

JOSEPH E. STIGLITZ

Princeton University

Technological Change, Sunk Costs, and Competition

THERE IS A PERSISTENT divergence between the widespread views of the virtues of capitalism and the models we use to justify those beliefs. While it is the dynamic properties of capitalism, the increases in the standards of living to which it has given rise, that constitute the basis of our confidence in its superiority to other forms of economic organization, the theory—at least the version we teach to students—is based on a model that assumes an unchanging technology.

This divergence is disturbing, and not only because of the intellectual dissonance to which it should give rise. Views concerning important policy issues are frequently based on simplistic models, regardless of how inappropriate those models are. For instance, considerations of the need for and consequences of antitrust policies, including policies aimed at restricting vertical restraints of trade, are frequently based on the standard competitive paradigm. In that context it is argued, for instance, that producers would impose vertical restraints only if the restraints enhanced efficiency. Yet markets in which technological change is important are never perfectly competitive, and in imperfectly competitive markets, vertical restraints may also serve to alter the degree of competition.¹

Helpful discussions with Partha Dasgupta, Steve Salop, and Richard Gilbert are gratefully acknowledged. I am also grateful to Mark Schwartz, Andrew Postewaite, Richard McMasters, Hal Varian, and the participants at the Brookings Conference for their helpful comments. This paper is based on research supported by the National Science Foundation, the Olin Foundation, and the Hoover Institution, Stanford.

1. In monopoly markets, vertical restraints could be adopted either because they enhance efficiency or because they enhance the ability of the monopolist to discriminate in pricing. Studying the polar cases of monopoly and competition may thus not provide

In *Capitalism, Socialism, and Democracy*, Joseph Schumpeter long ago argued that any analysis of capitalism must be based on models in which there is technological change, in which competition takes the form of developing new products and processes rather than, or in addition to, the price competition that is the subject of concern in conventional price theory. He argued that markets with technological change were inherently imperfectly competitive. But while Schumpeter recognized the relative unimportance of price competition, he did not fully explain it.

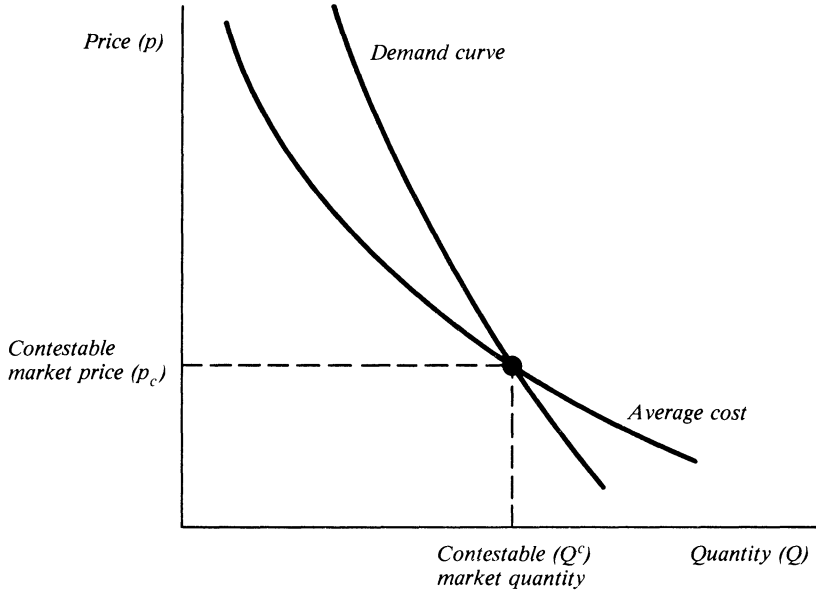
There are a number of reasons why the conventional theory of competition does not describe well the industrial sectors in which technological change is important, among them that technological change inherently entails an element of increasing returns and that expenditures on R&D are, for the most part, sunk costs.

As has long been recognized, when there are increasing returns, competition may not be viable: both market equilibrium and economic efficiency entail one firm, or at most a few. In recent years, however, arguments have been put forward that even markets in which there are a limited number of competitors may behave competitively because of the discipline provided by potential competition. This argument is called the contestability doctrine.² Industries with increasing returns will have one firm operating—as efficiency requires. And that firm operates at the highest output at which price equals average costs—that is, where profits are zero (figure 1). If the monopolist operated at a lower level of output, with a price higher than average cost, an entrant would contest the market, entering with a lower price and stealing all the customers away from the incumbent firm.

insights into the more prevalent case of imperfect competition, where such practices attempt to change the extent of competition. For further discussion of the role of vertical restraints in altering the degree of competition, see P. Rey and Joseph E. Stiglitz, "The Role of Exclusive Territories in Producers' Competition," unpublished paper (1987); and Rey and Stiglitz, "Imperfect Competition and Vertical Restraints," *European Economic Review* (forthcoming).

2. There are clearly antecedents of this view in the "Chicago School"; see, for example, Harold Demsetz, "Why Regulate Utilities," *Journal of Law and Economics*, vol. 11 (April 1968), pp. 55–66. Its most ardent current advocates include Sanford J. Grossman, "Nash Equilibrium and the Industrial Organization of Markets with Large Fixed Costs," *Econometrica*, vol. 49 (September 1981), pp. 1149–72; William J. Baumol, "Contestable Markets: An Uprising in the Theory of Industrial Structure," *American Economic Review*, vol. 72 (March 1982), pp. 1–15; and Baumol, John C. Panzer, and Robert D. Willig, *Contestable Markets and the Theory of Industry Structure* (Harcourt Brace Jovanovich, 1982).

Figure 1. Equilibrium in Contestable Markets



If the contestability doctrine were correct, it would radically alter our attitudes toward antitrust policy. It would imply that an important class of circumstances—those in which there was a natural monopoly (or oligopoly) because of increasing returns—which previously had been thought to give rise to market failure (Pareto inefficiencies), need not or would not do so. Government intervention, if this view were correct, would not be required, even when there was only one firm or a few firms. And indeed, the typical form of government intervention, which entails breaking up firms or limiting their agglomeration, could be positively harmful, since the potential economies of scale would not be fully exploited.

While traditional economic analysis argued that in the presence of increasing returns there was a trade-off between having many firms with less monopoly power but a loss in productive efficiency or a few firms with more power and more efficiency, the contestability doctrine claims there is no trade-off.

For countries concerned about their technological leadership, these issues are of particular importance. When the technological change is a result either of R&D or of learning by doing, there are likely to be

increasing returns and a high degree of concentration of market power. If potential competition suffices to ensure not only that the industry behaves in an efficient manner but that the efficiency gains are passed along to consumers, then government intervention to break up or limit the size of those firms that have gained their dominant position through technological advances may seem particularly inadvisable.

The purpose of this paper is to address the validity or generality of the premise that potential competition suffices to ensure economic efficiency. It is a paper as much about the models used to justify our beliefs about the design of economic policy as it is about what those policies should be. The latter is a far more difficult question, about which I have a few observations in the concluding section.

More precisely, I contend that the traditional arguments *against* such government policies to open up competition as antitrust regulations, regulations for fair trade practices, and so forth are at best of limited validity and relevance in most sectors of modern economies. These arguments hold that

- competition ensures efficiency, and the gains from technological change are passed along to consumers through low prices; and
- profits attract entry, ensuring that markets must behave competitively.

The contestability doctrine emphasizes that to assess the force of competition, one should not look at the number of existing firms but at potential entrants, and in most instances there are many of these. But just as with Bertrand (cutthroat) competition among firms in a market, in which it takes only two firms to obtain competitive solutions, so, the doctrine argues, it takes only one potential entrant, or at most a few entrants, to ensure both economic efficiency and zero profits.

By contrast, I argue that these results are highly sensitive to the assumption of *zero* sunk costs. With sunk costs, even very small ones, I argue:

- The existence of profits may not attract entry. Potential entrants will make a judgment about what will happen *after* they enter, and positive profits for the incumbent monopolist may entail negative profits for an entrant who attempts to contest that market. Thus potential competition does not suffice.

- Entry may not entail competition. The entrant or entrants and the incumbent firm may collude rather than compete. (And the collusion may be tacit rather than explicit, enforced by each firm's understanding

of what is in its self-interest.) Alternatively, entry may simply lead to one or another firm's exit—but not to competition.

—Incumbent firms will take action to deter entry; entering firms will take action to facilitate collusion. Among these actions is the choice of technology, including the decision about whether to use technologies that entail large sunk costs. Some of the entry-detering and collusion-facilitating actions reduce social welfare. Thus the fact that there is some competition or potential competition should not be taken as *prima facie* evidence that some action undertaken by a firm enhances efficiency.

—Even when competition is successful in reducing profits to zero, efficiency is not ensured. The kind of competition with which we are concerned here does not ensure that price equals marginal cost: there may be “rents” associated with being the first firm to produce a new product. But competition to be the first firm may be sufficiently fierce that those rents are largely dissipated. Competition for those rents may give capitalism its essential dynamism but need not lead to efficiency in the standard Paretian sense.

Because technological change inherently involves increasing returns and sunk costs, the contestability doctrine is particularly inapplicable to industries in which technological change is important; potential competition ensures neither economic efficiency nor zero profits.

The paper is divided into five parts. In the next part, I argue in the context of a static model that potential competition will not suffice to ensure economic efficiency when there are even small sunk costs, and that sunk costs are in fact pervasive. The subsequent section considers several important objections that have been raised to this argument. I show that they are either not valid or that they strengthen the concern about the limited efficacy of potential competition. I then extend the analysis to the problems of sunk costs associated with innovation. The paper concludes with some brief remarks concerning the interpretation of the results for economic policy.

Sunk Costs and Potential Competition

A major insight of the work on industrial organization over the past decade is that what matters for entry and the nature of competition is how firms will interact should entry occur, or—since these are generally not known—the beliefs of potential entrants about the nature of those

interactions.³ An incumbent firm's prices may be high today, but that does not necessarily result in entry if potential entrants believe that entry will result in fierce competition and lower prices. There is thus an intimate connection between *ex ante* competition, competition before entry, and *ex post* competition, the nature of competition after entry. Indeed, the possibility of excessively fierce *ex post* competition actually reduces the effectiveness of *ex ante* competition. The variables that determine the nature of *ex post* competition, or potential entrants' beliefs about the nature of *ex post* competition, are called state variables.

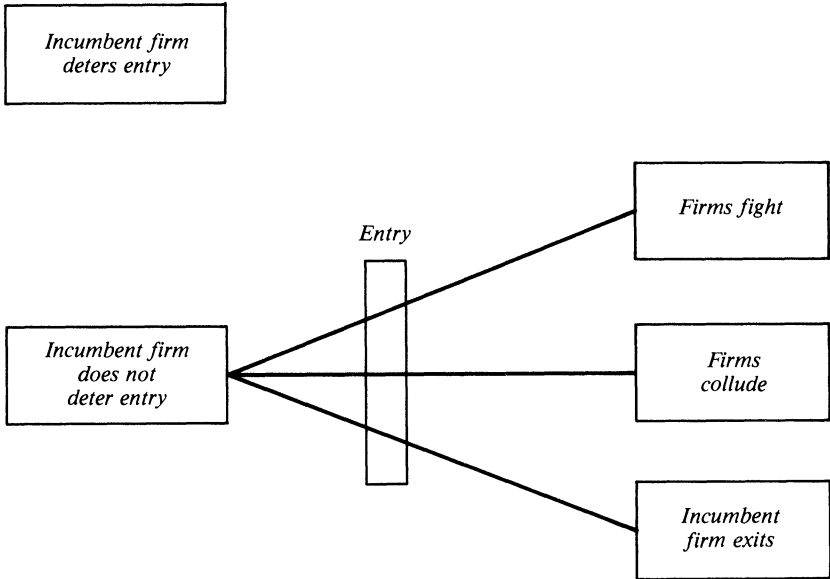
To examine these questions, I assume there is initially a single firm in the industry. I then divide the analysis into two stages: equilibrium after entry (should entry occur) and actions taken before entry, which affect the likelihood of entry and the equilibrium should entry occur. There are three possible responses to entry: the firms can compete; they can collude, sharing the monopoly profits; or the incumbent firm may exit. Competition in turn can take on a number of different forms—there is no agreed-upon theory of competition among duopolists. I focus on Bertrand (price) competition, because the concern here is to show that, even under the seemingly most favorable conditions where it is extremely effective in driving down prices, competition does not ensure either zero profits or economic efficiency. Later in the paper I show how the results are modified when *ex post* competition is less effective.

Thus one needs to ascertain the circumstances under which competition, collusion, and exit will occur. And this is precisely the same question facing the incumbent firm. It wants to take actions in the first stage that make the world most unattractive for any potential entrant. It seeks to deter entry. It tries to convince potential rivals that, should they enter, there will not be accommodation through collusion and a sharing of profits but war—sufficiently fierce competition that the entrant will regret its decision. Figure 2 shows the timing-decision structure.

3. Among the early contributors were Steven C. Salop, "Strategic Entry Deterrence," *American Economic Review*, vol. 69 (May 1979, *Papers and Proceedings*, 1978), pp. 335–38; Avinash Dixit, "The Role of Investment in Entry Deterrence," *Economic Journal*, vol. 90 (March 1980), pp. 95–106; A. Michael Spence, "Entry, Capacity, Investment, and Oligopolistic Pricing," *Bell Journal of Economics*, vol. 8 (Autumn 1977), pp. 534–44; Richard Gilbert and Joseph E. Stiglitz, "Entry, Equilibrium, and Welfare," paper presented to the NBER-NSF Conference on Industrial Organization, Toronto, October 1978; and Stiglitz, "Potential Competition May Reduce Welfare," *American Economic Review*, vol. 71 (May 1981, *Papers and Proceedings*, 1980), pp. 184–89.

Figure 2. Time-Decision Structure for Basic Model

Stage 1



I began this section by arguing that a major insight of recent work in industrial organization was that what mattered for entry was the nature of competition after entry, and that was determined by state variables, such as binding commitments and irreversible investments. Expenditures on sunk costs—costs that once expended cannot be recovered—are state variables and have, as we shall see, a critical effect on competition.

Investment expenditures need not be completely sunk. Because an airplane purchased by one firm can easily be resold to another, an investment in an airplane is not sunk. By the same token the number of airplanes an airline owns should not be viewed as a state variable. While much of investment is not sunk, however, there is a sunk cost element in almost all investments. An airline must advertise to obtain customers; it must solve complicated routing problems. It is therefore important to ascertain the nature of the competitive interactions that arise when there are sunk costs, particularly when there are small sunk costs.

The next two sections show that when there is strong ex post

competition even small sunk costs act as effective entry barriers. There are two ways in which sunk costs can be small: they may be small relative to marginal production costs or they may not be very sunk—a short-lived, irrecoverable investment is, it would seem, not very sunk. Even when sunk costs are small in either sense, potential competition may not be effective.

SMALL SUNK COSTS CAN CREATE COMPLETE ENTRY BARRIERS

Assume two firms, labeled 1 and 2. If a_1 and a_2 are the actions taken by them, respectively, their profits are denoted by $\pi_1(a_1, a_2)$ and $\pi_2(a_1, a_2)$. The firms are engaged in a noncooperative game. Suppose that (a_1^*, a_2^*) is the unique equilibrium point of the game and that $\pi_1(a_1^*, a_2^*) = \pi_2(a_1^*, a_2^*) = 0$, that is, equilibrium profits are zero.

Next assume that firm 2 can undertake a passive action, not entering the market, which is denoted by \hat{a}_2 . The best response by firm 1 to \hat{a}_2 is \hat{a}_1 . Assume that $\pi_1(\hat{a}_1, \hat{a}_2) > 0$, that $\pi_2(\hat{a}_1, \hat{a}_2) = 0$, and finally that there is no action available to firm 2 such that $\pi_2(\hat{a}_1, a_2) > 0$. That is, if the first firm takes action \hat{a}_1 , then \hat{a}_2 is the best response for firm 2.

Now embed this game in a larger game. Begin by regarding firm 1 as the incumbent and firm 2 the potential entrant. If firm 2 enters, the subsequent game played by the firms is the one above. But it costs ϵ (> 0) to enter, and this cost, one assumes, has to be sunk. Assume finally that if firm 2 does not enter, it is restricted to the passive action \hat{a}_2 . One can now easily confirm

proposition 1: If $\epsilon > 0$, there is a unique, subgame-perfect equilibrium in which firm 2 does not enter and firm 1 chooses \hat{a}_1 , thus earning $\pi_1(\hat{a}_1, \hat{a}_2) > 0$. If $\epsilon = 0$, there are two equilibria: one is the pair of actions (a_1^*, a_2^*) at which the firms earn no profits; the other is the equilibrium in which firm 2 does not enter and firm 1 chooses \hat{a}_1 .⁴

The argument is straightforward. If entry occurs, by assumption the profits of the entrant in the postentry period are zero. But to enter, firm 2 encounters a cost of ϵ ; its total profits from entering are thus $-\epsilon$. If it does not enter, its profits are zero. It is better for it not to enter.

