1988) and the presentation in Fudenberg and Tirole (1992). As with SPNEs using Nash reversion strategies, the full set of SPNEs grows as δ increases, making possible both more cooperation and more severe punishments. In fact, the result presented in Proposition 12.AA.5, known as the *folk theorem*, tells us that *any* feasible individually rational payoffs can be supported as the average payoffs in an SPNE as long as players discount the future to a sufficiently small degree.[34] (Feasibility simply means that there is some outcome path Q that generates these average payoffs.)

Proposition 12.AA.5: (*The Folk Theorem*) For any feasible pair of individually rational payoffs $(\pi_1, \pi_2) \gg (\underline{\pi}_1, \underline{\pi}_2)$, there exists a $\underline{\delta} < 1$ such that, for all $\delta > \underline{\delta}$, (π_1, π_2) are the average payoffs arising in an SPNE.

In comparison with Proposition 12.AA.3, Proposition 12.AA.5 tells us that as $\delta \to 1$ we can support any average payoffs that exceed each player's minimax payoff.[35] This limiting set of SPNE average payoffs is shown in Figure 12.AA.2.

Example 12.AA.1 gives some idea of how this can be done.

Example 12.AA.1: *Sustaining an Average Payoff of Zero in the Infinitely Repeated Cournot Game.* In this example, we construct an SPNE in which both firms earn an average payoff of zero in an infinitely repeated Cournot game. In particular, let the stage game be a symmetric Cournot duopoly game with cost function $c(q) = cq$, where $c > 0$, and a continuous inverse demand function $p(\cdot)$ such that $p(x) \to 0$ as $x \to \infty$. It will be convenient to write a firm's profit when both firms choose quantity q as $\pi(q) = [p(2q) - c]q$ and, as before, a firm's best-response profits when its rival

34. The theorem's name refers to the fact that some version of the result was known in game theory "folk wisdom" well before its formal appearance in the literature. See Fudenberg and Maskin (1986) and (1991) for a proof of the result. When there are more than two players, the result requires that the set of feasible payoffs satisfy an additional "dimensionality" condition. The original appearances of the result in the literature actually analyzed infinitely repeated games *without* discounting [see, for example, Rubinstein (1979)].

35. We may also be able in some cases to give each player exactly his minimax payoff. This is the case, for example, in the repeated Bertrand game, where the stage game's Nash equilibrium yields the minimax payoffs. In Example 12.AA.1, we show that we can also do this for large enough δ in the repeated Cournot duopoly game.

chooses quantity q as $\hat{\pi}(q)$.[36] Note that $\underline{\pi}_j = 0$ for $j = 1, 2$ here; if firm j's rival chooses a quantity at least as large as the competitive quantity q_c satisfying $p(q_c) = c$, then the best firm j can do is to produce nothing and earn zero, and firm j can never be forced to a payoff worse than zero.

Consider strategies for the players that take the following form:

(i) Both firms play quantity $\tilde{q}$ in period 1 followed by the monopoly quantity q^m in every period $t > 1$ as long as no one deviates, where quantity $\tilde{q}$ satisfies

$$\pi(\tilde{q}) + \frac{\delta}{1-\delta}\pi(q^m) = 0. \tag{12.AA.7}$$

(ii) If anyone deviates when $\tilde{q}$ is meant to be played, the outcome path described in (i) is restarted.

(iii) If anyone deviates when q^m is meant to be played, Nash reversion occurs.

Note that the outcome path described in (i), if followed by both players, gives both players an average payoff of zero by construction [recall (12.AA.7)].

By Proposition 12.AA.3, we know that for some $\underline{\delta} < 1$ we can sustain infinite repetition of q^m through Nash reversion for all $\delta > \underline{\delta}$. Thus, for $\delta > \underline{\delta}$, neither firm will deviate from the above strategies when q^m is supposed to be played. Will they deviate when $\tilde{q}$ is supposed to be played? Consider firm j's payoff from deviating from $\tilde{q}$ in a single period and conforming with the prescribed strategy thereafter. Firm j earns $\hat{\pi}(\tilde{q}) + (\delta)(0)$ because it plays a best response when deviating, and then the original path is restarted. Thus, this deviation does not improve firm j's payoff if $\hat{\pi}(\tilde{q}) = 0$ (it cannot be less than zero because $\underline{\pi}_i = 0$). This is so if $\tilde{q} \geq q_c$. But examining condition (12.AA.7), we see that as δ approaches 1, $\pi(\tilde{q})$ must get increasingly negative for (12.AA.7) to hold and, in particular, that there exists a $\delta_c < 1$ such that $\tilde{q}$ will exceed q_c for all $\delta > \delta_c$. Thus, for $\delta > \text{Max}\{\delta_c, \underline{\delta}\}$, these strategies constitute an SPNE that gives both firms an average payoff of 0.[37] ■

APPENDIX B: STRATEGIC ENTRY DETERRENCE AND ACCOMMODATION

In this appendix, we discuss an important example of credible precommitments to affect future market conditions in which an incumbent firm engages in pre-entry capacity expansion to gain a strategic advantage over a potential entrant and possibly to deter this firm's entry altogether [the original analyses of this issue are due to Spence (1977) and Dixit (1980)]. In what follows, we study the following three-stage game that is adapted from Dixit (1980).

36. We can make the strategy sets compact by noting that in no period will any firm ever choose a quantity larger than the level $\bar{q}$ such that $\pi(\bar{q}) + [\delta/(1-\delta)](\text{Max}_q\, \pi(q)) = 0$, because it would do better setting its quantity equal to zero forever. Then, without loss, we can let each firm choose its output from the compact set $[0, \bar{q}]$.

37. We have not considered any multiperiod deviations, but it can be shown that if no single-period deviation followed by conformity with the strategies is worthwhile, then neither is any multiperiod deviation (this is a general principle of dynamic programming).

