A model of duopoly suggesting a theory of entry barriers

Avinash Dixit
Professor of Economics
University of Warwick
Coventry, England

This paper analyzes a model of duopoly with fixed costs. Leadership by one "established" firm may yield an outcome in which the second is inactive, but entry prevention is not a prior constraint. We find that two aspects of product differentiation have distinct effects: an absolute advantage in demand for the established firm makes entry harder, but a lower cross-price effect facilitates it. In the basic model we maintain the same quantity after entry. An extension of the model deals with the case where the threat of a predatory output increase after entry is made credible by carrying excess capacity prior to entry.

1. Introduction

■ The subject of industrial entry barriers is one where existing theory fails to do justice to the riches of factual observations and taxonomy. The empirical studies (Bain, 1956; Scherer, 1970, chapter 8; Needham, 1976) emphasize scale economies, product differentiation, and cost differences among firms. Textbook treatments of oligopoly (Fellner, 1949) assume away all these features in their formal models, confining real-world complexities to vague informal remarks.[1] Some analyses of Cournot equilibria allow some of the features mentioned above, but they do not examine the question of entry (McManus, 1964; Roberts and Sonnenschein, 1976; Dasgupta and Maskin, 1977). Among theories specifically addressed to this aspect, there are two types of models (see Scherer, 1970, chapter 8). In the case where the prospective entrants are a fringe of small price-taking firms, the established firms are assumed to calculate their residual demand, given the entrants' supply curves, and then act to maximize their own profits. The fringe supply is thus a reaction function, and the established firms are acting as von Stackelberg leaders, allowing entry to the extent that suits their own best interests. This seems a proper solution concept, and can be extended to sophisticated dynamic settings (Gaskins, 1971; Wenders, 1971b), but neglects scale economies entailing large potential

---

My greatest debt is to Michael Waterson for valuable discussions during the Summer Research Workshop on Oligopoly Theory at Warwick University during July, 1977, and detailed comments on the first draft. I am also grateful to Victor Norman, Nicholas Stern, Alvin Klevorick, and participants in seminars at Stirling, Sussex, and Warwick, for useful suggestions.

[1] Scherer (1970, chapter 5) does allow scale economies and differentiated products, but fails to draw the important conclusions that hinge on these aspects.

entrants. In contrast, the well-known Bain-Sylos-Modigliani model (see Modigliani, 1958) takes entry prevention as a prior constraint on the established firms, and finds a limit price to achieve this. No account is taken of the costs of the action, and the possibility that permitting entry may be the more profitable choice for the established firms is not considered.[2]

In this paper I attempt a treatment of the case of large prospective entrants by using the solution concept of leadership by established firms. The model is that of duopoly, with one established firm and one prospective entrant. This is for expository convenience: it is not hard in principle to apply the method to the case of several established firms collusively facing several potential entrants. The real difficulty lies in the dynamics of the game of strategy. The established firm will maximize its net worth, taking into account the threat and the consequences of entry. By virtue of being established, it can exercise leadership in the dynamic sense of choice of strategy trees or complete contingent plans in a supergame.

To avoid the analytical difficulties of the full problem, I shall consider simplified versions by restricting permissible strategies. The simplest case is the familiar "Sylos postulate," where a level of output is chosen and maintained forever,[3] whether or not entry occurs. This reduces the problem to static duopoly with quantity-setting, and the solution concept is the point of leadership by the established firm. This yields some interesting results; in particular, it allows me to express Bain's famous classification of entry possibilities (1956, pp. 21–22) in terms of the underlying parameters of costs and demands and to carry out some comparative statics.

The Sylos postulate has obvious flaws; the difficulty lies in finding convincing alternatives (Scherer, 1970, pp. 228–229; Wenders, 1971b). If entry becomes an irrevocable fact, the established firm will find it best to accept some output reduction. Prior to entry, on the other hand, it would wish to threaten a response of a predatory output increase. The threat of a large enough post-entry output will make entry seem unprofitable, and then it need never be implemented. But such a threat has to be credible to be successful, and the credibility usually entails a cost. It has been suggested that one way to maintain credibility is to carry enough capacity from the outset to produce the threatened post-entry output (Wenders, 1971a; Spence, 1977a). This is the second case I discuss. Once again entry-prevention is not a prior constraint, and a classification of cases along the lines of Bain is possible.