4. Note that the second firm is indifferent to choosing between these two equilibria, but the first firm clearly is not.

Proposition 1 is disturbing because it says that even a negligible entry cost can fully cushion the incumbent against encroachment. As long as ϵ is positive, it is of little significance whether a potential entrant is waiting in the wings. Notice as well that the proposition does not deny that the incumbent firm may itself have had to incur a sunk cost of ϵ to enter in the first instance. As long as firm 1 had the first move, it would enter, provided ϵ was not too large, and as proposition 1 asserts, firm 2 would not find it profitable to enter.

This discussion has deliberately left the actions of the firms uninterpreted in economic terms. The actions may include signing contracts with customers to supply commodities. Provided costs have to be sunk to enter the competition, proposition 1 stands intact. The following discussion provides a specific economic example of proposition 1.

Consider a market for a homogeneous product. The market demand curve slopes downward and the unit cost of production is a constant $C \geq 0$. Firms 1 and 2 compete in the way suggested by Bertrand. Each takes the other's price as given. Bertrand competition implies cutthroat behavior by firms, with price being driven to marginal costs, C . I focus on it not only because it may describe competition in certain markets well—the recent models of contestable markets assume that such fierce competition is empirically important—but also as I suggested earlier, Bertrand competition provides, one might have thought, the most favorable circumstances for the doctrine that competition ensures efficiency and zero profits.

Now suppose that entry into the industry requires a cost $\epsilon > 0$ to be sunk, and firms 1 and 2 move sequentially as regards entry. Let π denote monopoly profits in this industry exclusive of entry costs. It is then easy to confirm

proposition 2: If $0 < \epsilon < \pi$, there is a unique, subgame-perfect equilibrium in which only firm 1 enters and it earns $\pi - \epsilon$ as net profit.⁵

5. For notational simplicity I am supposing that the interest rate is zero. If the firms move simultaneously in a two-stage game in which entry precedes price competition, there are three subgame-perfect equilibria: 1 enters and 2 does not (as in proposition 2); 2 enters and 1 does not; and a symmetric mixed-strategy equilibrium (where each firm enters with probability p) in which the expected profit of each firm is zero. Such symmetric outcomes are ignored in the text because I wish to explore the effect of potential entrants on an incumbent's behavior and hence on the eventual industrial structure.

Only the asymmetric equilibria are productively efficient. There is a probability of p^2

Notice again that even if ϵ is small, firm 2 poses no threat to firm 1; the outcome is as though there *is* no potential competition. The result is disturbing because the technology that I have postulated differs insignificantly from the standard assumption of constant returns to scale. There is a slight element of increasing returns, caused by the small sunk costs. Contestability suggests that, while even with this small degree of increasing returns there will be only one firm operating in the market, price should be just slightly above marginal cost; my analysis suggests that price should be equal to the monopoly price.

What are the policy implications? Any policy that gets the entrant into the industry may improve social welfare. Two such policies suggest themselves. One is to subsidize entry. In this case the traditional argument against such subsidies—that if it is socially desirable to enter, it must be privately desirable, and therefore subsidies are not needed—is simply incorrect. The entrant forces the market price down to marginal costs, but consumers capture social gains.

A second policy is minimum-price legislation (retail price maintenance). Let $D(q)$ denote market demand at price q . Now let

$$(1) \quad \bar{q} = C + [2\epsilon/D(\bar{q})]$$

(see figure 3). It is then an easy matter to confirm

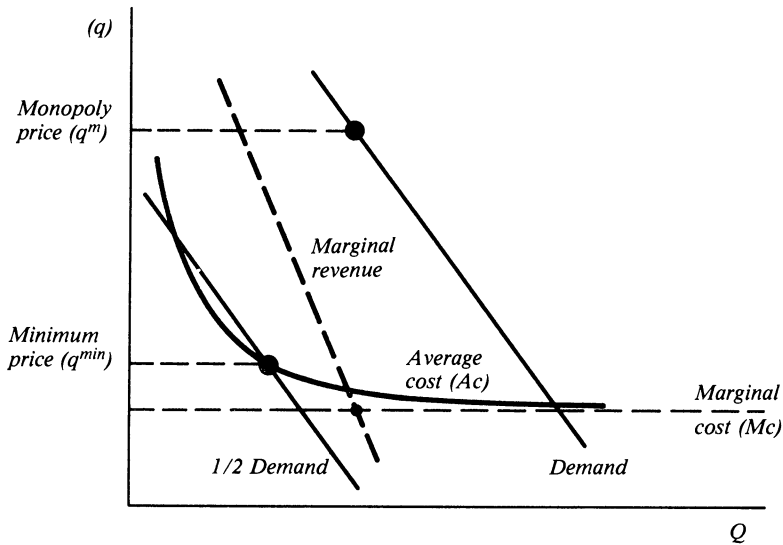
proposition 3: If the government sets a minimum price slightly above \bar{q} , there is a unique, subgame-perfect equilibrium in which both firms enter and charge the minimum allowable price.

The idea behind this policy is simple enough. Even a small entry cost provides complete protection to the incumbent in proposition 2 because the model envisages cutthroat competition should both firms enter. In proposition 3 the government prohibits cutthroat competition to allow firms to earn some profits should both enter, thus making it profitable to enter. Notice as well in equation 1 that q is only slightly in excess of C if ϵ is small. Thus if entry cost is small, the minimum-price legislation identified in proposition 3 sustains a nearly efficient outcome.

I am not convinced that there are many circumstances in which this

that there will be duplication of the sunk costs, and a probability of $(1 - p)^2$ that no firm enters. Of course, the asymmetric equilibria are inefficient in that price exceeds marginal costs. Even with a simple criterion, such as the sum of consumer plus producer surplus, the overall welfare comparisons are not obvious: they depend on the magnitude of sunk costs and consumer surplus.

Figure 3. Retail Price Maintenance



model provides a rationale for retail price maintenance, partly for reasons given later in this paper. Nor, because of the difficulties of determining the circumstances in which these policies would increase welfare, am I convinced that either of the two policies is generally desirable.⁶

Still, the analysis of this section should serve as a warning. If such a slight modification of the standard technological assumptions underlying analyses of competition leads to such a radical change in both market equilibrium and the desirability of government intervention, can we have confidence in policies derived from these analyses, and more particularly from the more recent variants based on the contestability doctrine?

6. Were these accepted as legitimate arguments for entry subsidies, potential entrants to a market would argue that such circumstances apply to them. Similarly, if the policies were accepted as legitimate arguments for retail price maintenance, firms in the industry would argue that theirs were precisely the circumstances under which, without RPM, there would be destructive competition. In the circumstances in which these policies were truly appropriate, there would be no rents (profits) even with these government policies, and hence no rent-seeking activity.

Of the two policies, the subsidy has the advantage that it requires a one-time action by the government, not long-run regulation. The effectiveness of the RPM policy hinges on the government's knowing the demand function and the technology (the value of sunk costs).

SUNK COSTS, DURABILITY, AND NATURAL MONOPOLY

The essential feature of sunk costs is that they represent expenditures that cannot be recovered. They frequently represent expenditures on nontransportable, nonconvertible plant and equipment, such as the site preparation work for a nuclear power plant. But even in the best of circumstances the plant and equipment are of finite duration; the planned life of a plant may be only thirty or fifty years. There is a natural connection between durability and sunk costs. If a machine was designed to last only five minutes, expenditures on it would not be considered investment and would not typically be viewed as representing significant sunk costs.

The previous sections showed how the presence of even an epsilon (ϵ) sunk cost could alter the equilibrium in a significant way. This section reinforces that conclusion by considering what happens as the durability of a machine goes to zero—again, there are negligible sunk costs. But while in the previous section potential competition had no effects, this section shows that although consumers are not benefited—prices remain at the monopoly level—profits are decreased: potential competition represents a Pareto worsening of welfare.

I investigate these questions in the context of a natural monopoly, one in which efficiency considerations dictate that there should be only one firm producing. Natural monopolies have always been at the center of discussions on economic regulation. Recall the earlier discussion of the contestability doctrine, which holds that with a natural monopoly (decreasing average costs), price will be set equal to the average cost and government regulation will be unnecessary.⁷ My analysis shows that conclusion is not valid so long as there are any durable sunk costs.

7. It is further argued that if the natural monopolist produces several commodities, it will use Ramsey prices, just as the government would. Though the monopolist will indeed not simply set marginal revenue equal to marginal cost (as in the simpler theory of monopoly), and the prices it charges will have a superficial semblance to Ramsey prices, in general the prices will not be the same as those that a government agency running the firm and maximizing social welfare would set. For an example of what is at issue, see David E. Sappington and Joseph E. Stiglitz, "Information and Regulation," in Elizabeth E. Bailey, ed., *Public Regulation: New Perspectives on Institutions and Policies* (London: MIT Press, 1987), pp. 3–43.

Consider the market for a homogeneous commodity facing a demand function $q(Q)$, where Q is the flow of output at any date and q the corresponding market-clearing price. In order to produce, facilities costing F have to be constructed. The lifetime of such facilities is precisely T years, and they depreciate all at once in year T . Once such facilities are installed, the commodity can be produced at marginal cost C (≥ 0). There are therefore no capacity limits. Let r (> 0) denote the rate of interest. For convenience let time be continuous. Firm revenues, R , net of variable costs are then

$$(2) \quad R(Q) \equiv q(Q)Q - CQ.^8$$

Let Q^m maximize $R(Q)$; Q^m is the output a monopolist would produce if there were no potential entrant in the industry. The technology and demand curves are such that one, but only one, firm can earn enough over time to cover the fixed cost of entry.⁹

If one assumes that F represents fixed costs that are not sunk, then the theory of contestable markets asserts that if the market in question is contestable, the threat of potential entry forces prices down to a level at which profits are zero—that is, the present discounted value of revenues net of variable costs is just equal to the fixed costs, with a new facility being built every T years. The contestable output, Q^c , is given by the solution of the zero-profit condition:

$$(3) \quad R(Q^c) = rF/(1 - e^{-rT}).$$

It is clear that $Q^c > Q^m$, that contestable output exceeds monopoly output.

Suppose now that F represents sunk costs. Assume that once production facilities have been constructed, there is Bertrand competition, and the equilibrium price of the product equals the marginal cost of production, C . If there is a potential entrant, the incumbent firm will deter entry not by raising output to Q^c but by constructing production facilities with greater frequency than T .

The incumbent sinks F at $T = 0$. The rival firm will achieve no

8. I assume that $R(Q)$ is concave in Q . In the more general case the analysis is more complicated, but the conclusions are unaltered.

9. This is called a strong natural monopoly. The condition for it is $2F/(1 - e^{-rT}) > R(Q_m)/r > F/(1 - e^{-rT})$. If this equation holds, then only one firm can be economically viable in the industry.

advantage by entering immediately, since over the life of the machine, Bertrand competition ensures that price will just equal C , and the entrant will be unable to recover sunk costs. There exists, however, some date before T such that it pays the rival to enter and sink F at that date if the incumbent has not already constructed a second set of production facilities. For although the entrant will earn zero profits while both firms are producing and until the original firm's plant wears out, after that date the entrant becomes the monopolist. But of course, if it pays a rival to enter at this date, it certainly pays the incumbent to sink F just before the date. By preempting its rival, the incumbent firm deters the entrant from entering and can keep price at the monopoly level. To deter entry, then, the incumbent constructs new production facilities every τ years, where $0 < \tau < T$; and since the monopolist is protected against entry, it produces at the monopoly rate, Q^m .

To compute the entry-detering τ , suppose the rival enters by sinking F at τ ($< T$) and announces that it will sink F every τ years. If this is carried out, the incumbent will depart at T . During the interval $[\tau, T]$ the duopolists will earn no operating profits, and from T onward the entrant will earn $R(Q^m)$ at each moment. From the vantage point of $t = 0$, the present value of profits accruing to the entrant is thus

$$(4) \quad [R(Q^m)e^{-rT}]/r - [Fe^{-r\tau}/(1 - e^{-r\tau})].^{10}$$

The expression is zero at a unique value of τ : call this τ^* . Thus τ^* is the replacement interval for which an entrant that subsequently becomes the single producer just breaks even. τ^* is the solution to

$$(5) \quad [R(Q^m)e^{-rT}/r] = [Fe^{-r\tau}/(1 - e^{-r\tau})].$$

The stationary subgame-perfect equilibrium of the model consists in the incumbent's building production facilities every τ^* years. If the incumbent ever delays, it will have lost the market to the rival, for if the incumbent does not construct the second set of production facilities by τ^* , the rival gains an advantage by entering. Once the rival enters by sinking F , the incumbent cannot find it profitable to construct facilities

10. The first term is just the present discounted value of the constant flow of operating profits $R(Q^m)$, beginning at date T when the incumbent firm leaves the industry. The second term is just the sum of the infinite geometric series, $F[e^{-r\tau} + e^{-2r\tau} + e^{-3r\tau} + \dots + e^{-nr\tau}]$, representing the present discounted value of expenditures on plant every τ periods.

yet again and then becomes just like any other potential entrant. The incumbent, knowing this, is forced to construct new facilities every τ^* years. And given that it does so, no rival enters. Since $\tau^* < T$, deterrence is achieved at a cost: the incumbent has to maintain idle facilities over the intervals of time between the construction of the new facility, τ , and the death of the old, T .

In the absence of competition the present value of the incumbent's profits is

$$(6) \quad [R(Q^m)/r] - [F/(1 - e^{-rT})].$$

We have seen that the threat of competition reduces the incumbent's profits to

$$(7) \quad [R(Q^m)/r] - [F/(1 - e^{-r\tau^*})].$$

Consumers are entirely unaffected by the presence of potential competition, which achieves only a reduction in the incumbent's profits. Potential competition thus leads to a Pareto inferior outcome.

Though profits are reduced, they are not driven to zero. Because $\tau^* < T$, a comparison of equations 5 and 7 shows that the incumbent's strategy to deter entry yields positive profits except in one limiting case. If the rate of interest is small, equation 5 reduces to $R(Q^m)\tau^* \approx F$, and profits in the interval during which the machine produces (before it is replaced by a new machine) must just equal the fixed costs. It follows then that the incumbent firm, which is forced to replace its machine at time intervals of τ^* , must also be making almost no profits. But even here the outcome does not resemble at all that of contestable markets. Profits per investment cycle are wiped out not by expanding output to Q^c but by constructing production facilities far too frequently. Thus

proposition 4: The threat of entry forces the incumbent firm to construct production facilities more frequently than is necessary. Entry is deterred not by expanding output but by maintaining idle capacity during certain intervals of time. Potential competition is deleterious to social welfare. Profits of the incumbent remain positive, except in the limiting case of a zero interest rate.

This model has obvious policy implications. Competition, actual or potential, is not enough to ensure the efficiency of a natural monopoly.

Government regulation, including such restrictions on entry as franchising, is required.¹¹

Now, what happens as sunk costs become less sunk, that is, as the durability of the equipment goes to zero? To obtain a meaningful comparison, assume that the present discounted value of expenditures on fixed (sunk) costs remains the same (if the machine is replaced when it wears out). Letting $F(T)$ denote the fixed costs associated with a technology in which machines last T years, assume that

$$F/(1 - e^{-rT}) = k \text{ for all } T.$$

Substituting in equation 5, one obtains $\tau(T)$, the length of time between the construction of the last machine and the construction of the next machine, as a function of the technology, T . Then $\tau(T)$ is the solution to

$$[R(Q^m) e^{-rT}]/(1 - e^{-rT}) = (kr e^{-r\tau})/(1 - e^{-r\tau}).$$

Clearly, as T goes to zero, τ goes to zero. For small t ,

$$\tau/T \approx kr/R(Q^m);$$

the ratio does not go to zero. The price remains at the monopoly price and profits do go to zero since the effective interest rate over a cycle is close to zero; but all of the potential monopoly profits are dissipated as excess expenditures on capacity duplication. Even though the expenditures on the plant are sunk for only one month, that is all that is required for the incumbent to maintain its monopoly position. Potential competition simply forces it to incur the sunk costs associated with a second plant (costs that a potential entrant would have to spend as well), say two weeks after the first plant has been constructed. Making costs less sunk in this way has simply increased the inefficiencies associated with potential competition. The welfare of consumers has not been increased.

EXIT AND ENDOGENOUS COMPETITION

The intensity of ex post competition is critical for determining the effectiveness of ex ante competition: the more intensive ex post competition, the less effective ex ante competition.

11. One needs to raise the same caveats noted in the discussion of the policy implications of the previous model: the government may have difficulty ascertaining the circum-

When a firm enters a market, the entrant and the incumbent might collude rather than compete. Alternatively, the incumbent might exit. In either case, entry will not have resulted in low prices for consumers. If collusion, tacit or explicit, is anticipated and in fact occurs, entry will be encouraged. The limited competition provided by entry will simply dissipate resources in excess entry. Consumers may be little better off.

Sunk costs are again crucial to determining the nature of the equilibrium. Remember that sunk costs can be thought of as costs that are *not* recovered upon exit. Indeed, there may even be additional costs, such as those associated with waste disposal, incurred upon exit.