Stage 1: An incumbent, firm I, chooses the capacity level of its plant, denoted by k_I. Capacity costs r per unit.

Stage 2: A potential entrant, firm E, decides whether to enter the market. If it does, it pays an entry cost of F.

Stage 3: If firm E enters, the two firms choose their output levels, q_I and q_E, simultaneously. The resulting price is $p(q_I + q_E)$. For firm E, output costs $(w + r)$ per unit: for each unit of output produced, firm E incurs both a capacity cost of r and a labor cost of w. For firm I, production must not exceed its previously chosen capacity level. Its production cost, however, is only w per unit because it has already built its capacity. If, on the other hand, firm E does not enter, then firm I acts as a monopolist who can produce up to k_I units of output at cost w per unit.

To determine the subgame perfect Nash equilibrium (SPNE) of this game, we begin by analyzing behavior in the stage 3 subgames and then work backward.

Stage 3: Quantity Competition

The subgames in stage 3 are distinguished by two previous events: whether firm E has entered and the previous capacity choice of firm I. We first consider the outcome of stage 3 competition following entry and then discuss firm I's behavior in stage 3 if entry does not occur. For simplicity, we assume throughout that firms' profit functions are strictly concave in own quantity; a sufficient condition for this is for $p(\cdot)$ to be concave. The concavity of $p(\cdot)$ also implies that firms' best-response functions are downward sloping.

Stage 3 competition after entry. Figure 12.BB.1 depicts firm E's best-response function in stage 3, which we denote by $b(q|w + r)$ to emphasize that it is the best-response function for a firm with marginal cost $w + r$. Firm E's stage 3 profits decline as we move along this curve to the right (involving higher levels of q_I) and, at some point, denoted Z in the figure, they fall below the entry cost F.

Now consider firm I's optimal behavior. The key difference between firm I and firm E is that firm I has already built its capacity. Hence, firm I's expenditure on this capacity is sunk (it cannot recover it by reducing its capacity), its capacity level is fixed, and its marginal cost is only w. Suppose we let $b(q|w)$ denote the best-response function of a firm with marginal cost w. Then firm I's best-response function in stage 3 is

$$b_I(q_E|k_I) = \text{Min}\{b(q_E|w), k_I\}.$$

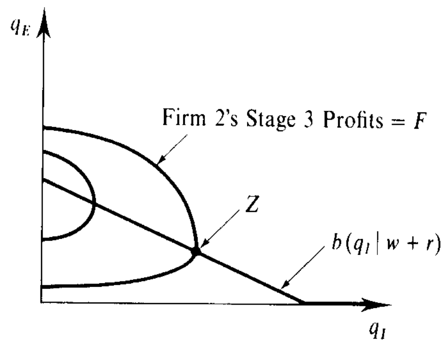


Figure 12.BB.1
Firm E's stage 3 best-response function after entry.

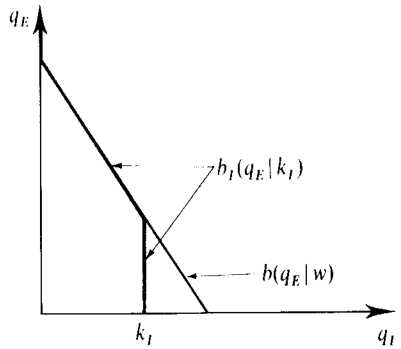


Figure 12.BB.2 (left)
Firm I's stage 3 best-response function after entry.

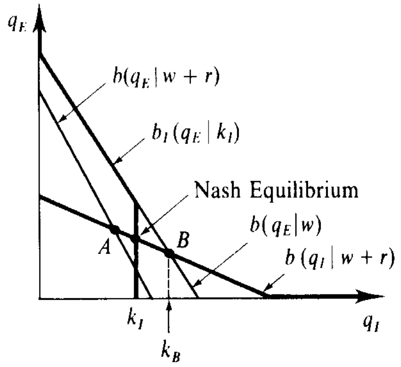


Figure 12.BB.3 (right)
Stage 3 Nash equilibrium after entry.

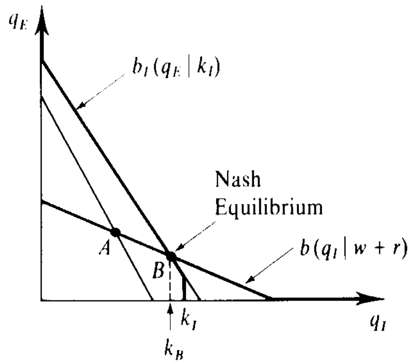


Figure 12.BB.4
A stage 3 equilibrium in which firm I does not use all of its capacity.

That is, firm I's best response to an output choice of q_E by firm E is the same as that for a firm with marginal cost level w as long as this output level does not exceed its previously chosen capacity. Figure 12.BB.2 illustrates firm I's best-response function.

We can now put together the best-response functions for the two firms to determine the equilibrium in stage 3 following firm E's decision to enter, for any given level of k_I. This equilibrium is shown in Figure 12.BB.3.

In Figure 12.BB.3, point A is the outcome that would arise if there were no first-mover advantage for firm I, that is, if the two firms chose both their capacity and output levels simultaneously. However, when firm I is able to choose its capacity level first, by choosing an appropriate level of k_I, it can get the post-entry equilibrium to lie anywhere on firm E's best-response function up to point B. Firm I is able to induce points to the right of point A because its ability to incur its capacity costs prior to stage 3 competition allows it to have a marginal cost in stage 3 of only w, rather than $w + r$. Note, however, that firm I cannot induce a point on firm 2's best-response function beyond point B, even though it might want to; if it built a capacity greater than level k_B, it would not have an incentive to actually use all of it. Figure 12.BB.4 depicts this situation. A threat to produce up to capacity following entry would in this case not be credible.