2. The basic model

■ The simplest case I consider is of a quantity-setting duopoly. In the context of entry, firm 1 is the established firm, and firm 2 the prospective entrant. The quantities of the products are x_1, x_2; the prices, p_1, p_2. There is a competitive numeraire sector whose output is x_0. The demands are assumed to arise from the utility function

$$u = U(x_1,x_2) + x_0. \tag{1}$$

[2] An exception is Osborne (1973), who considers a case where the von Stackelberg point occurs at a corner of the reaction function. But he neglects scale economies, which are a crucial aspect of the problem. For an independent critique of Osborne's paper, see Waterson (1977).

[3] The corresponding static assumption of price setting would be clearly inappropriate in the context of entry.

Since this has zero income effects on the duopoly industry, we can consider it in isolation. The inverse demand functions are the partial derivatives of the function U; thus

$$p_i = U_i(x_1,x_2), \qquad \text{for } i = 1, 2. \tag{2}$$

Profits of the firms are

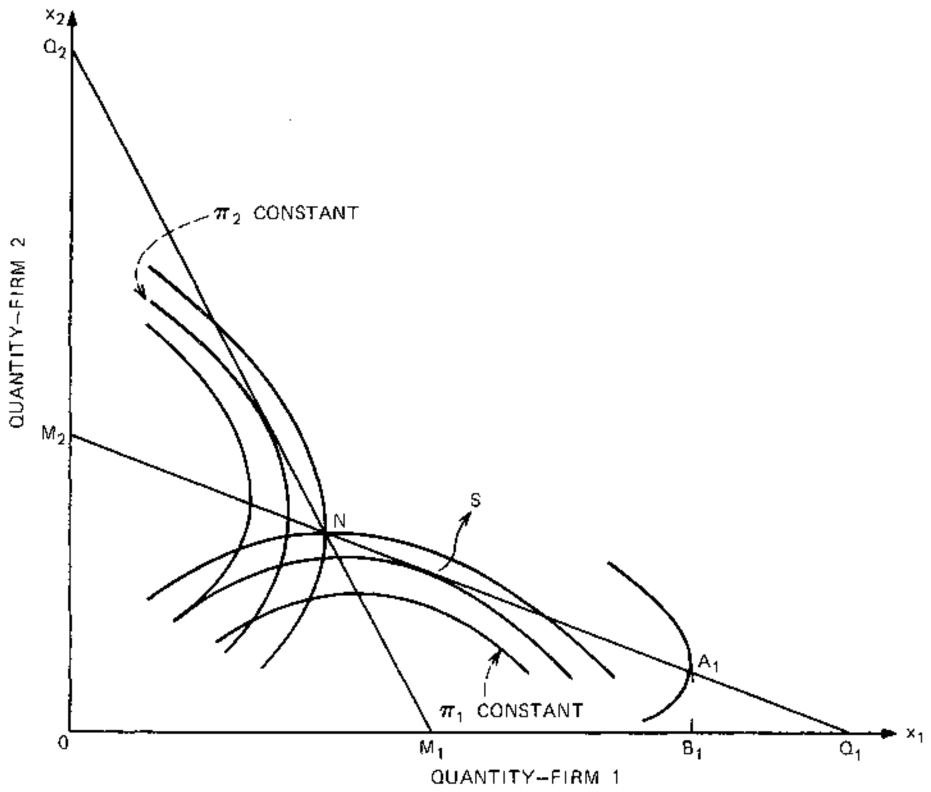
$$\Pi_i(x_1,x_2) = x_i U_i(x_1,x_2) - C_i(x_i), \qquad i = 1, 2, \tag{3}$$

where $C_i(x_i)$ are the total cost functions.