The choice of technology by the incumbent and the entrant affects whether there will be exit, collusion, or competition. And sometimes forces conflict. With large exit costs (high sunk costs) collusion is more likely, and this attracts entry. If the incumbent firm reduces sunk costs, its incentives to exit when a rival enters are increased, and this too makes entry attractive. The incumbent may not be able to deter entry, but entry may also not be effective in ensuring competition. That there are profits after exit does not lead to regret on the part of the exiting firm any more than the fact that profits before entry necessarily lead to entry: the relevant question is not what is the current level of profits, but what would profits have been? Profits after exit may be positive but would have been zero had the incumbent not exited, and it is this that drives exit.

Forcing exit of the incumbent. Assume, as before, that the entrant and the incumbent have the same marginal costs of production and that after entry there is a Bertrand equilibrium. Let F_e be the amount the entrant receives upon exit (the amount of its original expenditures that it recovers).¹² If a firm enters the market, the incumbent will be better off leaving if $F_e > 0$. So long as the incumbent firm can receive something when it exits, there is a perfect equilibrium in which the entrant enters and the incumbent leaves.

There is, of course, another perfect equilibrium in which the entrant

stances under which the model is applicable, that is, in which entry restrictions—with price regulations—are desirable. Some industries may attempt to obtain protection from competition even when such restrictions do not enhance welfare as long as they believe the price regulations are not so effective as to reduce profits to zero.

12. If the firm incurs costs upon exit, F_e is negative.

leaves, and a mixed-strategy equilibrium in which there is a given probability that each leaves. As I have modeled the problem, one cannot choose among the alternative perfect equilibria. Assume, however, that the greater the expenditure on sunk costs, the lower the production costs. Then whatever the technology chosen by the incumbent, the entrant will choose a technology with slightly lower production costs (higher sunk costs). Thus, in the Bertrand equilibrium, the entrant will have a slightly positive return on variable costs. If one assumes that there are at least some slight costs associated with continuing a production line (whether it produces or not), and denotes the present discounted value of those costs by C/r , then the *only* perfect equilibrium will entail the original incumbent's leaving, so long as $C/r + F_e > 0$.¹³

The persistence of monopoly and the advantageous position of incumbents. That the original firm will leave the market should be contrasted with the results of Partha Dasgupta and Joseph Stiglitz on the persistence of monopoly, showing that an existing monopolist could continue to maintain its monopoly position.¹⁴ Here, I have delineated a rather different set of circumstances in which the entering firm can choose its technology and an incumbent cannot maintain its monopoly position. Latecomers always have an advantage. This result seems more in accord with Schumpeter's vision of a succession of monopoly firms dominating each market.

But while competition seems more viable than in the analysis by Dasgupta and Stiglitz, competition still does not ensure efficiency (or zero profits). Prices remain at a monopoly level, and entry may have entailed socially unnecessary expenditures on sunk costs.¹⁵ These expenditures are simply the means by which the entrant "steals" profits away from the incumbent.

More potential competitors may entail less effective competition. Earlier, I argued that the more effective ex post competition was, the less effective potential competition would be. There are some

13. Appendix A provides an alternative and simpler formulation of the exit-entry game in which the threat of entry has no effect on price, either because there will be no entry or entry will immediately be followed by the exit of the incumbent.

14. Partha Dasgupta and Joseph E. Stiglitz, "Uncertainty, Industrial Structure, and the Speed of R&D," *Bell Journal of Economics*, vol. 11 (Spring 1980), pp. 1-28.

15. Because the entrant has chosen a technology with slightly lower marginal costs, however, price will be slightly lower after entry than before. Unnecessary expenditures on sunk costs will normally be the case if the incumbent does not recover all initial expenditures upon exit.

circumstances under which the easier entry is, the less effective potential competition will be.

There is a simple explanation of this seeming paradox. Potential entrants will enter only if they believe that they can earn profits. But if they believe that profits will be quickly stolen by subsequent entrants, they may be deterred from entry. It is not only the incumbent that deters entry: potential entrants themselves deter entry. Entry is motivated by the search for monopoly rents.

The model just presented illustrates this point. Assume there are many potential competitors but that, for simplicity, only one competitor can enter each period. Any potential entrant knows there can be a still later entrant that will steal the market away from it. If the fixed costs of entering (net of the costs recovered upon exit) are greater than the maximized value of operating profits for the one period during which the entrant would dominate the market, it will not pay to enter.

The incumbent, knowing this, may be able to design an appropriate strategy to deter entry. Let $N(C)$ be the value of nonrecoverable sunk costs associated with a technology that has variable costs C .¹⁶ Let $R(Q, C)$ be operating profits when marginal costs are C and output is Q , and let $Q^m(C)$ be the monopoly output corresponding to marginal costs C . The incumbent can deter entry by choosing a technology with low enough variable costs. The entry-detering technology is the solution to

$$(8) \quad R[Q^m(C^d), C^d] = N(C^d).^{17}$$

Potential entry has now simply distorted the production decision of the incumbent firm. The incumbent chooses a technology with a lower variable cost, and this leads to a lower price—but the price is not the competitive price, profits are not zero, and indeed, the technology chosen will in general be inefficient.¹⁸

Limiting case of small sunk costs. I now consider the limiting case in

16. That is, if the firm incurs a cost of ϵ upon entering and recovers F_e upon exit a period later, then $N = \epsilon - [F_e/(1 + r)]$, where r is the rate of interest.

17. The entry-detering technology given in this equation assumes that the entrant remains only one period. If the entrant remains permanently, the corresponding equation is $R[Q^m(C^b), C^b]/r = \epsilon(C^b)$, where $\epsilon(C)$ is the entry cost for a technology with variable costs C . If there are some costs recoverable upon exit, then $N(C) < \epsilon(C) > r\epsilon(C)$, provided the interest rate is less than unity. Thus in the general case, entry deterrence requires choosing $C = \min[C^d, C^b]$.

18. The efficient technology is the one that minimizes the present discounted value of costs at the equilibrium output level. That is, if $\epsilon(C)$ is the entry cost with technology

which sunk costs go to zero. At the same time, I consider what happens as the number of potential entrants increases. To do this, I modify the model slightly, moving to the straightforward continuous-time adaptation of the discrete time model. Let ϕ represent the probability of the arrival of a potential entrant per unit of time. For what value of C is it that the expected returns to entering are just zero? Assume for the moment that there is no entry deterrence. The expected net present discounted value of profits is $\{R[Q(C), C]/r\} - \epsilon$. The incumbent firm lowers this—and lowers its own profits—by lowering C . If it lowers C enough, the profits will be zero and entry will be deterred. Denote the technology for which this is true by C^d . Hence the first firm's expected profits (gross of entry costs) from pursuing the entry-detering strategy are just ϵ .

Alternatively, the incumbent can allow entry. The present discounted value of its profits gross of entry costs is then

$$\{R[Q(C^m), C^m] + \phi F_e\}/(r + \phi).^{19}$$

The incumbent is indifferent concerning whether to deter or not to deter entry if

$$\epsilon = \{R[Q(C^m), C^m] + \phi F_e\}/(r + \phi).$$

Now let the fraction of total costs that are not sunk be $\Phi (= F_e/\epsilon)$. Entry deterrence then occurs if

$$\Phi \leq 1 - (r/\phi) \cdot [R/\epsilon r - 1].$$

having variable costs C , then C is chosen to $\min \epsilon(C) + (CQ/r)$, which will, in the relevant cases, entail higher variable costs than those chosen by the entry-detering incumbent.

It may not, of course, pay the incumbent to deter entry—that is, it is possible that $\max R(Q, C) - N(C) > R[Q(C^d), C^d]/[r - \epsilon(C^d)]$, where C^d is the entry-detering technology, the solution to equation 8.

Later, in a slightly different context, I show that entry depends on whether there are an odd or an even number of firms. With an infinite number of potential entrants, I describe a mixed-strategy equilibrium. There is an equilibrium of this form here too: in each period there is a given probability that some firm will enter and displace the existing firm. Prices, meanwhile, only gradually get shaved down. Indeed, under the usual assumptions that an entrant can steal the entire market if it charges an amount just below the prevailing price, the price may remain near the monopoly price indefinitely.

19. If no potential entrant arrives until time T , the present discounted value of profits is $[R(1 - e^{-rT})/r] + e^{-rT}F_e$. The probability that the first potential entrant arrives at time T is $\phi e^{-\phi T}$.

Two results emerge. First, with small sunk costs (Φ near unity), so long as there is limited potential entry (ϕ is finite), price is set at the monopoly price; there is no limit pricing. One monopolist is succeeded by another, with infinitesimal decreases in prices. Second, as the number of potential entrants increases to infinity, the incumbent always engages in entry deterrence. Although this reaction results in what might be viewed as limit pricing, the price is not set at the competitive price and profits are not equal to zero. Only as entry costs (ϵ) go to zero will the competitive outcome be attained.

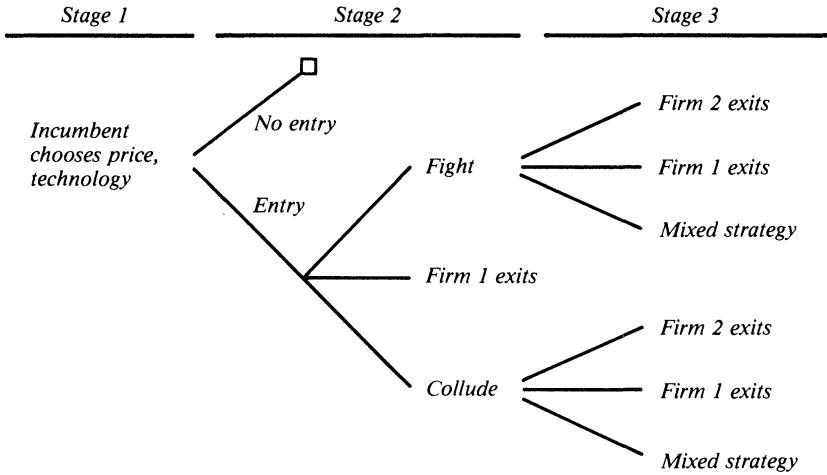
COLLUSIVE BEHAVIOR WITH EXIT COSTS

The previous sections showed how an entrant can drive the incumbent out of the market. Entry leads not to competition but to exit; or alternatively, the threat of entry leads to the choice of entry-detering technologies by the incumbent.

Rather than fighting or exiting, the incumbent can collude. And the entrant—again by choice of technology—can facilitate that collusion. Assume the incumbent has chosen a technology with high exit costs to persuade the potential entrant that it, the incumbent, will not leave. Now the entrant needs to persuade the incumbent that the incumbent is better off cooperating than fighting. The entrant can do this, as I shall now show by a simple extension of the model. Assume there are three periods, each divided into two parts. In the first part the firms must simultaneously make an exit-entry decision. In the second part, each having seen its rival's exit-entry decision, they simultaneously make a pricing decision.

The structure of the argument is simple. If the entrant enters, there are three equilibria in the third period: the incumbent leaves, the entrant leaves, or there is a mixed-strategy equilibrium. The entrant promises to leave in the third period—leaving the incumbent to enjoy the monopoly profit—if the incumbent shares the market in an appropriate way in the second period. In other words, the entrant bribes the incumbent to collude by holding out the prospect of exiting in the third period. If the incumbent does not cooperate, the entrant announces a mixed strategy, which entails some probability of not leaving. The promise of leaving if the incumbent cooperates is credible since the entrant's profits upon leaving are at least as great as if it stays in. And the threat not to leave (to play the mixed strategy) if the incumbent does not cooperate is also

Figure 4. Collusive Equilibrium



credible, since if the entrant does not leave, the Nash equilibrium entails playing the mixed strategy. Given these credible promises and threats, it pays the incumbent to cooperate. Figure 4 depicts the equilibrium.

I now investigate in greater detail the circumstances in which this collusive outcome can be sustained as part of a noncooperative equilibrium.

Assume that there is a time discount factor of $\delta = 1/(1 + r)$ and that monopoly profits, π_m , are growing at the rate g . The incumbent firm receives F_I upon departure and the entrant receives F_e upon departure (F_I and F_e can be positive or negative). The entrant announces that if the incumbent charges the monopoly price in the second period and lets the entrant have a fraction $(1 - \alpha)$ of the market, the entrant will leave in the third period. If the incumbent does not cooperate, the entrant will play the noncooperative-equilibrium mixed strategy the third period. This entails an entrant's leaving a fraction, β_e , of the time. When the entrant does not leave and its rival also does not leave, the entrant charges a price equal to the marginal cost of production. When its rival does leave, the entrant charges the monopoly price. The incumbent similarly leaves a fraction, β_I , of the time, and when it does not leave, it follows the same pricing policy as the entrant. Profits from leaving are F_i ($i = I, e$); expected profits from staying are, for the entrant, $\beta_I \pi_m$.

Thus β_I is chosen to make the entrant indifferent between staying and leaving:

$$(9) \quad \beta_I = F_e / \pi_m.$$

And β_e is chosen to make the incumbent indifferent between staying and leaving:

$$(10) \quad \beta_e = F_I / \pi_m.$$

The incumbent also announces that if the entrant cooperates in the second period, the incumbent will remain in the third; if the entrant does not cooperate, the incumbent will pursue the mixed strategy described above.

Viability of cooperation. First, one must check to see whether the entrant's promise to leave if the incumbent cooperates is credible. Assume that there has been cooperation in the second period, in which case the incumbent will not exit. The optimal action of the entrant is then to exit, provided only that $F_e \geq 0$. (If the entrant does not exit, its third-period profits are zero; whenever both firms are in the market, the Bertrand equilibrium prevails.) Similarly, if the entrant always exits, the incumbent is better off to stay in, provided only that $F_I \leq \pi_m$.

Next, I show that the threat to play the mixed strategy, if there has not been cooperation, is credible. The mixed strategy described above is a Nash equilibrium, and given that each firm believes the other is going to play the mixed strategy (if there is no cooperation in the second period), the rival will be better off playing the mixed strategy.

Consider now the second period. Assume the entrant has entered. If the incumbent firm cooperates, its profits must be greater than or equal to the profits if it does not, that is,

$$(11) \quad \alpha\pi_m + \delta g\pi_m \geq \pi_m + \delta F_I.$$

If the incumbent decides to cheat, it simply undercuts the entrant in the second period, stealing the entire monopoly profit; the incumbent then exits, obtaining F_I the third period.

Similarly, it must not pay the entrant to cheat, that is,

$$(12) \quad (1 - \alpha)\pi_m + \delta F_e \geq \pi_m + \delta F_e.$$

It immediately follows that if the entrant is not to cheat, $\alpha = 0$. But then

for the incumbent to cooperate, we require (rewriting equation 11) $F_I \leq (\delta g - 1)\pi_m/\delta$.

Note that once again, entry does not make consumers better off: the price is still the monopoly price. All that happens is that the monopoly profits are shared, and real social costs are incurred.

In appendix B to this paper, I show that if the number of periods gets extended, the range of cooperative equilibrium (sharing rules) for the second period is increased.²⁰

Entry deterrence. I have thus provided conditions under which, if entry occurs, there will be a collusive equilibrium. Entry will thus occur if the above condition is satisfied and if the entry costs, ϵ , are sufficiently small, so that $\epsilon \leq \pi_m + \delta F_e$. Thus to deter entry, the incumbent firm must choose a technology for which cooperation does not pay, one for which F_I is high.²¹

If $\delta = 1$ and $g = 1$, then for there not to be a cooperative equilibrium the incumbent need only set $F_I > 0$. But doing this ensures entry with some probability.²² Thus in this case there are but three possible

20. There is, in fact, a general theorem that as the number of periods becomes indefinitely large, if the discount factor is 1, then essentially all cooperative equilibria can be attained. The Folk theorem can be extended to finite period games if there are multiple equilibria, as here. See Jean-Pierre Benoit and Vijay Krishna, "Finitely Repeated Games," *Econometrica*, vol. 53 (July 1985), pp. 905–22.

Experimental evidence suggests the collusive behavior may be easier to achieve than suggested by the game theoretic analysis. The particular game theoretic formulation I have developed here, however, has been criticized in that future behavior is affected by past actions that do not affect the current state other than through the announced strategies. On the principle that bygones should be bygones, it is argued that whatever is a reasonable equilibrium beginning at date t should depend only on state variables at time t , or beliefs about state variables.

21. Though no cooperative equilibrium exists, several other perfect equilibria may, some of which entail entry and some of which do not. For instance, the incumbent firm announces its strategy that, in any period in which an entrant arrives, it will charge the competitive price, and in the following periods will play the mixed strategy described earlier. There then exists a perfect equilibrium with no entry, as long as $\epsilon \geq \delta F_e$, that is, as long as there are some nonrecoverable (in terms of discounted value) sunk costs. In this equilibrium the incumbent charges the monopoly price: the threat of entry is ineffective.