Stage 3 outcomes if firm E does not enter. If firm E decides not to enter, then firm I will be a monopolist in stage 3. Its optimal monopoly output is then the point where its best-response function hits the $q_E = 0$ axis, $b_I(0|k_I)$.

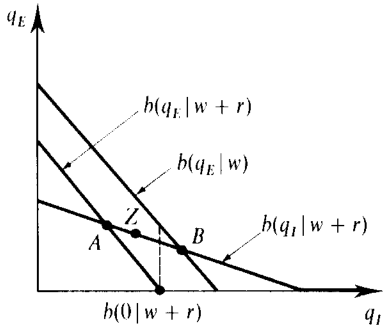


Figure 12.BB.5 (left)
Blockaded entry.

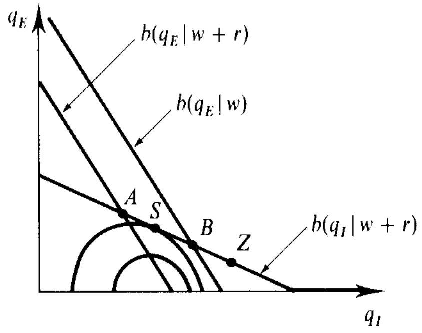


Figure 12.BB.6 (right)
Strategic entry accommodation when entry is inevitable.

Stage 2: Firm E's Entry Decision

Firm E's entry decision is straightforward: Given the level of capacity k_I chosen by firm I in stage 1, firm E will enter if it expects nonnegative profits net of its entry cost F. This means that firm E will enter when it expects that the post-entry equilibrium will lie to the left of point Z on its best-response function in Figure 12.BB.1.

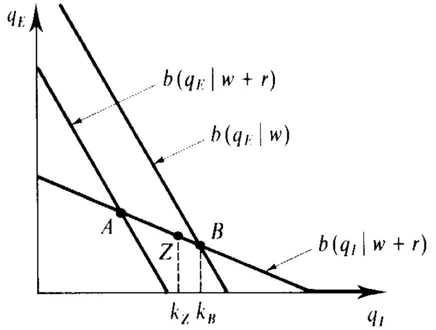
Stage 1: Firm I's Stage 1 Capacity Investment

Now consider firm I's optimal capacity choice in stage 1. There are three situations in which firm I could find itself: Entry could be blockaded, entry could be inevitable, or entry deterrence could be possible but not inevitable. Let us consider each in turn.

Entry is blockaded. One possibility is that the entry cost F is large enough that firm E does not find it worthwhile to enter even if firm I ignores the possibility of entry and simply builds the same capacity that it would if it were an uncontested monopolist, $b(0 \mid w + r)$. This situation, in which we say that *entry is blockaded,* is shown in Figure 12.BB.5. In this case, firm I achieves its best possible outcome: it builds a capacity of $b(0 \mid w + r)$, no entry occurs, and then it sells $b(0 \mid w + r)$ units of output.

Entry deterrence is impossible: strategic entry accommodation. Suppose that point Z is to the right of point B. In this case, entry deterrence is impossible; firm E will find it profitable to enter regardless of k_I. What is firm I's optimal choice of k_I in this case? In Figure 12.BB.6, we have drawn isoprofit curves for firm I; note that because these include the cost of capacity, they are the isoprofit curves corresponding to those of a firm with marginal cost $(w + r)$. Now recall that firm I can induce any point on firm E's best-response function up to point B through an appropriate choice of capacity. It will choose the point that maximizes its profit. In Figure 12.BB.6, this point, which involves a tangency between firm E's best-response function and firm I's isoprofit curves, is denoted as point S. This outcome corresponds to exactly the outcome that would emerge in a model of sequential quantity choice, known as a *Stackleberg leadership model* (see Exercise 12.C.18). Note that firm I's first-mover advantage allows it to earn higher profits than the otherwise identical firm E.

The point of tangency, S, could also lie to the right of point B. In this case, the optimal capacity choice will be $k_I = k_B$, and the outcome will not be as desirable for firm I as the Stackleberg point. Here firm I is unable to credibly



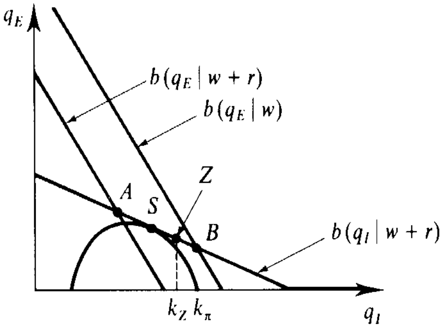


Figure 12.BB.7 (left)
Entry deterrence is possible but not inevitable.

Figure 12.BB.8 (right)
Entry deterrence versus entry accommodation.

commit to produce the output associated with point S, even if it builds sufficient capacity in stage 1.

Entry deterrence is possible but not inevitable. Suppose now that point Z lies to the left of point B but not so far that entry is blockaded, as shown in Figure 12.BB.7. Firm I can deter firm E's entry by picking a capacity level at least as large as point k_Z in the figure. The only question is whether this will be optimal for firm I, or whether firm I is better off accommodating firm E's entry. To judge this, firm I will compare its profits at point $(k_Z, 0)$ to those at point S (or at point B if point S lies to the right of B). This can be done by comparing the capacity level k_π in Figure 12.BB.8, the output level under monopoly that gives the same profit as the optimal accommodation point S, with k_Z. If $k_\pi > k_Z$, then firm I prefers to deter entry because its profits are higher in this case; but if $k_\pi < k_Z$, then it will prefer accommodation. Note that if deterrence is optimal, then even though entry does not occur its *threat* nevertheless has an effect on the market outcome, raising the level of output and welfare relative to a situation in which no entry is possible.