To provide a familiar point of departure, neglect scale economies for a moment. Figure 1 shows the conventional isoprofit curves and the reaction functions of the two firms in (x_1,x_2) space. I make all standard assumptions that yield downward-sloping reaction functions with a stable intersection; this is to isolate and highlight the new features to be introduced. Firm 2's reaction function begins at M_2, the point where Π_2 is maximized, given $x_1 = 0$, which is just the monopoly output for firm 2. Under mild restrictions on U, the reaction function will meet the x_1-axis, say at Q_1 where $\Pi_2 = 0$. I shall assume this to simplify some exposition; nothing important hinges on it. Similarly we have firm 1's reaction function M_1Q_2 in obvious notation. The point of intersection of the two is the Cournot-Nash equilibrium $N = (N_1,N_2)$, and the point $S = (S_1,S_2)$, where an isoprofit curve for firm 1 is tangent to the reaction function of firm 2, is the von Stackelberg point where firm 1 is the leader and firm 2 is the follower.

FIGURE 1

REACTION FUNCTIONS AND ISOPROFIT CURVES



It is possible that M_2Q_1 and M_1Q_2 do not meet, i.e., one of the firms is inactive in the Cournot-Nash equilibrium, but this is a trivial case. More subtly, there may be no tangency between 2's reaction function and 1's isoprofit curves, so that the von Stackelberg leadership point is at the corner Q_1. This is the case considered by Osborne (1973): even without scale economies, the best strategy for the established firm may be to deter entry by producing the limit quantity Q_1 and correspondingly charging the limit price $P_1 = U_1(Q_1,0)$. Being concerned here to highlight the richer possibilities that arise with scale economies, I shall neglect this case for the time being and return to it later.

Now introduce scale economies in the form of fixed costs. The isoprofit curves are unaffected in shape, but each one corresponds to a lower level of profit. In particular, Π_2 reaches zero at some point before Q_1, for example at A_1, as shown in the figure. Let B_1 be the point on the x_1-axis vertically below A_1. If x_1 is set in the segment B_1Q_1, the optimum response for firm 2 is no longer given by the appropriate point on A_1Q_1; it is better to secure zero profit by staying out. Hence firm 2's reaction function is now discontinuous, made up of the two segments M_2A_1 and B_1Q_1 and including the end points of both segments. The position of the discontinuity depends on the level of firm 2's fixed costs. If these are so high that firm 2 cannot even make a profit as a monopolist, then its reaction function is simply OQ_1; I shall ignore this case.

Similarly, fixed costs for firm 1 will give rise to a discontinuity in its reaction function. If there are more general scale economies, again there will be discontinuities in the reaction functions. But the nature of the discontinuities—a single downward jump to zero—will be preserved so long as the scale economies are moderate enough to keep marginal costs falling no faster than marginal revenue. I shall confine the discussion to the case of fixed costs for simplicity of exposition.

With discontinuous reaction functions, the nature of the equilibria changes. I begin by looking at the Nash equilibrium. If both fixed costs are small, the points of discontinuity lie in irrelevant regions and the Nash equilibrium is unaffected. If the fixed cost for firm 2 is large enough to take the point B_1 to the left of M_1, then we have a new meeting point at M_1 for the two reaction functions, i.e., a new Nash equilibrium. If the fixed cost is still larger, making $B_1 < N_1$, then the Nash equilibrium at N is eliminated. Similar remarks apply to the fixed cost for firm 1. Thus, depending on the values of the two fixed costs, there can be one, two, or three Nash equilibria.