22. There is a perfect equilibrium for which the entrant enters, the incumbent leaves, and prices are set at the monopoly level. There is another perfect equilibrium in which the entrant enters with some probability and the incumbent leaves with some probability. When they are both in the market, prices are Bertrand prices; when only one is in the market, prices are set at the monopoly level. The probabilities of entry and exit are chosen to ensure zero expected profits for the incumbent and expected profits equal to the entry costs for the entrant.

outcomes: with F_I negative, entry occurs and there is collusion; the threat of entry does not act as a discipline on current price. If F_I is positive, entry *and* exit may occur, with price remaining at the monopoly price; or, alternatively, occasional price wars may break out. Either equilibrium may occur, regardless of the size of entry and sunk costs—as sunk costs go to zero, prices do not necessarily converge to the competitive level. Even if there are price wars, the effects on expected price of reducing sunk costs are ambiguous. As sunk costs go to zero as a fraction of total entry costs, there is still a finite fraction of the time that prices in the second period will be at the monopoly level. And even when prices in the second period are at the competitive level, there is a finite probability that prices the third period will be at the monopoly level.²³

ENDOGENOUS SUNK COSTS

Sunk costs (entry and exit costs) affect the nature of both *ex ante* and *ex post* competition. Since both entrants and the incumbent know this, they naturally take it into account when making investment decisions. Put another way, sunk costs represent a commitment. In this section I consider a simple example in which the incumbent firm has a choice of alternative technologies, some of which entail sunk costs.

I wish to make two points. First, much of the literature on contestability seems to depend on the existence of technologies with no sunk costs. I show that in fact more is required: technically efficient technologies cannot exist without sunk costs. Second, to deter entry the incumbent firm may sink more costs than is economically efficient; thus again, potential competition may result in a Pareto inferior equilibrium.²⁴

Consider a situation in which a firm has available to it two technologies, one involving sunk costs and the other not. The natural assumption is that by making investments nonconvertible (sunk) a firm gains something in variable costs. In the absence of uncertainty firms will thus choose

23. Indeed, as sunk costs are decreased, if entry occurs, it becomes increasingly likely that one of the two firms, and possibly both, will exit in the third period.

24. The analysis abstracts from uncertainty. The presence of uncertainty concerning future demands and technologies provides an additional reason why it may not be economically efficient (from a social point of view) to sink costs. As a method of deterring entry, sinking costs may have additional costs besides those discussed here.

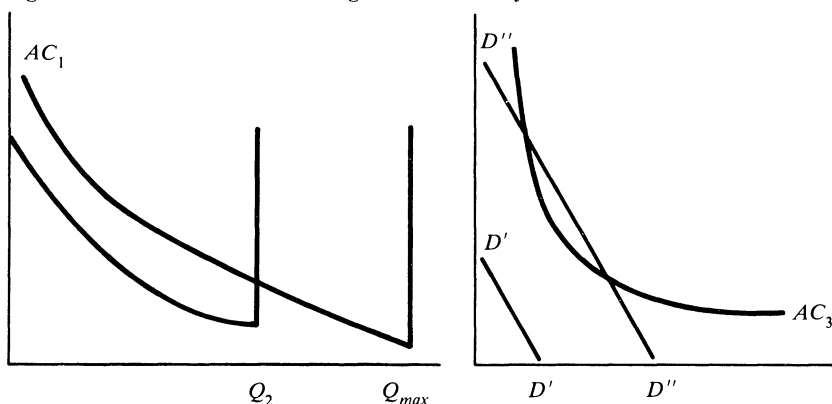
the more efficient sunk-cost technology (if output is large enough). So long as there is any cost associated with making capital usable at other locations or in other industries, the appropriate equilibrium will entail sunk costs. The availability of a technology with no sunk costs, an assumption probably belied by fact, does have one important implication: it limits the extent to which a firm can exercise monopoly power. It cannot charge a price higher than that at which a firm with no sunk costs could enter and make a profit. But this in turn has an important implication: the incumbent firm may choose a technology that can deter entry even of firms with no sunk costs. Thus, assume that the firm had available to itself two technologies. One entails a higher sunk cost but zero marginal cost of production up to some large scale, Q_{max} (see figure 5). The other entails a smaller sunk cost but positive marginal cost up to some output, Q_2 . Assume that at the monopoly level of output the second technology is more efficient than the first. (This is assumed true even at the monopoly output for the technology with zero marginal cost.) Consider now a third technology that entails no sunk costs, the contestability technology. A firm will enter using this technology, provided it can make a profit or at least break even. If the incumbent chooses the first technology, it would, in the Bertrand equilibrium, dump on the market Q_{max} if an entrant entered the market, while if it chooses the second technology, it would dump on the market Q_2 . Denote the residual demand curves facing the potential entrant in two situations by $D'D'$ and $D''D''$ respectively, and assume that $D''D''$ intersects the average cost curve, while the average cost curve is everywhere above $D'D'$ (see figure 5). Thus by choosing the inefficient technology (with large sunk costs), the firm is able to deter entry even of firms with no sunk costs. Whether this is more profitable depends on a comparison of monopoly profits with inefficient technology and profits with the equilibrium prices in the contestable equilibrium with the second technology. Either is possible.²⁵

When firms resort to inefficient entry deterrence devices, they will

25. The equilibrium price is the solution to $D(Q_2 + Q_c) = A(Q_c)$, where $A(Q)$ is the average cost curve of the contestable technology and Q_c is the output of the firm employing that technology. Thus the profits of the incumbent over two periods (ignoring discounting) with technology 2 are $R[Q(C_2), C_2] + [A(Q_c) - C_2]Q_2 - \epsilon_2$, and profits with the entry-detering technology are $2R[Q(C_1), C_1] - \epsilon_1$, where C_i is marginal costs with technology i , ϵ_i is sunk costs, $R(Q, C_i)$ is operating profits from a monopolist operating technology i with variable costs C_i at output Q , and $Q(C)$ is the profit-maximizing level of output.

This, I take it, is a correct version of the argument for excess capacity as an entry

Figure 5. Use of Inefficient Technologies to Deter Entry



naturally ask themselves whether other strategic policies would enable them to deter entry as effectively, and which ones would cost the least to implement. One could imagine a large variety of contracts that a firm could sign with third parties, contracts saying that the firm will pay the third party a large fixed amount on the condition that the third party does not retaliate in the event of entry. Such contracts make fierce competition

deterrent put forward by Michael Spence, who did not distinguish between sunk and fixed costs and did not formulate a strategic equilibrium model. Avinash Dixit correctly criticized him on the grounds that, with a Cournot quantity setting postentry equilibrium in which capacity changes only the maximum output, not the marginal cost of production, the excess capacity would not be used after entry and thus would not constitute an effective entry. See Avinash Dixit, "A Model of Duopoly Suggesting a Theory of Entry Barriers," *Bell Journal of Economics*, vol. 10 (Spring 1979), pp. 20–32. But the Cournot quantity setting postentry equilibrium is not the only possible equilibrium. Indeed, in the presence of excess capacity, there is likely to be price competition. Hence I have modeled the postentry equilibrium as entailing price competition. However, if there are small sunk costs with inelastic demand, entry deterrence can occur with arbitrarily small excess capacity. The potential entrant firm knows that if it enters, price will be bid down to the marginal cost of production, and the firm will be unable to recuperate its sunk costs. This case, in which there is some technology with no sunk costs, is one in which significant excess capacity may be created to deter entry.

More generally, even with Cournot equilibrium, if an increase in capacity decreases short-run marginal costs at all outputs, then there is a gain from excess capacity (that is, from choosing a capital stock larger than which minimizes the firm's cost of producing its equilibrium output). See Marius Schwartz and Michael Baumann, "Entry-Deterrence Externalities and Relative Firm Size," *International Journal of Industrial Organization* (forthcoming).

in a postentry game credible and thus serve to deter entry. And since, under the nonstochastic assumptions of my analysis, entry never occurs, these contracts have no costs associated with them. However, there is an incentive for the third party and the incumbent to renegotiate the contract in the second period should entry occur, for the third party will realize that unless it renegotiates, it will receive nothing, while if it renegotiates, it can at least get something, and the original incumbent can gain from renegotiation as well. Knowing this, the entrant may not really believe such contracts are credible.

There are two alternative ways of interpreting why one does not see such contracts. Sunk costs may be sufficiently important that they serve as a costless entry barrier, and hence these contracts are unnecessary. There may also be important stochastic elements.

The general point remains: the incumbent firm's choice of technology will be affected by the threat of entry. It may choose a technology that will deter entry. It may choose a technology that will make another firm's entry unprofitable if the two firms compete vigorously. The incumbent firm may choose a technology that will make the entrant believe that it will not exit. Or it may choose a technology that will make the entrant believe it will not cooperate. The incumbent will choose to sink costs: the contestability doctrine requires that no technology with sunk costs be available, a clearly inadmissible assumption. And the technology chosen in response to the threat of competition will not, in general, be the efficient technology.

RENT DISSIPATION

A general lesson that emerges from the preceding sections is that potential competition does not suffice to ensure zero profits.

The first firm in an industry has an advantage that successors cannot completely eliminate. Not surprisingly, there may be a race to become the first incumbent firm in an industry, and this race may or may not completely dissipate the potential rents.

If there are several competitors, all equally well situated, then the race to become the first firm can be viewed as if it were a patent race, and in some cases this situation does result in rent dissipation.²⁶ It may

26. Drew Fudenberg and others, "Preemption, Leapfrogging, and Competition in Patent Races," *European Economic Review*, vol. 22 (June 1983), pp. 3-31.

even pay firms to obtain patents on products that are not currently commercially viable. In such circumstances the early research is also not socially desirable and should be postponed until the product is commercially viable. The early research is a form of rent dissipation.²⁷ In instances in which there is learning by doing, it may pay to begin producing while price is below marginal cost (earning a negative profit flow).²⁸

However, firms are seldom in identical positions to begin with, and the firm that is in an advantageous position may forestall its competitors with only limited expenditures on R&D, as we shall see later. First, however, a few objections that have been raised to my basic results need to be addressed.

Objections

The previous section argued that even with small sunk costs potential competition would not suffice either to ensure economic efficiency or to ensure that the gains from efficient production would be passed along to consumers—that is, that the profits would be driven to zero. Moreover, even entry itself may not ensure effective competition, since the incumbent firm may accommodate itself to the entry and collusion may result.

Advocates of the view that potential competition will ensure economic efficiency and zero profits have raised objections to this analysis, which fall into two categories: that I have assumed that competition is too fierce or that I have not taken account of all the possibilities of competition (that is, that I have assumed competition is insufficiently fierce).

If competition is less fierce than Bertrand competition with homogeneous products, then after entry, firms may profit and these profits may induce further entry. By the same token, of course, neither potential nor limited actual competition suffices to ensure zero profits or economic efficiency. There *may* be grounds for government intervention.

There are three forms that limited competition may take: lags in adjustment, imperfect substitutability, and Cournot (rather than Bertrand) responses.

27. Partha Dasgupta, Richard J. Gilbert, and Joseph E. Stiglitz, "Invention and Innovation under Alternative Market Structures: The Case of Natural Resources," *Review of Economic Studies*, vol. 69 (October 1982), pp. 567–82.

28. Partha Dasgupta and Joseph E. Stiglitz, "Learning by Doing, Market Structure, and Industrial and Trade Policies," *Oxford Economic Papers* (forthcoming).

LAGS

If the incumbent firm cannot immediately lower its price upon the arrival of an entrant, then potential competition—the threat of entry—may affect the firm's pricing decision. If the firm charges too high a price, an entrant can come in and steal customers before it can respond. Because competition after entry is limited, there are incentives to enter; this is what enables potential competition to exert its disciplinary role.

The presence of these response lags is in itself evidence of sunk costs. If firms cannot change their prices for a fixed period of time, then the pricing decision is like a sunk, temporarily irreversible investment decision. Is it plausible that the adjustment costs associated with pricing decisions are larger than the adjustment costs associated with production decisions, that it takes longer to alter the former than the latter? In the one celebrated example in which sunk costs are alleged to be relatively unimportant—the airline industry—prices typically fall very quickly after entry and rise quickly after exit. Obviously, the airlines (and potential entrants) do not consider that preentry prices need prevail after entry. Indeed, it has been argued that when the existing airlines failed to respond by lowering prices, they did so solely because of fear of antitrust prosecution. Thus it is government policy that restricts firms from responding. But one of the objectives of this study is to ask, what should government policies be toward such price responses?

Since I believe that response lags in pricing decisions, particularly concerning entry and exit, are typically much shorter than those in production, I have focused attention on models that reflect those judgments concerning relative adjustment speeds. But in this section I explore the consequences of long price responses. These limitations in response lags are supposedly one way of limiting ex post competition (and thus increasing ex ante competition), but they are both an implausible and an ineffective way of limiting ex post competition.

To examine the effects of price lags, I modify slightly the model presented earlier to allow for the fact that firms may not be able to respond instantaneously to entry by lowering their price. There are two alternative versions of the rigid-price hypothesis: in the version explored next a firm must announce any change in prices L periods in advance but can vary its price as frequently as it wishes.²⁹ In the second version,

29. The first announcement cannot be made until $T = 1$.

explored later, a firm is allowed to vary its price only every L periods. In the first version there are two possible patterns of equilibrium: entry accommodation and entry deterrence. It is easiest if we reformulate our model in continuous time. The incumbent has a secure monopoly position from time zero to time 1; entry occurs, if ever, at or after time 1.³⁰

Entry deterrence. After L periods the market degenerates to a conventional Bertrand equilibrium, with price equal to marginal cost. Thus the entrant makes profits only during the L periods when the incumbent's price is fixed at q_1 . Let $\pi(q)$ be the profits (ignoring sunk costs) associated with charging price q , when a rival charges a higher price or when there is no rival. Then the present discounted value of the entrant's profits is

$$[\pi(q) \cdot (1 - e^{-rL})]/r = V(q, L, r).$$

If the incumbent sets its price so that $V(q, L, r) = \epsilon$, entry will be deterred. Hence

$$q_1 = \min(q_m, q),$$

where q_m is the monopoly price, that is, $\pi'(q_m) = 0$. For each value of $q_1 = \min(q_m, q)L$, there is a critical value of ϵ , which is ϵ^* , such that for $\epsilon > \epsilon^*$, $q_1 = q_m$. This is just the set of (ϵ, L) satisfying $V(q_m, L, r) = \epsilon$.

For large ϵ the threat of competition has absolutely no effect on price. For $\epsilon < \epsilon^*$ the threat of competition lowers the price below the monopoly price but not to the competitive level (marginal cost of production). Only as L approaches infinity does the price approach that at which the entrant firm breaks even were it to be the sole producer forever. (Thus in region II in figure 6 the threat of competition has absolutely no effect on the incumbent; in region I it has some effect but a limited one. Note that for short reaction lags the incumbent may be able to act as a monopolist even with small entry costs.)

The threat of competition is also ineffective in eliminating profits. If $\pi_1(q) \equiv \pi_2(q) = \pi(q)$, net profits for firm 1, the incumbent, are

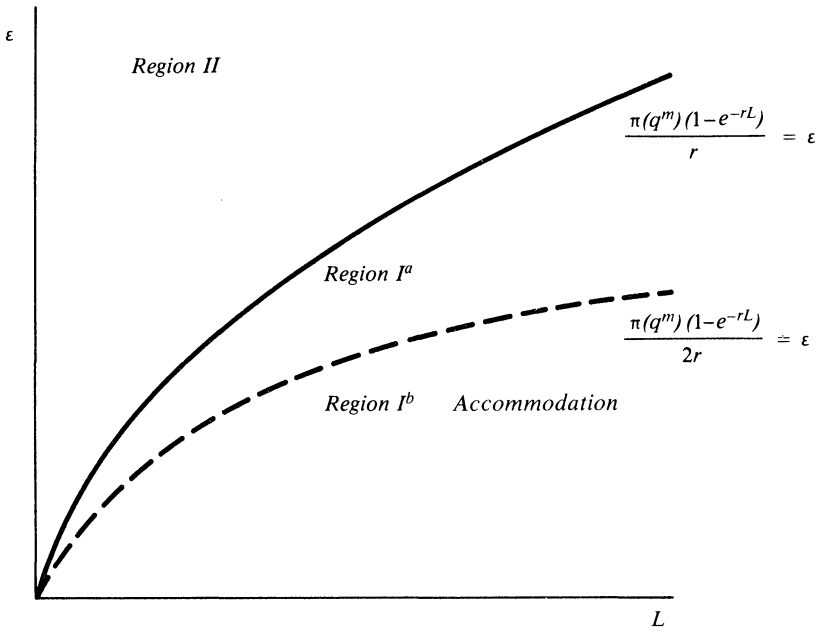
$$\pi(q/r - \epsilon_1) = \epsilon_2/(1 - e^{-rL}) - \epsilon_1,$$

where ϵ_i is the i th firm's entry costs. If $\epsilon_1 = \epsilon_2 = \epsilon$, the present discounted value of profits is $\epsilon e^{-rL}/(1 - e^{-rL}) \rightarrow 0$ only as $\epsilon \rightarrow 0$ or as $rL \rightarrow \infty$.