Exercise 12.BB.1: Show that when entry deterrence is possible but not inevitable, if point S lies to the right of point Z, then entry deterrence is better than entry accommodation.

REFERENCES

Abreu, D. (1986). Extremal equilibria of oligopolistic supergames. *Journal of Economic Theory* **39**: 191–225.

Abreu, D. (1988). On the theory of infinitely repeated games with discounting. *Econometrica* **56**: 383–96.

Abreu, D., D. Pearce, and E. Stachetti. (1990). Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica* **58**: 1041–64.

Baumol, W., J. Panzar, and R. Willig. (1982). *Contestable Markets and the Theory of Industry Structure.* San Diego: Harcourt, Brace, Jovanovich.

Bertrand, J. (1883). Théorie mathématique de la richesse sociale. *Journal des Savants* **67**: 499–508.

Bulow, J., J. Geanakoplos, and P. Klemperer. (1985). Multimarket oligopoly: strategic substitutes and complements. *Journal of Political Economy* **93**: 488–511.

Chamberlin, E. (1933). *The Theory of Monopolistic Competition.* Cambridge, Mass.: Harvard University Press.

Cournot, A. (1838). *Recherches sur les Principes Mathématiques de la Théorie des Richesses.* [English edition: *Researches into the Mathematical Principles of the Theory of Wealth,* edited by N. Bacon. London: Macmillan, 1897.]

Dixit, A. (1980). The role of investment in entry deterrence. *Economic Journal* **90**: 95–106.

Dixit, A., and J. E. Stiglitz. (1977). Monopolistic competition and optimal product diversity. *American Economic Review* **67**: 297–308.

Edgeworth, F. (1897). Me teoria pura del monopolio. *Giornale degli Economisti* **40**: 13 31. [English translation: The pure theory of monopoly. In *Papers Relating to Political Economy,* Vol. I, edited by F. Edgeworth. London: Macmillan, 1925.]

Friedman, J. (1971). A non-cooperative equilibrium for supergames. *Review of Economic Studies* **28**: 1–12.

Fudenberg, D., and E. Maskin. (1986). The folk theorem in repeated games with discounting or with incomplete information. *Econometrica* **52**: 533–54.

Fudenberg, D., and E. Maskin. (1991). On the dispensibility of public randomizaton in discounted repeated games. *Journal of Economic Theory* **53**: 428–38.

Fudenberg, D., and J. Tirole. (1984). The fat cat effect, the puppy dog ploy, and the lean and hungry look. *American Economic Review, Papers and Proceedings* **74**: 361–68.

Fudenberg, D., and J. Tirole. (1992). *Game Theory.* Cambridge, Mass.: MIT Press.

Green, E., and R. Porter. (1984). Noncooperative collusion under imperfect price information. *Econometrica* **52**: 87–100.

Hart, O. D. (1985). Monopolistic competition in the spirit of Chamberlin: A general model. *Review of Economic Studies* **52**: 529–46.

Kreps, D. M., and J. Scheinkman. (1983). Quantity precommitment and Bertrand competition yield Cournot outcomes. *Rand Journal of Economics* **14**: 326–37.

Mankiw, N. G., and M. D. Whinston. (1986). Free entry and social inefficiency. *Rand Journal of Economics* **17**: 48 58.

Osborne, M. J., and A. Rubinstein. (1994). *A Course in Game Theory.* Cambridge, Mass.: MIT Press.

Rotemberg, J., and G. Saloner. (1986). A supergame-theoretic model of business cycles and price wars during booms. *American Economic Review* **76**: 390–407.

Rubinstein, A. (1979). Equilibrium in supergames with the overtaking criterion. *Journal of Economic Theory* **21**: 1 9.

Salop, S. (1979). Monopolistic competition with outside goods. *Bell Journal of Economics* **10**: 141–56.

Shapiro, C. (1989). Theories of oligopoly behavior. In *Handbook of Industrial Organization,* edited by R. Schmalensee and R. D. Willig. Amsterdam: North-Holland.

Spence, A. M. (1976). Product selection, fixed costs, and monopolistic competition. *Review of Economic Studies* **43**: 217 35.

Spence, A. M. (1977). Entry, capacity investment, and oligopolistic pricing. *Bell Journal of Economics* **8**: 534 44.

Stigler, G. (1960). A theory of oligopoly. *Journal of Political Economy* **72**: 44–61.

Tirole, J. (1988). *The Theory of Industrial Organization.* Cambridge, Mass.: MIT Press.

EXERCISES

12.B.1[A] The expression $[p^m - c'(q^m)]/p^m$, where p^m and q^m are the monopolist's price and output level, respectively, is known as the monopolist's *price–cost margin* (or as the *Lerner index of monopoly power*). It measures the distortion of the monopolist's price above its marginal cost as a proportion of its price.

(a) Show the monopolist's price–cost margin is always equal to the inverse of the price elasticity of demand at price p^m.

(b) Also argue that if the monopolist's marginal cost is positive at every output level, then demand must be *elastic* (i.e., the price elasticity of demand is greater than 1) at the monopolist's optimal price.

12.B.2^B Consider a monopolist with cost function $c(q) = cq$, with $c > 0$, facing demand function $x(p) = \alpha p^{-\varepsilon}$, where $\varepsilon > 0$.

(a) Show that if $\varepsilon \leq 1$, then the monopolist's optimal price is not well defined.

(b) Assume that $\varepsilon > 1$. Derive the monopolist's optimal price, quantity, and price–cost margin $(p^m - c)/p^m$. Calculate the resulting deadweight welfare loss.

(c) (Harder) Consider a sequence of demand functions that differ in their levels of ε and α but that all involve the same competitive quantity $x(c)$ [i.e., for each level of ε, α is adjusted to keep $x(c)$ the same]. How does the deadweight loss vary with ε? (If you cannot derive an analytic answer, try calculating some values on a computer.)