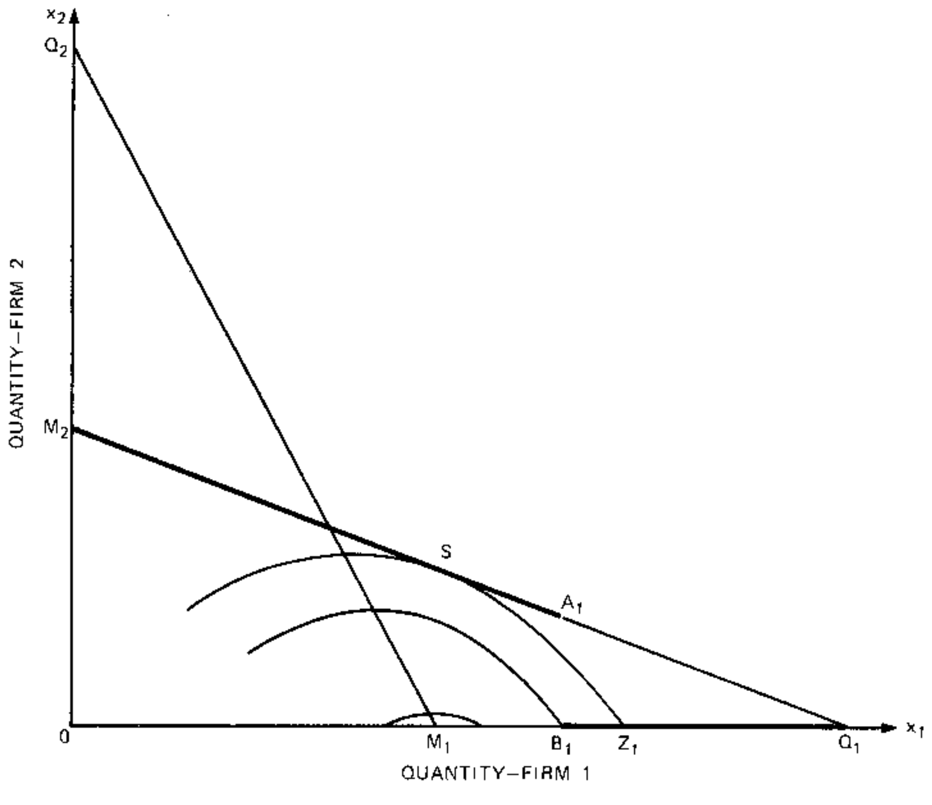
When there are multiple Nash equilibria, we cannot point to a deterministic outcome, even if we believe in a process of successive reactions leading to an equilibrium. Depending on where the two firms started, they might end up in a Nash equilibrium at N where both were active, or at M_1 or M_2 where only one of them survived to enjoy a monopoly. This suggests that we should pay more attention to historical or even purely accidental factors when economies of scale are important, since they can affect industrial structure in a significant way.

Turning to the question of entry, I begin with the case where the Sylos postulate is maintained, i.e., the established firm maintains its output at the same level whether or not entry occurs. The prospective entrant reacts to this, and the solution point is just the static leadership point for firm 1. We must of course see how this is affected by fixed costs, causing the discontinuity in firm 2's reaction function. Fortunately this difficult problem of constrained optimization allows a very simple geometric solution.

Figure 2 reproduces the relevant aspects of Figure 1, with the added point Z_1, where the isoprofit curve for firm 1 which is tangent to the line M_2Q_1 meets the x_1-axis. If firm 2's fixed cost is so small that the point B_1 of discontinuity in its reaction function lies to the right of Z_1, the best choice for firm 1 remains at S, and it is optimal for the established firm to allow entry. If the fixed cost for firm 2 is so large that B_1 lies to the left of M_1, the best point for firm 1 is M_1, i.e., it can ignore firm 2 altogether, and exercise unrestrained monopoly. The intermediate case where B_1 lies between M_1 and Z_1 needs more attention, and this is the case explicitly shown in Figure 2. Now firm 1 can do better than the old von Stackelberg point S by setting its output somewhat below Z_1, so that firm 2 stays out. This profit can be increased by further lowering x_1 up to any value slightly greater than B_1. If x_1 is set actually equal to B_1, firm 2 is indifferent between staying out and entering to yield the point A_1. However, its entry would lower firm 1's profits substantially. Therefore, so long as firm 1 thinks that there is a positive probability of entry at $x_1 = B_1$, there is a discontinuous downward jump in its expected profit as its output is lowered to B_1. In a technical sense, no optimum exists. However, we can sensibly think of a solution where firm 1 keeps its output only slightly greater than B_1. Then B_1 is the limit-output, and there is a corresponding limit price $P_1 = U_1(B_1, 0)$. The conclusion is that in this intermediate case, firm 1 finds it profitable to prevent entry, but cannot exercise unrestrained monopoly power.