Entry accommodation. Rather than attempt to deter entry, the in-

30. For simplicity, the incumbent is allowed to announce a price reduction only after the entrant actually enters; obviously, to the extent that the incumbent can anticipate the date of entry, the effective response lag is reduced.

Figure 6. Response Lags and Entry



cumbent can decide to accommodate it. The firm announces a strategy of the following form: it will charge price q_m so long as the entrant does so; if the entrant charges a lower price, the incumbent will revert to the Bertrand strategy of undercutting its rival. A perfect equilibrium exists when the entrant announces a similar strategy. The present discounted value of the profits if both cooperate is $\pi(q_m)/2r$. Since, when both use the Bertrand strategy, profits are zero, both firms always benefit from cooperation.

Whether the incumbent firm chooses to accommodate or deter entry then depends on the relative value of profits in the two situations, that is, on whether $\epsilon_2/(1 - e^{-rL})$ is greater or less than $\pi(q_m)/2r$. The locus of (ϵ_2, L) at which the firm is simply indifferent is the dotted line in figure 6. Thus for very low values of entry costs, the price jumps back to the monopoly price. The firm accommodates the potential competitor, and the two firms act collusively to divide the market.³¹

31. When there are many potential entrants, the perfect equilibrium may entail no entry when entry can occur at any time. Assume a firm enters at a particular date and cannot change its price for L periods thereafter. Another potential entrant would then find

An alternative interpretation of price lags. In the alternative interpretation of reaction lags, the incumbent firm can change its price only every L periods. Considerations of symmetry dictate that the entrant too can change its price only every L periods. The entrant must then decide when to enter. In making its decision it needs to take into account the fact that how the incumbent will respond will depend on the entrant's choice of price as well as the date at which the firm enters. The incumbent has a choice of either fighting or cooperating, that is, dividing the market with the entrant.

Assume the entrant enters L^* periods after the incumbent has set a price. I first show that accommodation then is not possible. Assume the incumbent fights by lowering its price just slightly below that of the entrant firm, which responds by lowering its price just below that of the incumbent when its turn for adjusting its price comes. The present discounted value of the incumbent's profits is then

$$(13) \quad \pi(q^*) (1 - e^{-rL^*})/r(1 - e^{-rL}).$$

If the incumbent firm cooperates by charging a price just equal to that of the entrant, the present discounted value of its profits is

$$(14) \quad \pi(q^*)/2r.$$

It must set (L^*, q^*) to induce cooperation rather than competition. Moreover, when the entrant has the opportunity to reset its price, cooperation must be more beneficial than fighting. The present discounted value of profits to the entrant if it fights is then

$$(15) \quad \pi(q^*)[1 - e^{-r(L-L^*)}]/r(1 - e^{-rL}),$$

while if it cooperates, they are given by equation 14. Thus to sustain cooperation,

$$0.5(1 + k) \leq \Gamma \leq 2k/(1 + k)$$

is required, where $k = e^{-rL}$ and $\Gamma = e^{-rL^*}$. But this inequality cannot be satisfied because $k \leq 1$. Hence long-run cooperation is not viable. It thus pays the entrant to enter just after the incumbent firm has set its price, and on each occasion that the entrant has to reset its price, to reset the

it beneficial to enter a second later, charge an ϵ lower price, and steal the market. Thus the first entrant will have profits too small to cover its fixed cost of entry, and again one sees the seemingly paradoxical result that more competition results in less competition.

price just below the incumbent's. Thus even with price lags the market is characterized at any moment by a dominant seller. And though prices fall, the lags in adjustment imply that they fall very slowly. For all practical purposes, prices remain near the monopoly price indefinitely.

Plausibility of price reaction lags. Though I have argued that the presence of price reaction lags is not enough to restore the validity of the contestability doctrine, I would argue further that significant price reaction lags are not very plausible. (By contrast, production lags may plausibly be significant.) Indeed, it is easy for the incumbent firm to announce a policy of meeting the competition, charging a price just below that of the competitor, down to a price equal to the marginal cost of production. Firms can and do build these policies into sales contracts. Such a strategy ensures there will be no entry. And the strategy is credible, for given that the rival has announced a price in excess of marginal cost, the incumbent has a choice of either losing the sale or meeting the competition. It clearly pays the incumbent to meet the competition. (Again there is the seeming paradox of an apparent increase in competition actually resulting in a reduction in effective competition.)³²

The issue, of course, is not whether there are or are not lags in pricing decisions: there undoubtedly are. The issue is, what is the length of these lags relative to the lags in entry, exit, and other production decisions? I would argue that the pricing lags are relatively short, and that this idealization provides a better description of the market than the other idealization, that they are of the same order of magnitude as exit, entry, and production decisions.

IMPERFECT SUBSTITUTES

The analysis so far, like the analyses in much of the contestability literature, has focused on firms producing perfect substitutes. In such situations Bertrand competition drives prices to marginal cost, and it is

32. The consequences of meeting-the-competition (or most favored nation) strategies were earlier noted by Steven C. Salop, "Practices That (Credibly) Facilitate Oligopoly Coordination," in Joseph E. Stiglitz and G. Frank Mathewson, eds., *New Developments in the Analysis of Market Structure* (MIT Press, 1986), pp. 265-90. They have been further analyzed by Thomas E. Cooper, "Facilitating Practices and Most-Favored Customer Pricing" (Ph.D. dissertation, Princeton University, 1984).

this that serves to deter entry. But many, if not most, markets are characterized by firms producing products that are imperfect substitutes.³³ Imperfect substitutes raise a number of interesting issues.

Imperfect substitutes and entry deterrence. One consequence of imperfect substitutability is that ex post competition will be less fierce and hence entry will occur more easily. Assume, for instance, the family of demand functions derived from the indirect utility function of the form $a(q_2/q_1)I/q_1$, where I is the individual's income and $a'q_2/qa_1 = 0.5$ at $q_2 = q_1$. Then it can be shown that in the postentry game the equilibrium price will be

$$q_1 = q_2 = C(\sigma + 1)/(\sigma - 1),$$

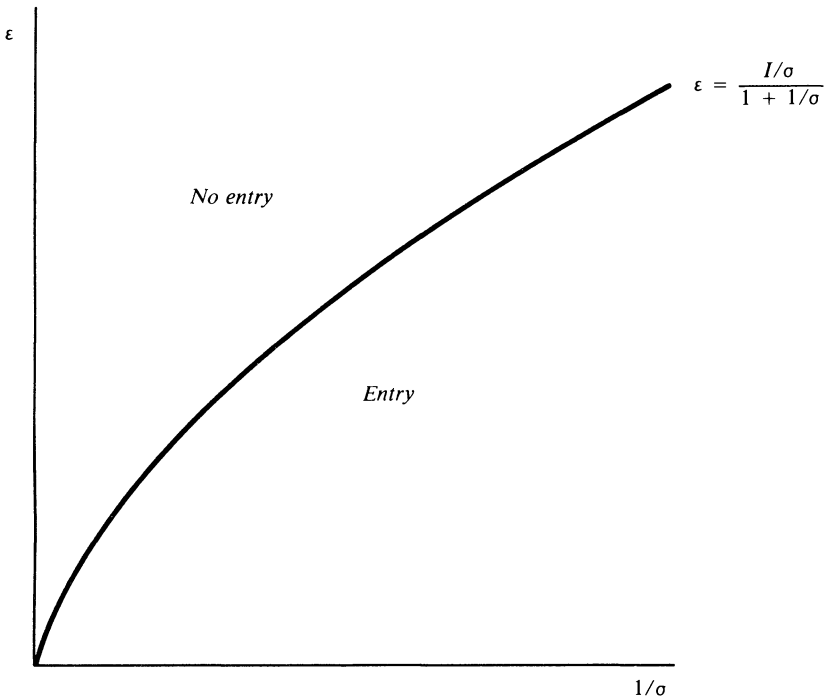
where σ is the elasticity of substitution between the two commodities. It is clear that price exceeds marginal cost of production, provided σ is less than infinite. Figure 7 shows the critical value of ϵ , above which entry will not occur, as a function of σ .

Multiproduct firms and Ramsey prices. Once one shifts attention from markets for homogeneous products to markets with related commodities, a whole range of market structures needs to be investigated. One firm could produce all products within the industry. A second firm could then enter by competing against only one product or a whole range of products. Such questions are among those that face managers responsible for corporate strategy, both those considering entering a market and those attempting to protect themselves against entry.

One reason for product differentiation is the fixed cost associated with producing any product. Since nonconvexities are at the heart of the analysis of markets with differentiated products, contestability doctrine would seem to be particularly useful for analyzing how such markets function. And advocates of the contestability doctrine have indeed considered such industries. They have argued that even when the

33. To the extent that this is true, the appropriate model of the economy, of course, is one of imperfect or monopolistic competition, not the perfect, competitive Arrow-Debreu model. There are important differences between monopolistically competitive economies and perfectly competitive economies, not the least of which is that the fundamental theorems of welfare economics do not, in general, apply to them. For a recent survey of these models, see Joseph E. Stiglitz, "Towards a More General Theory of Monopolistic Competition," in M. H. Preston and Richard E. Quandt, eds., *Prices, Competition, and Equilibrium* (Deddington, England: Philip Allan, 1986), pp. 22-69.

Figure 7. Entry and Imperfect Substitutability



technology is such that only one firm produces, contestability ensures that it operates with zero profits, that it raises the revenue to pay for the fixed costs by charging Ramsey prices, and that accordingly it behaves exactly as a government enterprise that was constrained to break even would behave.³⁴ This argument, while seeming persuasive, is in fact inadequate. For if entry of competitors is permitted, there are important constraints on the prices such an incumbent firm can charge, constraints

34. Frank Ramsey, in his classic contribution, analyzed how the government should raise revenue by excise taxes in such a way as to minimize deadweight loss. In the case of separable demand functions with horizontal supply functions, tax rates were inversely related to the elasticity of demand. Under the stipulated conditions on demand and technology, Ramsey prices thus entail prices in excess of marginal cost, with the percentage deviation being inversely related to the elasticity of demand. For a more extensive discussion, see, for instance, Joseph E. Stiglitz, "The Theory of Pareto-Efficient and Optimal Redistributive Taxation," in Alan J. Auerbach and Martin Feldstein, eds., *Handbook of Public Economics*, vol. 2 (North-Holland, 1987), chap. 15; or Anthony B. Atkinson and Joseph E. Stiglitz, *Lectures on Public Economics* (McGraw-Hill, 1980).

that a government-run monopoly charging Ramsey prices would not face. And if entry of competitors is not permitted, in the absence of regulation the discriminatory pricing policies the monopolist would engage in to extract as much consumer surplus as possible would result in pricing policies far different from those the government would have used.³⁵

Exit. Earlier I showed how the possibility of exit affected entry. A firm might enter believing it could drive out the incumbent. With differentiated products there are greater incentives for the incumbent firm to stop producing one of the products it is producing. Thus there are greater incentives for entry. But entry does not result in competitive prices (prices equaling marginal costs). Worse still, the threat of entry induces the incumbent to take entry-detering actions that reduce social welfare.

Assume two symmetric products with a given elasticity of substitution between them.³⁶ Assume further that the incumbent is producing both commodities and that exit is impossible. Then, as my earlier analysis established, if there are ϵ sunk costs, there will be no entry (if there is Bertrand competition). Assume now that the incumbent firm receives

35. For a fuller discussion of this, see Sappington and Stiglitz, "Information and Regulation." These also note the further problems, both theoretical and practical, caused by the definition of the boundaries of the industry in the presence of pervasive interdependencies. If a broad definition is employed, the view of the market economy as a decentralized allocative mechanism is vitiated. If a narrow definition is employed, the set of commodities, taxes against which can be used to provide the revenues required to finance the fixed costs, is unnecessarily circumscribed.

It should also have been clear that if firms could use nonlinear prices, even if limited to charging a fixed service fee, the optimal set of prices would bear little relation to Ramsey-Boiteaux prices. See Anthony B. Atkinson and Joseph E. Stiglitz, "The Design of Tax Structure: Direct versus Indirect Taxation," *Journal of Public Economics*, vol. 6 (July–August 1976), pp. 55–75.

The popularity of the view that with nonconvexities but free entry, market equilibrium would be Pareto efficient is surprising. Even before the publication of Baumol, Panzer, and Willig's contribution, *Contestable Markets*, Avinash K. Dixit and Joseph E. Stiglitz, "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, vol. 67 (June 1977), pp. 297–308, had shown that the zero profit condition ensured constrained Pareto optimality (that is, in which the government could not provide lump-sum subsidies to firms) only under extremely stringent conditions that were unlikely to be satisfied. (All firms had to face constant elasticity demand curves for their products, with the same elasticity.)

36. A variant of this model is presented in Kenneth Judd, "Credible Spatial Preemption," *Rand Journal of Economics*, vol. 16 (Summer 1985), pp. 153–66.

F_1 upon exit from one of the product lines. Let π^d be the profits if it exits from one of the two product lines, π^b if it does not.³⁷ The incumbent will then benefit from exiting from one of the two product lines if $\pi^d + F_1 > \pi^b$. But if this equation is satisfied, entry will occur provided only that $\epsilon < \pi^d$. The incumbent knows this and will thus have an incentive to choose technologies for which exit costs are large; these technologies serve as an entry deterrent. The firm may choose such a technology, even though costs of production are higher. The incumbent thus has the advantage of a first mover arising from its ability to commit itself not to leave.³⁸

NASH-QUANTITY EQUILIBRIA (COURNOT)

In the analysis so far I have assumed that the two firms play Bertrand strategies in the postentry game. Given the absence of any capacity constraints and given the constant-returns-to-scale technology, this seems reasonable. I investigate briefly, however, the consequences of the two firms playing quantity-setting strategies in the postentry game.

Cournot competition is less fierce than Bertrand competition. This means that postentry profits will be larger than under Bertrand competition, which in turn means that sunk costs will serve as a less effective barrier to entry.

Assume, for instance, that the industry demand curve is constant elasticity:

$$Q = kc^{\mu-1}p^{-\mu}.$$

Then postentry equilibrium price is given by

$$p = c/[1 - (1/2\mu)],$$

while postentry profits are given by

$$\pi = k[1 - (1/2\mu)]^{\mu-1}/2\mu.$$

For each value of μ , there is a critical value of ϵ above which entry will

37. The firm will not earn any profits on the product line in which there is direct (Bertrand) competition, and the lower price on that product line lowers profits on the other product line.

38. The incumbent may, however, face a dilemma. Earlier, I showed that a low value of F_1 (high exit costs) might induce an entrant firm to believe that should it enter, the incumbent firm would cooperate (since it would not benefit from leaving) to split the monopoly profits. There may not exist any value of F_1 that will deter entry.

not occur. The critical value is monotonic in μ (for relevant values of μ , that is, for $\mu \geq 0.5$), with ϵ approaching zero as the elasticity approaches infinity. For finite elasticities of demand, there is a finite level of sunk costs below which entry occurs. But the postentry price is not the competitive price: even actual competition is not enough to ensure that price equals average, let alone marginal, cost.

When there are a large number of potential competitors, profits net of entry costs are driven to zero; the equilibrium number of firms, n^* , solves $\epsilon = k[1 - (1/n\mu)]^{\mu-1}/n\mu$. Only as ϵ goes to zero do n go off to infinity and prices converge to marginal cost. Moreover, with free entry, though some gains from competition are passed on to consumers in the form of lower prices, some are also dissipated as excess entry expenditures. The total resources wasted as a result of competition are $(n^* - 1)\epsilon$. As ϵ goes to zero, this goes to k/μ .³⁹ As a fraction of first-best total expenditures on the commodity, this waste amounts to $1/\mu$; that is, for reasonable values of demand elasticity, the fraction wasted is nonnegligible. Though profits are zero, the economy is far from efficient.

The same question can be asked that was posed earlier in the case of Bertrand competition: What happens if firms are unable to change the level at which they produce for L periods? In the earlier case, I noted that if the incumbent responded to the threat of potential entry by attempting to deter it, potential competition would lower the price during the preentry period. If the incumbent responded to the threat of competition by accommodating it, potential competition would have no effect on preentry prices. Now, potential competition may increase prices during the preentry period if the incumbent seeks to accommodate entry. For if the firm anticipates entry, the optimal output during the postentry period will be lower than the firm's output when it was the single producer; it sets its preentry output at a level between the monopoly and duopoly levels.

CONTRACTS

While the preceding discussion considered reasons why ex post competition may be less fierce than envisioned earlier, so ex ante competition may be more fierce than envisioned. I now consider two reasons why it is argued that potential competition may be effective even

39. That is, $n\epsilon = k[1 - (1/n\mu)]^{\mu-1}/\mu$, which converges to k/μ .

with effective ex post competition. One criticism of the earlier results is that I have ignored an important dimension of competition—contracting—and the other is that implicitly I must have assumed only limited potential competition. A closer examination of both issues shows that potential competition may be less, not more, effective than suggested by the analysis in the preceding section, “Sunk Costs and Potential Competition.”