12.B.3^B Suppose that we consider a monopolist facing demand function $x(p, \theta)$ with cost function $c(q, \phi)$, where θ and ϕ are parameters. Use the implicit function theorem to compute the changes in the monopolist's price and quantity as a function of a differential change in either θ or ϕ. When will each lead to a price increase?

12.B.4^B Consider a monopolist with a cost of c per unit. Use a "revealed preference" proof to show that the monopoly price is nondecreasing in c. Then extend your argument to the case in which the monopolist's cost function is $c(q, \phi)$, with $[c(q'', \phi) - c(q', \phi)]$ increasing in ϕ for all $q'' > q'$, by showing that the monopoly price is nondecreasing in ϕ. (If you did Exercise 12.B.3, also relate this condition to the one you derived there.)

12.B.5^B Suppose that a monopolist faces many consumers. Argue that in each of the following two cases, the monopolist can do no better than it does by restricting itself to simply charging a price per unit, say p.

(a) Suppose that each consumer i wants either one or no units of the monopolist's good and that the monopolist is unable to discern any particular consumer's preferences.

(b) Suppose that consumers may desire to consume multiple units of the good. The monopolist cannot discern any particular consumer's preferences. In addition, resale of the good is costless and after the monopolist has made its sales to consumers a competitive market develops among consumers for the good.

12.B.6^A Suppose that the government can tax or subsidize a monopolist who faces inverse demand function $p(q)$ and has cost function $c(q)$ [assume both are differentiable and that $p(q)q - c(q)$ is concave in q]. What tax or subsidy per unit of output would lead the monopolist to act efficiently?

12.B.7^B Consider the widget market. The total demand by men for widgets is given by $x_m(p) = a - \theta_m p$, and the total demand by women is given by $x_w(p) = a - \theta_w p$, where $\theta_w < \theta_m$. The cost of production is c per widget.

(a) Suppose the widget market is competitive. Find the equilibrium price and quantity sold.

(b) Suppose, instead, that firm A is a monopolist of widgets [also make this assumption in **(c)** and **(d)**]. If firm A is prohibited from "discriminating" (i.e., charging different prices to men and women), what is its profit-maximizing price? Under what conditions do both men and women consume a positive level of widgets in this solution?

(c) If firm A has produced some total level of output X, what is the welfare-maximizing way to distribute it between the men and the women? (Assume here and below that Marshallian aggregrate surplus is a valid measure of welfare.)

(d) Suppose that firm A is allowed to discriminate. What prices does it charge? In the case where the nondiscriminatory solution in **(b)** has positive consumption of widgets by both men and women, does aggregate welfare as measured by the Marshallian aggregate surplus rise or fall relative to when discrimination is allowed? Relate your conclusion to your answer in **(c)**. What if the nondiscriminatory solution in **(b)** has only one type of consumers being served?

12.B.8[B] Consider the following two-period model: A firm is a monopolist in a market with an inverse demand function (in each period) of $p(q) = a - bq$. The cost per unit in period 1 is c_1. In period 2, however, the monopolist has "learned by doing," and so its constant cost per unit of output is $c_2 = c_1 - mq_1$, where q_1 is the monopolist's period 1 output level. Assume $a > c$ and $b > m$. Also assume that the monopolist does not discount future earnings.

(a) What is the monopolist's level of output in each of the periods?

(b) What outcome would be implemented by a benevolent social planner who fully controlled the monopolist? Is there any sense in which the planner's period 1 output is selected so that "price equals marginal cost"?

(c) Given that the monopolist will be selecting the period 2 output level, would the planner like the monopolist to slightly increase the level of period 1 output above that identified in **(a)**? Can you give any intuition for this?

12.B.9[C] Consider a situation in which there is a monopolist in a market with inverse demand function $p(q)$. The monopolist makes two choices: How much to invest in cost reduction, I, and how much to sell, q. If the monopolist invests I in cost reduction, his (constant) per-unit cost of production is $c(I)$. Assume that $c'(I) < 0$ and that $c''(I) > 0$. Assume throughout that the monopolist's objective function is concave in q and I.

(a) Derive the first-order conditions for the monopolist's choices.

(b) Compare the monopolist's choices with those of a benevolent social planner who can control both q and I (a "first-best" comparison).

(c) Compare the monopolist's choices with those of a benevolent social planner who can control I but not q (a "second-best" comparison). Suppose that the planner chooses I and then the monopolist chooses q.

12.B.10[B] Consider a monopolist that can choose both its product's price p and its quality q. The demand for its product is given by $x(p, q)$, which is increasing in q and decreasing in p. Given the price chosen by the monopolist, does the monopolist choose the socially efficient quality level?

12.C.1[A] In text.

12.C.2[C] Extend the argument of Proposition 12.C.1 to show that under the assumptions made in the text [in particular, the assumption that there is a price $\bar{p} < \infty$ such that $x(p) = 0$ for all $p \geq \bar{p}$], both firms setting their price equal to c with certainty is the unique Nash equilibrium of the Bertrand duopoly model even when we allow for mixed strategies.

12.C.3[B] Note that the unique Nash equilibrium of the Bertrand duopoly model has each firm playing a weakly dominated strategy. Consider an alteration of the model in which prices must be named in some discrete unit of account (e.g., pennies) of size Δ.

(a) Show that both firms naming prices equal to the smallest multiple of Δ that is strictly greater than c is a pure strategy equilibrium of this game. Argue that it does not involve either firm playing a weakly dominated strategy.

(b) Argue that as $\Delta \to 0$, this equilibrium converges to both firms charging prices equal to c.

12.C.4[B] Consider altering the Bertrand duopoly model to a case in which each firm j's cost per unit is c_j and $c_1 < c_2$.