FIGURE 2

LEADERSHIP SOLUTION WITH FIXED COSTS



All this is subject to the qualification that if firm 1's fixed cost is large enough, it may fail to make positive profits at some or all of these points. The industry will become a monopoly for firm 2, or even collapse altogether.

This completes the classification. In Bain's terminology, it can be stated compactly as follows:

(1) $B_1 < M_1$.	Entry is blockaded. Firm 1 has pure monopoly at $x_1 = M_1$.
(2) $M_1 < B_1 < Z_1$.	Entry is effectively impeded by limit pricing and $x_1 \cong B_1$.
(3) $Z_1 < B_1$.	Entry is ineffectively impeded, yielding the von Stackelberg duopoly equilibrium at S.

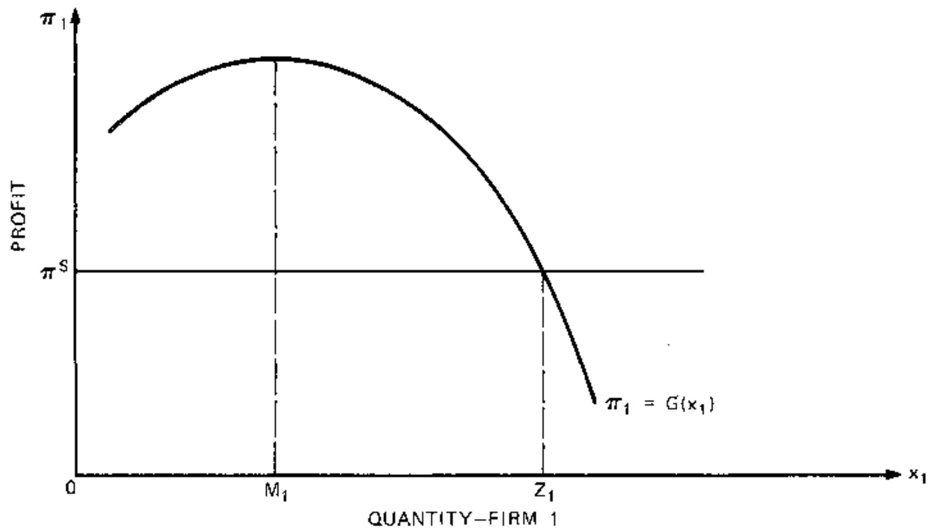
If the problem without fixed costs yields a corner von Stackelberg solution at Q_1, as in Osborne (1973), then with fixed costs there will be only two possibilities: if $B_1 > M_1$, entry will be effectively impeded with a limit pricing equilibrium at B_1, while if $M_1 > B_1$, entry will be blockaded with a pure monopoly equilibrium at M_1.

There is another diagrammatic exposition that is more convenient for later use. Figure 3 shows the approach for the basic case now being considered. Suppose for a moment that the output of firm 2 is held fixed at zero. The profit of firm 1 written as a function G of its output, i.e., $\Pi_1 = G(x_1)$, has the parabolic shape shown, and its peak is at M_1. The level of profit at the von Stackelberg point, say Π^s, is superimposed on this. The point of intersection to the right of M_1 is of course Z_1.