Those who claim that potential competition will ensure economic efficiency argue that, to the extent there are sunk production costs, competition occurs at the contracting stage. That is, firms compete for customers before costs are sunk. But even this activity involves some sunk costs. To establish an infrastructure for negotiating contracts involves learning something about the market, building some reputation, and so forth. Such sunk costs may not be large, but they are not absent.⁴⁰ And the argument of the preceding section established that only small sunk costs were required for potential competition to have limited efficacy.

In fact, economies in which long-term contracts can be written and enforced may act even less competitively than economies without contracts, provided there are some sunk costs associated with contracts. The fact that costs must be sunk can fragment a market, converting what would appear to be a large market with many suppliers and consumers into a large number of small markets, each of which has a limited number of suppliers. To put it another way, when there are contracts, the existing firm may have customers locked up, making successful entry difficult. I illustrate this by means of a highly stylized example in which equilibrium entails that almost all customers are locked up and potential competition is very ineffective.⁴¹

Time is discrete. At each date, N individuals are born. An individual lives for precisely T years. This is also the lifetime of a machine. A

40. In fact, of course, there may be other reasons that contract competition may be limited. In some R&D markets, in particular, it may be virtually impossible to design contracts fully specifying the nature of the new commodity to be delivered and the date and price of delivery.

41. As has been emphasized in Marius Schwartz, “The Nature and Scope of Contestability Theory,” *Oxford Economic Papers*, vol. 38, supplement (November 1986), pp. 37–57, there are other reasons that contracting may make it less, not more, likely that markets are effectively contestable. Contracts enable an incumbent firm to commit itself to competing fiercely should an entrant enter; they make it more credible that entry will not be profitable.

machine costs F to build and possesses the capacity to produce up to M units of a product at zero marginal cost. Each individual purchases precisely one unit of the product at each date during his lifetime. An individual's reservation price for the product is u (> 0) at each date. The example is simplified enormously if $M = N$. Then efficiency demands that one machine be constructed at each date, with the machine built at date t serving the generation born at t until both the machine and the cohort die (at $t + T$). Indeed, if T is large, there would be a large number of machines and firms, and one would be inclined to assume that the market acts competitively. This is not so if obtaining contracts involves sunk costs.

We are concerned here with the outcome of a market economy in which individuals cannot sign contracts with producers that are binding on their heirs. There are many potential producers. We suppose that the cost of signing a contract is ϵ (> 0). This, by hypothesis, is sunk. It is now easy to check whether there is a subgame-perfect equilibrium in which each firm charges the reservation price, u , and all individuals sign T -year contracts. Given that all producers—current and future—charge u , a representative individual loses nothing by committing himself to a T -year contract, and he gains $\epsilon(T - 1)$, which he would have had to incur had he entered the market at each date during his lifetime. Thus it is optimal for him to sign such a contract. Given that all existing individuals have signed such contracts, the only individuals in the market at any date are the newborn. If competition among firms is Bertrand competition, not more than one firm will enter at any date. Thus at each date a single firm enters, signs T -year contracts with all members of the latest cohort, and charges u for the product. The firm cannot recruit older cohorts because they have signed binding contracts with earlier suppliers.⁴² In spite of the large number of firms in the market, the equilibrium price is the monopoly price.

INCREASING THE NUMBER OF POTENTIAL COMPETITORS

In most of the models explored in the section on sunk costs and potential competition, I assumed only one potential competitor. Indeed,

42. I have not discussed the penalties associated with breaking a contract. Equilibrium contracts will entail penalty clauses at least strong enough to deter breaking the contract at any date. Since the individual knows that there will be no occasion, in equilibrium, for him to break the contract, he will not be reluctant to sign such a contract.

two competitors are all that is required with Bertrand competition to ensure that price falls to marginal costs of production. Hence most of the results would be unaffected by the presence of more competitors.

I have already noted that there are some circumstances, however, in which markets may be less competitive the greater the number of potential competitors. The reason is simple: a firm's incentive to risk entering a market to take away some of the profits of the incumbent firm will be diminished if it believes that success in doing so will encourage other firms to enter the market and take its profits away. Indeed, the firm might argue that if stealing customers away from the current incumbent is worthwhile, a later firm will also find it worthwhile to steal its customers away. I will illustrate this point with two examples.

Lags in response. The first example is a slight modification of the model developed earlier with lags in response. With many potential entrants, each will worry whether a subsequent entrant will steal the market from him. Assume, for instance, that there are two potential entrants. Each firm would know that if it entered first, it would be immediately followed by its rival, and it would be unable to recover its sunk costs. It would thus postpone entering until a period sufficiently close to the date at which the incumbent is free to vary its price, so that its rival would not find it worthwhile to enter.⁴³ If there are three potential entrants, however, one entrant enters immediately, for it knows that once there are only two left, neither will enter until just before the incumbent firm is once again free to vary its price. The argument generalizes: if there are an even number of potential entrants, then no entry occurs until shortly before the incumbent firm is free to vary its price. If there are an odd number, one firm enters immediately. In either case the equilibrium price remains at the monopoly level.⁴⁴

43. That is, if L is the interval between price changes for any firm, λ is the time of entry of the entrant. The entrant will then dominate the market over the interval $L - \lambda$. The time of entry is chosen to produce zero profits, that is, $\epsilon = \pi(q^m)(1 - e^{-r(L-\lambda)})/r(1 - e^{-rL})$.

44. These are not, of course, the only perfect equilibria. Consider, for instance, the following. Each firm announces that if an odd number of firms has entered, it will not enter, but if an even number (or zero) has entered, it will. Then one firm enters. The firm knows that once it enters, no one else will enter, since now an odd number of firms has entered. The strategies are perfect. Given the strategies announced by all other firms, if one firm should enter when an odd number of firms has entered, it will immediately be followed by further entry and will thus suffer a loss. (Conversely, any firm would do well to enter if there is an even number of firms, for it knows that as soon as it enters, no further entry

If there are an infinite number of potential entrants, a mixed-strategy equilibrium exists, such that the probability of some firm entering per unit time, say $\Omega(N)$, is a function of the number of firms already in the market (and the dates at which they entered, which then specifies when the incumbent firms will be free to vary their prices once again). The unit of time, Ω , is such that upon entry a firm's expected return is just enough to cover its sunk costs. Consider the limiting case where L^* is infinite. Then Ω is a constant and solves $\epsilon = \pi_m/(r + \Omega)$.

Imperfect information and entry deterrence. Though the models have so far assumed that entrants know perfectly what will happen if they enter, information is in fact imperfect. They can only guess at what the incumbent's costs are. But the reaction of the incumbent to the first firm that attempts to compete reveals considerable information, which is of value to the other potential entrants. If there were only a single potential entrant, entry would be beneficial. With many potential entrants, there is no entry.

Assume that whether the incumbent firm has a high or low marginal cost of production is unknown,⁴⁵ but that following entry, Bertrand competition will break out. For simplicity, assume that the potential competitors all have identical cost functions; there is an ϵ sunk cost and a constant marginal cost of c_e , up to some capacity level, Q_e . Thus if the incumbent's marginal cost is less than c_e , the entrant loses its sunk costs ϵ . When there is only one potential entrant, if the incumbent's marginal cost is $c_h > c_e$, price equals c_h , and the entrant makes a profit of $(c_h - c_e)Q_e/r - \epsilon$, where r is the interest rate.⁴⁶ Thus, provided there is a sufficient probability that the incumbent has high marginal costs, there will be entry. Note that in this model, entry will reveal information. With an unlimited number of potential competitors, as soon as it becomes evident that the incumbent's marginal cost is c_h , entry will occur until price is driven down to average cost. Hence the total return to the first entrant is just $(c_h - c_e)Q_e - \epsilon$. The existence of later potential entrants

will result.) There is another equilibrium in which no entry occurs; the equilibrium strategies are the same as described above, except odd is replaced by even and conversely.

There also exist mixed-strategy equilibria, as the discussion in the text will show.

45. Salop, "Strategic Entry Deterrence," has argued that under these circumstances, it may pay the incumbent to signal its marginal costs by the price it charges.

46. I have reverted to a discrete time model in which there is a one-period lag in the dissemination of information.

deters entry, and potential competition is thus less effective in disciplining the market.⁴⁷

Sunk Costs and Innovation

If, as I have argued, sunk costs are pervasive, and if, as seems the case in many contexts, *ex post* competition is, if not described by a Bertrand model, at least reasonably fierce, it is no wonder that firms do not like to compete on prices. Firms may not understand the multistage-game theoretic models I have formulated here, but they know the outcome of price wars is unattractive. They therefore seek to focus on those forms of competition for which there are potential profits: on the development of new products and processes. It may, indeed, be the ineffectiveness of price competition that directs managerial energies to R&D, as well as providing the resources to finance it.⁴⁸

In this section I extend the results of the sections on sunk costs to sectors of the economy in which innovation is important. This section is divided into two parts. First, I argue that those aspects of technology on which I focused earlier—increasing returns and sunk costs—characterize technological change. I then show that a small expenditure on sunk costs may enable an incumbent firm to maintain its monopoly position without profits being driven to zero and with a negligible effect on the pace of innovation. Potential competition may be almost totally ineffective.

WHY MARKETS WITH TECHNOLOGICAL CHANGE ARE GENERALLY IMPERFECTLY COMPETITIVE

The standard approach to technological change treats the production of change much like the production of any other good or service. There is a production function describing the relationship between inputs of conventional factors and output of, say, some good, $Q = F(K, L, A)$, in

47. The problem here is, of course, the inability of the first entrant firm to appropriate the returns to the information produced as a result of its action.

48. Recall the earlier discussion on the limitations on equity and credit markets that constrain the ability of firms to finance R&D by means other than retained earnings.

which a variable, A , represents the stock of knowledge just as K represents the stock of physical capital. There is then a production function relating increments to the state of knowledge to inputs—university professors, for example—used in the production of knowledge,

$$dA/dt = G(K_R, L_R, A_R, A),$$

where A_R represents the state of knowledge about how to produce knowledge.⁴⁹

There is a naive but fashionable approach that, noting the similarity between the modeling of the production of conventional goods and that of information, argues that the “market” for knowledge works just like the market for conventional goods. This argument is wrong in several important respects. First, technological change is characterized by nonconvexities in its production.⁵⁰ Second, if technology exhibits, say, constant returns to scale at a fixed technology (the function F is a constant return to scale in K and L), it exhibits increasing returns to scale when inputs devoted to R&D are included.⁵¹ From the perspective of the firm, this means that the value of information may increase with the scale of production. A discovery that reduces the cost of producing widgets by \$1.00 is worth \$1,000 a year if 1,000 widgets are produced each year, but \$10,000 a year if 10,000 widgets are produced each year. The consequence of these nonconvexities is that markets with and for technological change are inherently imperfectly competitive. They are, in a sense, natural monopolies.⁵²

49. The importance of the state of knowledge concerning learning, or learning to learn, has recently been stressed by Joseph E. Stiglitz, “Learning to Learn, Localized Learning and Technological Process,” in Partha Dasgupta and Paul Stoneman, eds., *Economic Policy and Technological Performance* (Cambridge University Press, 1987), pp. 125–53. Obviously, we now need an equation describing increments in A_R . For simplicity, in the remainder of this paper A_R is assumed to be fixed and is suppressed in the representation of the function G .

50. R. Radner and Joseph E. Stiglitz, “A Nonconcavity in the Value of Information,” in Marcel Boyer and Richard Khilstrom, eds., *Bayesian Models in Economic Theory* (Elsevier, 1984), pp. 33–52.

51. With even slight increasing returns there will be a natural monopolist. The section on sunk costs and potential competition showed that with even small sunk costs such a natural monopolist may be able to charge monopoly prices.

52. There are several other respects in which technological change differs from conventional commodities. The production of technological change is essentially the

I argued earlier that the nature of equilibrium with a natural monopoly—and in particular the discipline provided by potential competition—depends on the existence of sunk costs. Most expenditures on R&D are, by their very nature, sunk costs. The resources spent on a scientist to do research cannot be recovered. Once his time is spent, it is spent.⁵³ And I showed that natural monopolies, with even small sunk costs, behaved in a far from competitive manner. I also argued that the recognition of the effect of sunk costs on the nature of ex post competition resulted in firm's sinking costs or incurring more sunk costs than they otherwise might. Sunk costs, I argued, were endogenous. This provides a further incentive for R&D expenditures because such expenditures are sunk costs that may serve to deter entry. Later in this section, I will show that the incumbent, by sinking a relatively small amount in a research program, can deter potential rivals from even entering the R&D competition.

Without using modern vocabulary, Schumpeter long ago argued for the importance of R&D competition. He did not have an explicit welfare

production of information. Hence the list of reasons that technological change differs from conventional commodities parallels the list of reasons that, more generally, markets for information differ from markets for conventional commodities. These arguments have been set forth in greater length in Joseph E. Stiglitz, "On the Microeconomics of Technical Progress," in George M. Katz, ed., *Technology Generation in Latin American Manufacturing Industries* (St. Martins Press, 1987), pp. 56–77; and Stiglitz, *Information and Economic Analysis* (Oxford University Press, forthcoming).

While conventional theory focuses on markets for homogeneous commodities, each piece of new knowledge must be different from previously produced knowledge. Moreover, one cannot show the good that is being sold in the market in the way that a firm producing chairs can show what it has produced. As a result, information is difficult to sell. At the same time, it is difficult to appropriate fully the returns to technical change. (Though patent laws attempt to facilitate the appropriation, they do so imperfectly.) Indeed, knowledge has many of the properties of public goods—being both difficult and undesirable to exclude.

Finally, while investments in physical capital goods are often financed by borrowing and using the physical capital good as collateral, expenditures on R&D are not for the most part collateralizable investments. Given limitations on firms' abilities to raise equity, the causes of which are discussed in Bruce Greenwald, Joseph E. Stiglitz, and Andrew Weiss, "Informational Imperfections in the Capital Market and Macroeconomic Fluctuations," *American Economic Review*, vol. 74 (May 1984, *Papers and Proceedings*, 1983), pp. 194–205, capital constraints (credit rationing) are thus likely to be more important than they are for conventional investments.

53. Technological change is also often produced as a by-product of production (learning by doing). These production decisions too are, by their nature, like sunk costs. The importance of these sunk costs for market equilibrium has been stressed by Dasgupta and Stiglitz, "Learning by Doing."

criterion with which to evaluate the effectiveness of the competition, but in his writings one can perhaps detect three propositions:

—R&D competition results in profits being driven to zero.

—R&D competition results in a succession of firms; competition occurs over time, not at a moment in time.

—Competition results in an efficient level of expenditure on R&D, with the gains from new innovations more than offsetting the static inefficiencies resulting from the temporary monopolies to which R&D gives rise.

In earlier work, Partha Dasgupta and Joseph Stiglitz and Richard Gilbert and David Newbery suggested that all three propositions were incorrect.⁵⁴ They showed that the threat of entry would indeed accelerate R&D, perhaps beyond a socially efficient level, but that the profits of the incumbent firm would still be positive and that it had an incentive to pace itself at such a rate as to deter entry, so that monopoly persisted.

That work ignored, however, the important role of sunk costs. Dasgupta and Stiglitz later suggested that with sunk costs the incumbent firm could preempt its potential rivals by an expenditure of only a limited amount of money.⁵⁵ Thus R&D competition might be relatively ineffective: profits would be even larger than suggested in the earlier analyses, and the pace of R&D would differ little from that engaged in by a pure monopolist. Potential competition need not exercise an effective discipline on the market.

SMALL, ENTRY-DETERRING R&D EXPENDITURES CAN OFFER COMPLETE PROTECTION

Consider the race for a patent in which the winner obtains all property rights to the use of the invention. There is no uncertainty in the R&D technology. To capture the patent, a firm has to complete a fixed number, N , of experiments before any other firm does. I assume that a firm has

54. Partha Dasgupta and Joseph E. Stiglitz, "Uncertainty, Industrial Structure, and the Speed of R&D"; and Richard J. Gilbert and David M. G. Newbery, "Preemptive Patenting and the Persistence of Monopoly," *American Economic Review*, vol. 72 (June 1982), pp. 514–26.

55. The earlier models (with no sunk costs) were not really models of a patent race but of a once-and-for-all decision about the allocation of resources to R&D. See Dasgupta and Stiglitz, "Uncertainty, Industrial Structure, and the Speed of R&D."

always to commit itself to its R&D decision over two periods and that capacity constraints limit each firm to M experiments, where $N > M$. Let $[N/M]$ denote the largest integer less than or equal to N/M . Write $n^* \equiv [N/M]$. Let $k^* \equiv N - n^*M \geq 0$.