(a) What are the pure strategy Nash equilibria of this game?

(b) Examine a model in which prices must be named in discrete units, as in Exercise 12.C.3. What are the pure strategy Nash equilibria of such a game? Which do not involve the play of weakly dominated strategies? As the grid becomes finer, what is the limit of these equilibria in undominated strategies?

12.C.5^{B} Suppose that we have a market with I buyers, each of whom wants at most one unit of the good. Buyer i is willing to pay up to v_i for his unit, and $v_1 > v_2 > \cdots > v_I$. There are a total of $q < I$ units available. Suppose that buyers simultaneously submit bids for a unit of the output and that the output goes to the q highest bidders, who pay the amounts of their bids. Show that every buyer making a bid of v_{q+1} and the good being assigned to buyers $1, \ldots, q$ is a Nash equilibrium of this game. Argue that this is a competitive equilibrium price. Also show that in *any* pure strategy Nash equilibrium of this game, buyers 1 through q receive a unit and buyers $q + 1$ through I do not.

12.C.6^{A} In text.

12.C.7^{B} In text.

12.C.8^{C} Consider a homogeneous-good J-firm Cournot model in which the demand function $x(p)$ is downward sloping but otherwise arbitrary. The firms all have an identical cost function $c(q)$ that is increasing in q and convex. Denote by Q the aggregate output of the J firms, and let $Q_{-j} = \sum_{k \neq j} q_k$.

(a) Show that firm j's best response can be written as $b(Q_{-j})$.

(b) Show that $b(Q_{-j})$ need not be unique (i.e., that it is in general a correspondence, not a function).

(c) Show that if $\hat{Q}_{-j} > Q_{-j}$, $q_j \in b(Q_{-j})$, and $\hat{q}_j \in b(\hat{Q}_{-j})$, then $(\hat{q}_j + \hat{Q}_{-j}) \geq (q_j + Q_{-j})$. Deduce from this that $b(\cdot)$ can jump only upward and that $b'(Q_{-j}) \geq -1$ whenever this derivative is defined.

(d) Use you result in **(c)** to prove that a symmetric pure strategy Nash equilibrium exists in this model.

(e) Show that multiple equilibria are possible.

(f) Give sufficient conditions (they are very weak) for the symmetric equilibrium to be the only equilibrium in pure strategies.

12.C.9^{B} Consider a two-firm Cournot model with constant returns to scale but in which firms' costs may differ. Let c_j denote firm j's cost per unit of output produced, and assume that $c_1 > c_2$. Assume also that the inverse demand function is $p(q) = a - bq$, with $a > c_1$.

(a) Derive the Nash equilibrium of this model. Under what conditions does it involve only one firm producing? Which will this be?

(b) When the equilibrium involves both firms producing, how do equilibrium outputs and profits vary when firm 1's cost changes?

(c) Now consider the general case of J firms. Show that the ratio of industry profits divided by industry revenue in any (pure strategy) Nash equilibrium is exactly H/ε, where ε is the elasticity of the market demand curve at the equilibrium price and H, the *Herfindahl index of concentration*, is equal to the sum of the firms' squared market shares $\sum_j (q_j^*/Q^*)^2$. (*Note:* This result depends on the assumption of constant returns to scale.)

12.C.10^{B} Consider a J-firm Cournot model in which firms' costs differ. Let $c_j(q_j) = \alpha_j \tilde{c}(q_j)$ denote firm j's cost function, and assume that $\tilde{c}(\cdot)$ is strictly increasing and convex. Assume that $\alpha_1 > \cdots > \alpha_J$.

(a) Show that if more than one firm is making positive sales in a Nash equilibrium of this model, then we cannot have productive efficiency; that is, the equilibrium aggregate output Q^* is produced inefficiently.

(b) If so, what is the correct measure of welfare loss relative to a fully efficient (competitive) outcome? [*Hint:* Reconsider the discussion in Section 10.E.]

(c) Provide an example in which welfare decreases when a firm becomes more productive (i.e., when α_j falls for some j). [*Hint:* Consider an improvement in cost for firm 1 in the model of Exercise 12.C.9.] Why can this happen?

12.C.11^C Consider a capacity-constrained duopoly pricing game. Firm j's capacity is q_j for $j = 1, 2$, and it has a constant cost per unit of output of $c \geq 0$ up to this capacity limit. Assume that the market demand function $x(p)$ is continuous and strictly decreasing at all p such that $x(p) > 0$ and that there exists a price $\tilde{p}$ such that $x(\tilde{p}) = q_1 + q_2$. Suppose also that $x(p)$ is concave. Let $p(\cdot) = x^{-1}(\cdot)$ denote the inverse demand function.

Given a pair of prices charged, sales are determined as follows: consumers try to buy at the low-priced firm first. If demand exceeds this firm's capacity, consumers are served in order of their valuations, starting with high-valuation consumers. If prices are the same, demand is split evenly unless one firm's demand exceeds its capacity, in which case the extra demand spills over to the other firm. Formally, the firms' sales are given by the functions $x_1(p_1, p_2)$ and $x_2(p_1, p_2)$ satisfying [$x_i(\cdot)$ gives the amount firm i sells taking account of its capacity limitation in fulfilling demand]

$$\text{If } p_j > p_i: \quad x_i(p_1, p_2) = \text{Min}\{q_i, x(p_i)\}$$

$$x_j(p_1, p_2) = \text{Min}\{q_j, \text{Max}\{x(p_j) - q_i, 0\}\}$$

$$\text{If } p_2 = p_1 = p: \quad x_i(p_1, p_2) = \text{Min}\{q_i, \text{Max}\{x(p)/2, x(p) - q_j\}\} \quad \text{for } i = 1, 2.$$

(a) Suppose that $q_1 < b_c(q_2)$ and $q_2 < b_c(q_1)$, where $b_c(\cdot)$ is the best-response function for a firm with constant marginal costs of c. Show that $p_1^* = p_2^* = p(q_1 + q_2)$ is a Nash equilibrium of this game.