The output of firm 2 will in fact be zero if x_1 exceeds the barrier level B_1. The profit levels that firm 1 can actually attain are therefore given by all the points on the curve $\Pi_1 = G(x_1)$ to the right of B_1 attained by barring entry and the von Stackelberg level Π^s attained by allowing entry. It remains to pick the best of these, and that depends on the position of B_1. If $B_1 < M_1$, the best policy for

FIGURE 3

PROFITS AND ENTRY POSSIBILITIES



firm 1 is to set $x_1 = M_1$, and entry is irrelevant. If $M_1 < B_1 < Z_1$, the best policy is to keep x_1 just above the limit quantity B_1, thereby preventing entry. If $Z_1 < B_1$, the profit at $x_1 = B_1$ with entry prevented is not so high as that at S with entry allowed. This gives us an alternative view of the possible cases.

3. Some comparisons

■ The three critical quantities M_1, Z_1, and B_1 depend on the underlying parameters of demand and cost, so the classification scheme can in principle be expressed in terms of these basic magnitudes. Most importantly, we can examine how various changes in these underlying parameters affect the critical magnitudes and hence the entry possibilities. Such comparative statics will give us a better understanding of the forces that deter entry. Any change that raises B_1 can be said to make entry easier: if initially entry is blockaded, it moves closer to being merely effectively impeded, etc. Similarly, any change that raises M_1 or Z_1 can be said to make entry more difficult.

The simplest case is that of an increase in the fixed cost for the prospective entrant. This lowers B_1 while leaving M_1 and Z_1 unaltered, thus making entry more difficult. This is as it should be. An increase in the established firm's fixed cost has no effect on any of the three critical quantities. However, a sufficiently large fixed cost may make the whole enterprise unprofitable for the established firm, as was mentioned before.

Further comparative static analysis proves very difficult at the level of generality used so far. I shall therefore turn to a case involving linear demand and cost functions that yields some clear results.

Suppose the utility function is quadratic,

$$u = x_0 + \alpha_1 x_1 + \alpha_2 x_2 - \tfrac{1}{2}(\beta_1 x_1^2 + 2\gamma x_1 x_2 + \beta_2 x_2^2), \tag{4}$$

yielding linear inverse demands

$$\left.\begin{aligned} p_1 &= \alpha_1 - \beta_1 x_1 - \gamma x_2 \\ p_2 &= \alpha_2 - \beta_2 x_2 - \gamma x_1 \end{aligned}\right\} . \tag{5}$$

This can be valid only over a limited range of quantities, but these restrictions will be automatically satisfied at all relevant equilibria. Concavity of u requires

$$\beta_1 > 0, \qquad \beta_2 > 0, \qquad \gamma^2 \leq \beta_1\beta_2,$$

and the commodities are substitutes if

$$\gamma > 0.$$

They are identical if $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2 = \gamma$; a special case which will be mentioned occasionally. With nonidentical products, an absolute advantage in demand enjoyed by one of the firms will be reflected in a higher α for it, while γ measures the cross-price effects. These two aspects of product differentiation will be seen to have different effects.

The total costs for the two firms are

$$C_i = f_i + v_i x_i, \qquad i = 1, 2. \tag{6}$$

Thus f_i are the fixed costs and v_i the constant marginal (or average variable) costs. Write $\theta_i = \alpha_i - v_i$, reflecting the net absolute advantage for firm i.

It is easy to calculate several important quantities explicitly. The two monopoly outputs are

$$M_i = \theta_i/(2\beta_i), \qquad i = 1, 2. \tag{7}$$

The points where the "conventional" reaction functions—the ones with fixed costs ignored—meet the axes are

$$Q_1 = \theta_2/\gamma \qquad \text{and} \qquad Q_2 = \theta_1/\gamma. \tag{8}$$

The conventional Nash equilibrium has coordinates

$$\left.\begin{aligned} N_1 &= (2\beta_2\theta_1 - \gamma\theta_2)/(4\beta_1\beta_2 - \gamma^2) \\ N_2 &= (2\beta_1\theta_2 - \gamma\theta_1)/(4\beta_1\beta_2 - \gamma^2) \end{aligned}\right\} . \tag{9}$$