Suppose to begin with that there is a single firm facing no actual or potential competition. Let the interest rate be positive. If demand for the product in question is growing, say at a constant rate, there is an optimal date, T^* , at which the incumbent would wish to complete N experiments. Then because the interest rate is positive, it pays the firm to postpone its research as much as it can toward T^* . It will conduct $e^*(t)$ experiments over periods $(t, t + 1)$ and $(t + 1, t + 2)$, where

$$(16) \quad \begin{aligned} e^*(t) &= 0 \quad \text{for } t = 1, \dots, T^* - 2n^* - 4 \\ &k^* \quad \text{for } t = T^* - 2n^* - 2 \\ &M \quad \text{for } t = T^* - 2n^*, T^* - 2n^* + 2, \dots, T^* - 2. \end{aligned}$$

Suppose now that there are instead two firms and that they alternate: firm 1 chooses its R&D policy at odd dates and firm 2 at even dates. Thus firm 1 has the first move. Now if the cost of each experiment (or anyhow, the first experiment in the sequence) is small enough, firm 1, faced with potential competition, will preempt.⁵⁶ With the threat of entry the unique subgame-perfect equilibrium consists of firm 1 committing itself to k experiments at $t = 1$, rather than at $t = T^* - 2n^* - 2$, but otherwise leaving its research program, $e^*(t)$, unaffected. Firm 2, of course, does not do any research at all.

This equilibrium is sustained by the following strategies. The incumbent firm commits itself at $t = 1$ to k experiments during the interval $(1, 3)$ and announces at the same time that at every odd date it will commit itself to the number of experiments its rival will have committed itself to at the previous even date. For suppose firm 2 chooses m experiments at some even date. Unless firm 1 matches it at the next date, firm 2 will no longer be behind. It then will pay firm 2 to pursue exactly the same strategy, matching whatever the incumbent does. If firm 2 does that, it will at worst share the patent, which, I assume, is also profitable. If it does not, it will obtain no return on its investment. Thus if firm 1 fails to match its rival, it loses a considerable amount of profit. Hence it

56. This assumes that the patent is worth a great deal more than having to share the invention and that having to share is in turn superior to losing.

is rational for the incumbent to match its rival. Because firm 2 knows this, the announced strategy is credible, and so it will not enter. We sum up by way of

proposition 5: Faced with a potential R&D competitor, the incumbent firm's optimal strategy, which sustains a subgame-perfect equilibrium, consists of preempting the potential entrant by committing itself to $e^{**}(t)$ experiments at each odd date t , in which

$$(17) \quad \begin{aligned} e^{**}(t) &= k, t = 1, \\ &0 \text{ if } S_1(t) < S_2(t) - k, \\ &M \text{ if } S_2(t) + k > S_1(t) > S_2(t) - k, \\ &\min [M, S^*(t) - S_1(t)] \text{ if } S_1(t) > S_2(t) + k, \end{aligned}$$

where $S_i(t) = \sum e_i(\tau)$, the sum of the experiments done to date, and $S^*(t) = \sum e^*(\tau)$, the sum of the experiments done through date t by a monopolist. (The equation is just the formal representation of the strategy I have described in words.)

Notice that if $(T^* - 2n^*)$ and M are large then equation 17 describes what is in effect an epsilon preemption: the incumbent takes a tiny R&D lead, and this is enough to keep the rival out of the race.⁵⁷

Concluding Remarks

Traditional economic theory has emphasized the importance of fixed costs and the nonconvexities to which they give rise. The presence of these nonconvexities limits the extent of competition. There is, however,

57. The literature on the possible persistence of monopoly has suggested that to deter entry an incumbent firm would have to pursue as intensive a research program as the most aggressive of its potential rivals, and that the incumbent would benefit by doing so. See, for example, Dasgupta and Stiglitz, "Uncertainty, Industrial Structure, and the Speed of R&D"; and Gilbert and Newbery, "Preemptive Patenting and the Persistence of Monopoly."

By way of contrast, in Dasgupta and Stiglitz, pp. 11–12, the idea of an epsilon preemption was introduced to suggest that if R&D has to be conducted over a number of periods, the fact that it is a sunk expenditure gives the incumbent firm (provided it has the first move) much more monopoly power than one might think. Proposition 5 confirms this conjecture.

The analysis of this section should be contrasted with that of Fudenberg and others, "Preemption, Leapfrogging, and Competition in Patent Races." In their analysis, firms must move simultaneously rather than sequentially. The firms focus on the mixed-strategy

a widespread belief that even if actual competition is absent, potential competition may be sufficient to discipline firms in an industry and to hold prices down. This argument demands that scale economies be due to costs that are not sunk. But it has not been recognized precisely how sensitive the conclusions are to the assumptions. I have shown that the presence of arbitrarily small sunk costs can serve as an absolute barrier to entry and make potential competition completely ineffective as a discipline device.

I have also suggested why it is that so much of the competition in modern industrial economies is focused on innovative activity rather than on the price competition on which conventional microeconomic analysis focuses—there is at least a hope that efforts put in this direction will yield profits—and why potential competition may be particularly ineffective in those sectors of the economy where R&D is important—expenditures on R&D are by their nature sunk costs.

Much of the analysis in this paper has employed the assumption of Bertrand competition. I have done this for two reasons. First, with Bertrand (but not with Cournot) competition price will equal marginal cost in markets in which the number of actual participants is limited. Bertrand competition thus provides a natural benchmark: one should not be surprised that markets without Bertrand competition are not efficient. Second, the standard “stories” in the contestability literature have firms undercutting each other, that is, the authors seem to presume that competition takes the Bertrand form.

I noted in the beginning that this was a paper partly about economics, partly about the models we employ to analyze economies. I have made use here of the standard concept of the perfect-equilibrium solution to

perfect equilibrium, in which the firm that is behind either does no research or a lot of research, while the firm that is ahead either maintains the monopoly pace or accelerates it slightly. There is thus some probability, when the firm that is ahead maintains its monopoly pace and the firm that is behind has an accelerated research program, that the firm that is behind leapfrogs ahead of its rival and becomes the new leader. This leapfrogging is not possible in our model. Which model is more appropriate depends on the lags in observation and reaction. If the incumbent firm can see the experiments its rival is undertaking and responds quickly, then the analysis of this section, in which firms announce that they will match their rivals' actions but in the interim maintain a slightly modified monopoly pace, seems plausible. Christopher Harris and John Vickers, “Perfect Equilibrium in a Model of a Race,” *Review of Economic Studies*, vol. 52 (April 1985), pp. 193–209, have extended the analysis of the sequential model to consider more general technologies. The results are similar to those reported here.

the almost standard game-theoretic formulations of firm interactions, with results that, to some, may seem disquieting. That standard literature has, I think, three distinct messages: that what is essential for entry is entrants' beliefs about the nature of ex post competition; that ex post competition depends on certain state variables; and that entrants' beliefs about the nature of ex post competition are arrived at by a process of introspection, or backward induction—they arrive at them by calculating what it is rational for the other firm to do, given that there is common knowledge of both the structure of the game and the rationality of the participants. The first two hypotheses I find eminently sensible; the third is far more questionable.⁵⁸ The world is too complicated to be well described by our simple models. Potential entrants in the market may be either not rational or not confident about the rationality of their rivals. If potential entrants form their expectations concerning the incumbent's responses by a process of extrapolation of past behavior, then entry, the incumbent's activities to deter it, and the incumbent's response to it may markedly differ from those hypothesized here. If potential entrants think there is even a small chance that the incumbent may not respond in the ruthless way hypothesized, they may be willing to take the risk of entering, provided sunk costs are not too small. Moreover, equilibria with small sunk costs may not look that much different from equilibria with zero sunk costs.⁵⁹ These more realistic models serve to qualify the results described here, but they also qualify those results that rely on similarly strong rationality hypotheses to ensure the efficiency of market economies.

More generally, I have observed five principles, the validity of which extends beyond the simple models investigated here.

—An increase in ex post competition reduces the effectiveness of ex ante competition, and conversely. Thus government policies that reduce ex post competition, such as price regulation, may actually enhance

58. Even the logical consistency of the third hypothesis, which underlies much of modern game theory, has recently been called into question by the work of P. J. Reny, "Common Knowledge and Extensive Form Games," *Journal of Economic Perspectives* (forthcoming); and K. Binmore, "Modeling Rational Players," unpublished paper (London School of Economics, 1985). For surveys of this work, see K. Binmore and A. Brandenburger, "Common Knowledge and Game Theory," *Journal of Economic Perspectives* (forthcoming).

59. Richard McMasters has explored these questions in a Ph.D. dissertation presently being completed at Princeton University.

social welfare, while policies aimed at increasing ex post competition may decrease it.

—Both the incumbent and the entrant have at their disposal mechanisms that affect the nature of ex post competition. In particular, the incumbent may choose to sink costs. In the R&D problem discussed in the section on sunk costs and innovation, the incumbent firm may deter entry by spending a relatively small amount earlier than it would if it were an unconstrained monopolist.

—Potential competition may decrease welfare as firms take actions to deter entry.⁶⁰ Thus the early expenditures on R&D do not result in an innovation occurring earlier than it otherwise would, and consumers are no better off. But these entry-detering expenditures do lower the present discounted value of the incumbent firm's profits. Similarly, in the case of natural monopoly, entry deterrence takes the form of excess capacity but with no benefits conferred on consumers. Firms may also sink costs when it is not efficient to do so, again simply to deter entry.

—Actual entry may not enhance welfare. Rather, the entrant may be accommodated through a cooperative arrangement maintained as a noncooperative perfect equilibrium. The incumbent's monopoly profits are shared (at the social cost of excessive expenditures on entry), but consumers are no better off.

—Sunk transactions costs may further fragment markets, so that the effective degree of competition may be much less than a naive look at the number of firms and customers might suggest.

I have shown how these basic principles cast light on some of the informal arguments used to buttress the contestability doctrine. In particular, the argument that (when there are sunk production costs) competition occurs in going after contracts becomes less persuasive once the inevitable sunk costs associated with writing contracts are taken into account. Similarly, price lags may serve to reduce ex post competition and thus encourage entry and make ex ante competition more effective. But with small lags in response, even small entry costs will deter entry. And with infinite lags in response, entry will occur, but prices may remain at the monopoly level: resources are simply wasted on excessive entry.

60. It is important not to read any policy implications into this result. In particular, I am not advocating a reduction in potential competition (and indeed, it is not apparent how one could go about the reduction). The concern here is only with ascertaining the effect of potential competition on prices, profits, and welfare.

These models leave the policy analyst in a quandary. They teach that reliance on the market may not suffice. They make it clear that in sectors in which increasing returns are important, deregulation may lead neither to lower consumer prices nor to more efficient production.⁶¹ They teach that when practices such as exclusive territories or vertical restraints seem to restrict competition—for instance, by deterring entry or facilitating collusion—they may in fact be doing just that.⁶² Certainly, the models provide no support for those who would loosen the enforcement of antitrust laws on the grounds that there are important returns-to-scale effects in R&D, and potential competition will ensure that even if a monopolist or dominant firm emerges, that firm will have to engage in rapid R&D to maintain its position.

But the models also suggest that government actions designed to increase competition may have their costs. As with any natural monopoly, there are economies of scale. There is an obvious solution: employ a welfare-based antitrust policy. But such a solution is an illusion: the history of antitrust law enforcement strongly suggests that the legal system does not provide a framework within which the necessary distinctions can and will be made.

The advantages of simple rules seem clear. The *per se* doctrine represented one such rule. There are those today who suggest quite different, more permissive rules. The question is one of trade-offs, in particular between the inefficiencies associated with the inability to exercise fully the economies of scale and those associated with the exercise of monopoly power to which, we have seen, sunk costs give rise.⁶³ If one believes, with Schumpeter, that the spur to innovation

61. For a discussion of the applicability of the contestability doctrine to rail or water carrier mergers, see William B. Tye, "The Applicability of the Theory of Contestable Markets to Rail/Water Carrier Mergers," *Logistics and Transportation Review*, vol. 21 (March 1985), pp. 57–76.

62. These practices, however, may have efficiency benefits as well, in which case one has to offset such benefits with the costs associated with decreased competition. At the same time, it should be emphasized that these practices may have efficiency costs rather than benefits.

63. At a number of points in this paper, I have called attention to the inefficiencies to which entry-deterring strategies give rise. It is not apparent, however, whether these activities will be greater with a strict or a loose antitrust policy. With a strict policy a large firm may be reluctant to engage in such activities, lest it be subject to antitrust prosecution; moreover, the larger number of firms in the market may mean that entry-deterring activities are less effective. With a loose antitrust policy, potential as well as actual competition may be decreased, and thus the need for entry-deterring policies may be reduced.

provided by competition is of central importance and if one gives credence to those studies purporting to show the importance of small firms in the innovation process, then the losses associated with a strict antitrust law may be outweighed by the gains. But this, I am sure, will remain a question of debate.

Appendix A: Alternative Formulation of Entry-Exit Game, Single Commodity

In an alternative formulation of the entry-exit game, within any period there are four moves. First the entrant firm decides whether to enter or, if it is already in the market, decides whether to leave. Then the incumbent firm decides whether to leave. Next, the entrant sets a price. Finally the incumbent sets a price. The ordering of moves is somewhat artificial. The results are similar to the simultaneous-move game described in text but somewhat simpler to see. It is clear that in any period in which the firms coexist, the equilibrium price is the Bertrand price; for if the entrant sets any price above the Bertrand price, it gets no sales because the incumbent will undercut.⁶⁴ It is also clear that in the third period, if the entrant does not exit, the incumbent will find exit beneficial if and only if $F_I \geq 0$. (If $F_I = 0$, the incumbent is indifferent to exiting.) The entrant, knowing this, will exit in the third period if and only if $F_I \leq 0$ and $F_e \geq 0$).

Moving back a period, if $F_I > 0$, if entry occurs in the second period, the incumbent will leave immediately. Hence entry will occur, but one monopolist is replaced with another. If $F_I < 0$, no entry will occur. Again, the threat of entry has no effect on prices.

Appendix B: Multiperiod Game with Exit

When the number of periods is extended, one can obtain a far richer set of cooperative equilibria. By the same token, it becomes more difficult for the incumbent to deter entry.

64. The entrant is actually indifferent with respect to the price it charges, since it knows it will either be undercut or it will make zero profits over variable costs. But if the firm incurs a cost, whether it sells the first unit or not (that is, there are costs to continuing as a firm, even if it does not produce), then the only equilibrium is the one described in the text.

To see this, consider the four-period model. The conditions for cooperation in the third period are the same as those analyzed earlier, and it is assumed they are satisfied. Let α_2 be the share of second-period monopoly profits allotted to the incumbent firm. It pays for the firm to cooperate, so long as

$$\alpha_2\pi + \delta^2g^2\pi \geq \pi + \delta F_I,$$

that is, so long as

$$\alpha_2 \geq 1 - \delta^2g^2 + (F_I\delta/\pi).$$

Similarly, for the entrant to cooperate requires that

$$(1 - \alpha_2)\pi + \delta g\pi + \delta^2F_e \geq \pi + \delta F_e,$$

that is

$$\alpha_2 \leq g\delta - \delta(1 - \delta)F_e/\pi.$$

It is easy to verify that the two conditions may be consistent, and provide a range of values of α_2 consistent with perfect equilibrium. For instance, if $g = \delta = 1$, cooperation requires that $1 \geq \alpha_2 \geq F_I/\pi$.

Comments and Discussion

Daniel McFadden: The main theme of Joseph Stiglitz's paper is that sunk costs are ubiquitous and may be used by incumbents to preempt entrants. In particular, investments in R&D can behave as sunk costs that permit an incumbent to charge monopoly prices without fear that other firms will enter. The paper reviews some of the main contentions on sunk costs and contestability. First, if there are positive sunk costs and instantaneous Bertrand rivalry is credible, then contestability fails.¹ Second, when sunk excess capacity is needed to make the threat of Bertrand rivalry credible, equilibrium may require the accommodation of small entrants and the practice of limit pricing.² Third, eliminating real sunk costs by preemptive contracting is not feasible if contract negotiation itself involves sunk costs. I would add that contracts that make an incumbent's response slow and encourage entry will be unattractive to them. They will seek instead contracts that precommit them to Bertrand rivalry in case of entry.

In the most innovative part of his paper, Stiglitz considers R&D as a

1. See Partha Dasgupta and Joseph E. Stiglitz, "Uncertainty, Industrial Structure, and the Speed of R&D," *Bell Journal of Economics*, vol. 11 (Spring 1980), pp. 1–28; Richard J. Gilbert and David M. G. Newbery, "Preemptive Patenting and the Persistence of Monopoly," *American Economic Review*, vol. 72 (June 1982), pp. 514–26; and Joseph Farrell and Garth Saloner, "Installed Base and Compatibility: Innovation, Product Preannouncements, and Predation," *American Economic Review*, vol. 76 (December 1986), pp. 940–55.