(b) Argue that if either $q_1 > b_c(q_2)$ or $q_2 > b_c(q_1)$, then no pure strategy Nash equilibrium exists.

12.C.12^B Consider two strictly concave and differentiable profit functions $\pi_j(q_j, q_k)$, $j = 1, 2$, defined on $q_j \in [0, q]$.

(a) Give sufficient conditions for the best-response functions $b_j(q_j)$ to be increasing or decreasing.

(b) Specialize to the Cournot model. Argue that a decreasing (downward-sloping) best-response function is the "normal" case.

12.C.13^B Show that when $v > c + 3t$ in the linear city model discussed in Example 12.C.2, a firm j's best response to any price of its rival p_{-j} always results in all consumers purchasing from one of the two firms.

12.C.14^C Consider the linear city model discussed in Example 12.C.2.

(a) Derive the best-response functions when $v \in (c + 2t, c + 3t)$. Show that the unique Nash equilibrium in this case is $p_1^* = p_2^* = c + t$.

(b) Repeat **(a)** for the case in which $v \in (c + \frac{3}{2}t, c + 2t)$.

(c) Show that when $v < c + t$, the unique Nash equilibrium involves prices of $p_1^* = p_2^* = (v + c)/2$ and some consumers not purchasing from either firm.

(d) Show that when $v \in (c + t, c + \frac{3}{2}t)$, the unique symmetric equilibrium is $p_1^* = p_2^* = v - t/2$. Are there asymmetric equilibria in this case?

(e) Compare the change in equilibrium prices and profits from a reduction in t in the case studied in **(d)** with that in the equilibria of **(a)** and **(b)**.

12.C.15^B Derive the Nash equilibrium prices of the linear city model where a consumer's travel cost is quadratic in distance, that is, where the total cost of purchasing from firm j is $p_j + td^2$, where d is the consumer's distance from firm j. Restrict attention to the case in which v is large enough that the possibility of nonpurchase can be ignored.

12.C.16^B Derive the Nash equilibrium prices and profits in the circular city model with J firms when travel costs are quadratic, as in Exercise 12.C.15. Restrict attention to the case in which v is large enough that the possibility of nonpurchase can be ignored. What happens as J grows large? As t falls?

12.C.17^B Consider the linear city model in which the two firms may have different constant unit production costs $c_1 > 0$ and $c_2 > 0$. Without loss of generality, take $c_1 \le c_2$ and suppose that v is large enough that nonpurchase can be ignored. Determine the Nash equilibrium prices and sales levels for equilibria in which both firms make strictly positive sales. How do local changes in c_1 affect the equilibrium prices and profits of firms 1 and 2? For what values of c_1 and c_2 does the equilibrium involve one firm making no sales?

12.C.18^B (*The Stackleberg leadership model*) There are two firms in a market. Firm 1 is the "leader" and picks its quantity first. Firm 2, the "follower," observes firm 1's choice and then chooses its quantity. Profits for each firm i given quantity choices q_1 and q_2 are $p(q_1 + q_2)q_i - cq_i$, where $p'(q) < 0$ and $p'(q) + p''(q)q < 0$ at all $q \ge 0$.

(a) Prove formally that firm 1's quantity choice is larger than its quantity choice would be if the firms chose quantities simultaneously and that its profits are larger as well. Also show that aggregate output is larger and that firm 2's profits are smaller.

(b) Draw a picture of this outcome using best-response functions and isoprofit contours.

12.C.19^C Do Exercise 8.B.5.

12.C.20^B Prove Proposition 12.C.2 for the case of a general convex cost function $c(q)$.

12.D.1^B Consider an infinitely repeated Bertrand duopoly with discount factor $\delta < 1$. Determine the conditions under which strategies of the form in (12.D.1) sustain the monopoly price in each of the following cases:

(a) Market demand in period t is $x_t(p) = \gamma^t x(p)$ where $\gamma > 0$.

(b) At the end of each period, the market ceases to exist with probability γ.

(c) It takes K periods to respond to a deviation.

12.D.2^B In text.

12.D.3^B Consider an infinitely repeated Cournot duopoly with discount factor $\delta < 1$, unit costs of $c > 0$, and inverse demand function $p(q) = a - bq$, with $a > c$ and $b > 0$.

(a) Under what conditions can the symmetric joint monopoly outputs $(q_1, q_2) = (q^m/2, q^m/2)$ be sustained with strategies that call for $(q^m/2, q^m/2)$ to be played if no one has yet deviated and for the single-period Cournot (Nash) equilibrium to be played otherwise?

(b) Derive the minimal level of δ such that output levels $(q_1, q_2) = (q, q)$ with $q \in [((a - c)/2b), ((a - c)/b)]$ are sustainable through Nash reversion strategies. Show that this level of δ, $\delta(q)$, is an increasing, differentiable function of q.

12.D.4[B] Consider an infinitely repeated Bertrand oligopoly with discount factor $\delta \in [\frac{1}{2}, 1)$.

(a) If the cost of production changes, what happens to the most profitable price that can be sustained?

(b) Suppose, instead, that the cost of production will increase permanently in period 2 (i.e., from period 2 on, it will be higher than in period 1). What effect does this have on the maximal price that can be sustained in period 1?

12.D.5[C] [Based on Rotemberg and Saloner (1986)] Consider a model of infinitely repeated Bertrand interaction where in each period there is a probability $\lambda \in (0, 1)$ of a "high-demand" state in which demand is $x(p)$ and a probability $(1 - \lambda)$ of a "low-demand" state in which demand is $\alpha x(p)$, where $\alpha \in (0, 1)$. The cost of production is $c > 0$ per unit. Consider Nash reversion strategies of the following form: charge price p_H in a high-demand state if no previous deviation has occurred, charge p_L in a low-demand state if no previous deviation has occurred, and set price equal to c if a deviation has previously occurred.