2. See Steven C. Sallop, "Strategic Entry Deterrence," *American Economic Review*, vol. 69 (May 1979, *Papers and Proceedings*, 1978), pp. 335–38; Avinash Dixit, "The Role of Investment in Entry Deterrence," *Economic Journal*, vol. 90 (March 1980), pp. 95–106; and Richard J. Gilbert, "Preemptive Competition," in Joseph E. Stiglitz and G. Frank Matthews, eds., on *New Developments in the Analysis of Market Structure* (MIT Press, 1986), pp. 90–123.

sunk cost. He argues that sinking even a small amount to obtain a slight technological advantage generates a credible threat that deters entry. This is then a model of entry to a patent race.³ The key features of Stiglitz's model are that incumbent and entrants move sequentially, incumbent first. There is no exogenous uncertainty, and each player has complete information on all previous moves. The actions of the players are R&D investment levels, with a known cumulative investment capturing a patent that conveys property rights to production of the product. The model also assumes that R&D costs cannot be recovered by divesting labs or extracting rents based on patent agreements signed by a firm's scientists; hence these costs are sunk. Finally, the model assumes that R&D investment capacity is limited, so two or more periods are required to obtain a patent.

Stiglitz concludes in this model that the incumbent has a superior strategy that blocks entry: an initial R&D investment, followed by matching any R&D investment of the entrant so as to guarantee winning the patent race. Because of this, a subgame-perfect equilibrium exists with preemption through a small sunk R&D investment by the incumbent. Because entry does not occur, the final R&D necessary to complete the patent is never done.

A natural question to ask is whether Stiglitz's conclusions hinge on the special assumptions of this model. First, consider the possibility of simultaneous moves (or, equivalently, sequential moves in which the second player does not know the move of the first player). In this case the players will follow randomized strategies, but preemption by R&D investment still holds.⁴

Second, consider the technology for R&D. A patent resulting from R&D investment may reduce the marginal cost of a product, increase product quality or introduce differentiation, thus increasing product demand, or convey property rights to production of the product. Stiglitz's model considers the last case, in which there is no direct social return to

3. Related models have been developed by Dasgupta and Stiglitz, "Uncertainty, Industrial Structure, and the Speed of R&D"; Gilbert and Newbery, "Preemptive Patenting and the Persistence of Monopoly"; and Drew Fudenberg and others, "Preemption, Leapfrogging, and Competition in Patent Races," *European Economic Review*, vol. 22 (June 1983), pp. 3–31.

4. Fudenberg and others, "Preemption, Leapfrogging, and Competition in Patent Races."

R&D. There is thus a social cost to the early preemptive R&D investment as well as to the monopoly pricing the incumbent can practice when the preemptive investment deters entry. If the marginal cost of a product is reduced or product quality is improved, there may be both an appropriate and a net direct social return to R&D, and both the behavior of the incumbent and the social consequences require further examination.

If R&D activities involve setup costs, then long-term expenditures at a steady rate will be more productive than the same cumulative investment cycled on and off. Such cycling makes preemption more expensive to the incumbent. If the patent has a private return, this may tip the balance and encourage early completion of R&D. Diseconomies of scale and scope in R&D experiments may, however, discourage catch-up by entrants.⁵

Exogenous uncertainty about the outcome of experiments is a feature of most R&D activities. How does uncertainty affect the possibilities of R&D preemption? Intuitively, the possibility of preemption is greatly restricted. The knife edge in which a lead in R&D gives a sure win to the incumbent disappears. And while the probable effect of early R&D by the incumbent firm is to give it an edge, the entrant may nevertheless find the expected profit from entering the patent race higher than its opportunities in other industries. This situation may require the incumbent to undertake a sufficiently large R&D preemption to lower the expected payoff to entrants to a point at which the payoff is less attractive than alternative opportunities. Equilibrium in this case will also depend critically on the information of the players. If the degree of success of the incumbent's experiments is known to the potential entrant, then the entrant can condition entry on circumstances in which the incumbent is disadvantaged. This in turn will force the incumbent to take actions that increase the probability of outcomes that are *ex post* preemptive.

A final issue in the technology of R&D concerns the place of patents and preemption in multiproduct markets. One may think of technology as having the shape of a tree, with fundamental processes as the trunk that supports the production of many products, and specialized processes as the branches that affect only a few products at the extremities. Patents enhance these processes and permit the patent holders to appropriate

5. Some of the issues surrounding the productivity of long-term contracting for R&D are considered in *ibid.*

the resulting returns by monopolizing or licensing the production of the products affected. How will an incumbent controlling a technology and facing entrants in some branches of it deploy R&D effort? Does preemption deflect R&D effort toward the threatened branches? How are the quality and cost of products that are not threatened by entry affected?

A final constellation of issues surrounding contestability concerns cases in which R&D is socially productive, say in lowering marginal cost. A perfectly discriminating monopolist who can appropriate all the benefits of the R&D will make the investment if and only if it is socially desirable (with lump-sum redistribution assumed if necessary). Suppose this firm is now faced with potential entry. Will R&D in the industry still come close to socially optimal levels? The following line of reasoning suggests that it will not. If R&D cannot be used effectively to preempt entry and contestability results, the incumbent firm will have difficulty ensuring that it can appropriate the benefits of R&D, and thus a reduced incentive to perform it. The point at which this occurs depends on the information available to the opponents and on the credibility of their threats, and is unlikely to coincide with the point at which R&D becomes socially productive. Further, the benefit to an incumbent of preempting entry may lead to R&D in situations in which it is not socially productive, as in the Stiglitz model. However, if R&D is socially productive and limited R&D permits preemption, it may nevertheless be the case that the existence of the potential entrant limits the ability of the incumbent to appropriate its benefits, leading to less R&D than is socially optimal. This suggests that policies that promote contestable market structures may be very inefficient in distinguishing socially productive and unproductive R&D.

Sam Peltzman: I find two ways to view this paper. One is as an inquiry into the minimal conditions that are logically required for contestability results. The second is as an elaboration of things that are in practice likely to cause wedges between price and marginal cost. The author does not tell me which view to take, but the paper does come across as a logical exercise that also purports to say something substantial. As a logical exercise it causes few problems, though I think it pushes conclusions too far and begins tripping over itself. I have stronger reservations, however, about its substantive message, although these reservations are more methodological than particular to this piece.

The author's argument rests on the knife-edge result about ease of entry in standard theory. If prospective profits are positive epsilon, everybody wants to enter the market. If epsilon is negative, potential entrants will stay out. Now assume there are sunk entry costs of epsilon. Add that to a situation of pure contestability and endow the entrant with a belief that competition will be so fierce that price will fall to marginal cost immediately after entry. Epsilon profits are then negative and entry can be deterred. Importantly, the pure monopoly price can be sustained. This is what was once called ruinous competition. It is ruinous competition theory written with epsilons, but ruinous nevertheless. Here the prospect of ruinous competition ($P < \text{ex ante } AC$) fully deters entry and supports the monopoly price. The conclusion is that the simplest sustainable result ($P = AC$) rests on this knife edge. It does not, as previously believed, recede slowly as sunk costs become larger.

The more general conclusion Joseph Stiglitz draws from this—one that has also been around for awhile—is the prospect of fierce, instant, price-equals-marginal-cost competition that can reduce actual competition and, more generally, deter investments that increase social welfare. I find nothing objectionable about the particular result on sustainability or the more general implication about the potential problems with price-equals-marginal-cost competition. The latter is, for example, a useful antidote to such bromides as the undesirability of exclusive use of knowledge that appear early in the paper and underlie a variety of other possible situations illustrating the potential welfare gains from slowing down rivalry in the short run to induce Pareto superior resource allocation. Examples include the old infant industry arguments for protectionism and modern sustainability arguments for entry restriction. All these cases raise doubts about how each of these restrictions may work in practice. But no case is implausible on its face.

The author's more particular result on sustainability serves as a useful reminder that, because entry-exit is one of the few areas of theory where a knife-edge result is important, one should not just wave one's hands about continuity results.

If I were advising Stiglitz's Princeton colleagues to reply, it would be that they should say, "Well, we always wave hands at such problems; what is special about this one?" For example, my students sometimes comment, "If profits are epsilon, anybody would like to enter this industry. But then everyone will enter, and if everyone enters, profits

will be negative. So if they think ahead one step, nobody will move. What do you say to that?" Well, I say, "Economics is frictionless, except where it is inconvenient." They are not all going to move. That is why you get all the nice properties or competition.

Toward the end of the paper Stiglitz tries to anticipate replies and needlessly overextends himself. First, he anticipates the rejoinder that the monopoly-price equilibrium can be undercut by long-term contracts. He does this by cobbling together a very special model in which the only contract offered is the monopoly price. To me, that only invites a wasteful competition in theories. Immediately I would invoke a Coase theorem about theories. If one model generates unexploited gains, another model will come along with some set of deals that realizes those gains. In this case there are buyers who otherwise face the monopoly price forever, and potential sellers who can serve them profitably at a lower price. So one has to expect a rejoinder in which a deal in principle breaks the monopoly price, to the mutual benefit of the buyers and potential sellers.

I cannot anticipate what the deal might be. Maybe one takes a bit of a buyers' coalition to underwrite the sunk costs here, adds a dash of vertical integration there, shakes well, and that will do the trick. But I do not think Stiglitz should anticipate that reply, either. The proper response, would I think go something like this. If breaking the monopoly price requires some gimmick such as long-run contracts, then sustainability becomes a special case in which the transaction costs required for whatever the appropriate deal is are trivial.

I have a similar reaction to his discussion of response lags. Response lags are not, as long-term contracts are, an extraneous consideration that Stiglitz can properly ignore. They are, after all, what motivated Bertrand's critique of Cournot. Stiglitz ends up being palpably uncomfortable in dealing with them: sometimes they help limit monopoly power and sometimes they do not; and then on the next page they probably are not important, anyway.

There are two difficulties with this approach. One is theoretical. The response lags here are grafted to the case arbitrarily. Stiglitz cannot make up his mind about exactly how arbitrarily they ought to be grafted on. Given that procedure, their role is going to remain as fuzzy as it seems to be here. Ultimately, restrictions on the speed of response must be derived from uncertainty and the costs this creates for detecting or

evaluating the consequences of a rival's move. They cannot be arbitrary. More importantly, simply dismissing the importance of response lags undercuts the larger message of the paper, with which I agree. Response lags are just an example of a limitation on the fierceness of postentry competition, and nothing would be lost if Stiglitz simply said, "That is exactly my point. Some such friction is often what is needed for effective competition."

Now, let me get to what the point of this paper might be for how actual markets behave. Here, I have to exercise some imagination because the paper only hints at applications for actual behavior. One possible inference is that entry deterrence is in fact important, that it is a pervasive feature of small-numbers rivalry because sunk costs are so pervasive. I am unaware, however, of any convincing evidence that this is true, and my own cursory look at the data does not support the idea. One formulation that occurred to me was, is there less entry the closer these markets approach monopoly? This modification is required because the monopoly of the theory is never found in the real world except in regulated markets. In any case, I identified the monopoly in his theory with small-numbers rivalry and then ran down the standard industrial classification. I took the first two high-concentration and first two low-concentration—the criterion was a four-firm concentration ratio higher than 60 or lower than 20—four-digit industries in each two-digit industry. Then I calculated the entry rate. Here are the results.

	<i>High (CR > 60)</i>	<i>Low (CR < 20)</i>
$100 \times \bar{\Delta} \log N$	4.6 (24.8)	2.0 (32.2)
$100 \times \bar{\Delta} \log VS$	71.6 (45.0)	76.5 (28.8)

The first row is the average log change in the number of firms from 1972 to 1982. The standard deviation is in parentheses. The second row shows the average growth rate of value of shipments for the same industries. High- and low-concentration industries grew about equally fast. So the room for new entrants is about the same in both segments, and the average rate of entry is about the same. Also, the odds of high entry rates, as reflected in the standard deviation, look about the same. Now, nothing so obviously crude can be definitive; all I am saying is that the case for strategic entry deterrence in the real world needs to be made.

Stiglitz says that it takes actual entry to lower price, not just potential entry, and cites the airline case. My reading of the literature on the

experience with airline deregulation so far suggests he is right. It does seem to take new entrants to lower fares significantly. A more indirect but more thorough test is afforded by the main body of the old literature on profits concentration that also tends to support his point. But Stiglitz's Princeton colleagues can also point to that literature as supporting some kind of working contestability story.

Suppose, for example, that a highly concentrated industry—one that has an 80 percent concentration ratio—is subject to a flood of new entrants that succeed in capturing half the large firms' business. A reasonably typical profits-concentration study would imply that, given this flood of entries, price would fall about 5 percent. Such a conclusion is not too bad for those who support contestability, even though it pushes in the direction that Stiglitz seems to want one to believe the world goes.

Or he may just be saying that small-number industries that approach monopoly markets work better because they can hope to avoid fierce competition. Again, the empirical literature on industrial organization tends to be supportive: high and rising concentration seems to be associated with dynamically superior price and productivity performance. But significant decreases in concentration also tend to be associated with superior price and productivity performance. So the data may simply mean that great technological progress upsets established structures.

Thus the broad tenor of the empirical literature is that market structure does seem to matter, roughly in the way suggested by this paper, but it does not matter nearly as much as the barrage of theory fired at the problem would suggest.

Nevertheless, perhaps none of all this is germane to the paper, which can support the negative of almost every inference I have just drawn from it. Does actual entry lower prices? One model says that entry leads to collusion. Another says it leads to wasteful duplication of investment. Does successful deterrence of entry conduce to efficiency? A series of models suggests it leads to the opposite—excess capacity, overly rapid obsolescence, and so forth. Stiglitz may reply—properly in some sense—that each outcome is possible, although in different circumstances. I have no objection to that. My objection is rather that we are given no guidance as to when to expect one outcome or another.

So, my quarrel here is not so much with this paper as with its claim, early on, that a major insight of the work on industrial organization during the past decade is its emphasis that what matters for the nature

of competition are variables that will affect how firms will interact strategically. It is indisputably clear that strategic interaction has attracted much model-building energy in the past decade. But what has not been shown is how or even whether the strategic considerations that elicit all this energy matter anywhere beyond the borders of the models in which they appear.

General Discussion

A major concern of the conference participants was the role the various models in this paper could have in guiding empirical work. The paper's meaning, Robert Willig proposed, is that incumbent firms may sometimes have devices available to help them deter entry and sometimes may not. The paper should be read as a guide to sorting out which circumstances cut which way, he suggested.

Carl Shapiro argued that this paper (and the whole body of literature on contestability theory) calls attention to two factors that are particularly important in predicting how markets will perform. One is the degree of sunk costs, which determines how hard it is to get into the market or to get back out; the other is the behavior that can be expected when there is more than one firm in a market. With the extreme assumptions of instantaneous price-cutting (Bertrand) competition, even small sunk costs are critical. The other extreme of no sunk costs and no price response by the incumbent produces a very different result. One role of the modeling exercise, he suggested, is to study the sensitivities of the results near the extremes.

Richard Levin noted that efforts to discriminate between the effects of these two classes of variables have already served as the basis for some empirical work on contestability. A recent study by Margaret Peteraf on monopoly airline markets, for example, attempted to measure the extent to which prices were driven by two sets of variables, one measuring the degree of sunk costs and the other attempting to predict price-cutting behavior on the basis of past reputations. Similarly, Clifford Winston noted that empirical work on airlines has found that potential competition seems to matter, but not nearly as much as actual competition.

Joseph Stiglitz responded that the paper was not intended to provide specific empirical implications but to provide a framework for classifying

markets according to certain kinds of exit costs and the kinds of ex post behavior those costs can be expected to produce. The models were intentionally oversimplified, with only three or four parameters in each, he argued, yet they produced very diverse behaviors. As a consequence, they provide a framework for sorting out which conditions lead to which reactions.

Timothy Bresnahan observed that the dispute about whether the paper provides testable empirical implications is, more deeply, about the role of theory in empirical work. Theorists should not necessarily be expected to produce empirical hypotheses, he argued. To do so may require them to be too literal-minded to produce good theory, while at the same time it encourages empirical researchers to be too literal-minded in their reading of theory. Sam Peltzman agreed that theorists should not necessarily have to produce precise empirical predictions. But his criticism was more general, he said, arguing that the whole body of work on strategic interaction has failed so far to produce very much in the way of new insights despite all the resources devoted to the enterprise.

Addressing this concern, Winston asked the participants to summarize the status of the contestability debate. Willig suggested that the debate has served to put the theorist in the role of a thoughtful potential entrant. It has forced them to consider the structural facts of the situation, the risks, and the reasonable expectations about incumbents' responses, and then to ask what kind of price-cost deviation would make a particular entry plan look profitable. Models such as those presented in this paper, he contended, help make such analyses sensitive to different conditions and different behaviors.

Stiglitz agreed that the game-theoretic modeling approach is essentially introspective in that it asks what the state variables are and how rational people will behave given these variables. Seen in this light, he conceded evidence that people rely on naive models of predicting their rivals' behavior, such as looking at their reputations for retaliation, is a bit disturbing for recent theoretical work.

So far, Shapiro observed, the debate has come down to whether there is any connection between prices before entry and pricing behavior after entry. If there is a strong connection, contestability theory will be applicable. If there is a weak connection, then potential entrants will not be able to influence current pricing.