(a) Show that if δ is sufficiently high, then there is an SPNE in which the firms set $p_H = p_L = p^m$, the monopoly price.

(b) Show that for some $\underline{\delta}$ above $\frac{1}{2}$, a firm will want to deviate from price p^m in the high-demand state whenever $\delta < \underline{\delta}$. Identify the highest price p_H that the firms can sustain when $\delta \in [\frac{1}{2}, \underline{\delta})$ (verify that they can still sustain price $p_L = p^m$ in the low-demand state). Notice that this equilibrium may involve "countercyclical" pricing; that is, $p_L > p_H$. Intuitively, what drives this result?

(c) Show that when $\delta < \frac{1}{2}$ we must have $p_H = p_L = c$.

12.E.1[B] Suppose that we have a two-stage model of entry into a homogeneous-good market characterized by price competition. If potential firms differ in efficiency, need the equilibrium have the most efficient firm being active?

12.E.2[B] Prove that π_J is decreasing in J under assumptions (A1) to (A3) of Proposition 12.E.1.

12.E.3[B] Calculate the welfare loss from the free-entry equilibrium number of firms relative to the socially optimal number of firms in the models discussed in Examples 12.E.1 and 12.E.2. What happens to this loss as $K \to 0$?

12.E.4[B] Consider a two-stage model of entry in which all potential entrants have a cost per unit of c (in additional to an entry cost of K) and in which, whatever number of firms enter, a perfect cartel is formed. What is the socially optimal number of firms for a planner who cannot control this cartel behavior? What are the welfare consequences if the planner cannot control entry?

12.E.5[C] Consider a two-stage entry model with a market that looks like the market in Exercise 12.C.16. The entry cost is K. Compare the equilibrium number of firms to the number that a planner would pick who can control (a) entry and pricing and (b) only entry.

12.E.6[B] Compare a one-stage and a two-stage model of entry with Cournot competition [all potential entrants are identical and production costs are $c(q) = cq$]. Argue that any (SPNE) equilibrium outcome of the two-stage game is also an outcome of the one-stage game. Show by example that the reverse is not true. Argue that we cannot, however, have more firms active in the one-stage game than in the two-stage game.

12.E.7[B] Consider a one-stage entry model in which firms announce prices and all potential firms have average costs of $AC(q)$ (including their fixed setup costs) with a minimum average cost of $\bar{c}$ reached at $\bar{q}$. Show that if there exists a J^* such that $J^*\bar{q} = x(\bar{c})$, then any equilibrium of this model produces the perfectly competitive outcome and, hence, the outcome is (first-best) efficient.

12.F.1[B] Show that in the Cournot model discussed in Section 12.F with demand function $\alpha x(p)$, a firm's best-response function $b(Q_{-j})$ is (weakly) decreasing in Q_{-j} provided α is large enough.

12.F.2[B] Suppose each of the I consumers in the economy has quasilinear preferences and a demand function for good ℓ of $x_{\ell i}(p) = a - bp$.

(a) Derive the market inverse demand function.

(b) Now consider a Cournot entry model with this market inverse demand function, technology $c(q) = cq$, and entry cost K. Analyze what happens to the equilibrium prices and output levels, as well as what happens to consumer welfare (measured by consumer surplus), as $I \to \infty$ for both a one-stage and a two-stage entry model.

12.F.3[B] Analyze the two-stage Cournot entry model discussed in Section 12.F when α remains fixed but $K \to 0$. Show, in particular, that the welfare loss goes to zero.

12.F.4[B] Consider the following two-stage entry model with differentiated products and price competition following entry: All potential firms have zero marginal costs and an entry cost of $K > 0$. In stage 2, the demand function for firm j as a function of the price vector $p = (p_1, \ldots, p_J)$ of the J active firms is $x_j(p) = \alpha[\gamma - \beta(Jp_j/\sum_k p_k)]$. Analyze the welfare properties as the size (α) and the substitution (β) parameters change.

12.G.1[B] Consider the linear inverse demand Cournot duopoly model and the linear city differentiated-price duopoly model with differing unit costs that you examined in Exercises 12.C.9 and 12.C.17. Find the derivative, with respect to a change in firm 1's unit cost, of firm 2's equilibrium quantity in the Cournot model and equilibrium price in the linear city model. In which model is this change in firm 2's behavior beneficial to firm 1?

12.AA.1[A] In text.

12.AA.2[C] Prove Proposition 12.AA.4. [*Hint:* Consider a strategy profile of the following form: the players are to play an outcome path involving some pair (q_1, q_2) in period 1 and (q_1^*, q_2^*) in every period thereafter. If either player deviates, this outcome path is restarted.]

12.BB.1[A] In text.

12.BB.2[B] Show that if the incumbent in the entry deterrence model discussed in Appendix B is indifferent between deterring entry and accommodating it, social welfare is strictly greater if he chooses deterrence. Discuss generally why we might not be too surprised if entry deterrence could in some cases raise social welfare.

12.BB.3[C] Consider the linear city model of Exercise 12.C.2 with $v > c + 3t$. Suppose that firm 1 enters the market first and can choose to set up either one plant at one end of the city or two plants, one at each end. Each plant costs F. Then firm E decides whether to enter (for simplicity, restrict it to building one plant) and at which end it wants to locate its plant. Determine the equilibrium of this model. How is it affected by the underlying parameter values? Compare the welfare of this outcome with the welfare if there were no entrant. Compare with the case where there is an entrant but firm 1 is allowed to build only one plant.