It is also useful to know the corresponding expressions for the socially optimal quantities, say X_1 and X_2 (not shown in the figures):

$$\left.\begin{aligned} X_1 &= (\beta_2\theta_1 - \gamma\theta_2)/(\beta_1\beta_2 - \gamma^2) \\ X_2 &= (\beta_1\theta_2 - \gamma\theta_1)/(\beta_1\beta_2 - \gamma^2) \end{aligned}\right\} . \tag{10}$$

I assume both these quantities to be positive; this fact will be used later. Of course, in the case of perfect substitutes with equal costs, only the sum $X_1 + X_2$ is relevant, and the separate solutions above become indeterminate.

The conventional von Stackelberg point S with firm 1 as the leader is

$$\left.\begin{aligned} S_1 &= (2\beta_2\theta_1 - \gamma\theta_2)/(4\beta_1\beta_2 - 2\gamma^2) \\ S_2 &= [(2\beta_1 - \gamma^2/(2\beta_2))\theta_2 - \gamma\theta_1]/(4\beta_1\beta_2 - 2\gamma^2) \end{aligned}\right\} . \tag{11}$$

Recall that I am assuming all these quantities to be positive. Given the assumption of positive quantities in (10), we can then verify that $M_1 > S_1$.

When fixed costs are introduced, we have the point of discontinuity in firm 2's reaction function

$$B_1 = [\theta_2 - 2(\beta_2 f_2)^{1/2}]/\gamma. \tag{12}$$

The point where firm 1's isoprofit curve through S meets the x_1-axis is more tedious to find. It can be shown to be:

$$Z_1 = M_1 + \{M_1^2 - [1 - \gamma^2/(2\beta_1\beta_2)]S_1^2\}^{1/2}. \tag{13}$$

We can now study the effects of various parameters on B_1 and M_1, thus finding how parameter changes affect entry conditions when the possible cases are effectively impeded entry and blockaded entry. An increase in firm 1's net absolute advantage raises M_1 while leaving B_1 unchanged. This makes blockaded entry more likely, i.e., entry becomes more difficult. An increase in θ_2 has the opposite effect. These results confirm our intuition. Making the products poorer substitutes i.e., lower γ, raises B_1 and so tilts the balance away from the blockaded case towards the effectively impeded case, i.e., makes entry easier. This is contrary to conventional views, and I shall return to this point later.

The comparison of effectively and ineffectively impeded entry is harder since the formula for Z_1 is very complicated. Matters are made easier if we observe that the real comparison is between firm 1's profits at S and B_1. Starting at a situation where the two profit levels are equal, we can then see how they respond to parametric shifts.

The general expression for firm 1's profit is

$$\Pi_1 = (\theta_1 - \beta_1 x_1 - \gamma x_2)x_1 - f_1. \tag{14}$$

Think of this as a function $\Pi_1(x_1,x_2,\sigma)$, where σ can be any relevant parameter. Its value at S, denoted Π^s, is

$$\Pi^s = \Pi_1(S_1,S_2,\sigma)\ . \tag{15}$$

It must be remembered that S_2 itself depends on S_1 and σ, say $S_2 = \phi_2(S_1,\sigma)$, according to firm 2's reaction function. Differentiating (15) with respect to σ and using envelope properties, we obtain

$$d\Pi^s/d\sigma = \Pi_{12}(S)\phi_{2\sigma}(S) + \Pi_{1\sigma}(S). \tag{16}$$

In particular, choosing different roles for σ,

$$d\Pi^s/d\theta_1 = S_1 \tag{16a}$$

$$d\Pi^s/d\theta_2 = -\gamma S_1/(2\beta_2) \tag{16b}$$

$$\begin{aligned} d\Pi^s/d\gamma &= \gamma S_1^2/(2\beta_2) - S_1S_2 \\ &= -S_1(\beta_1\theta_2 - \gamma\theta_1)/(2\beta_1\beta_2 - \gamma^2) \end{aligned} \tag{16c}$$

after some simplification.

For the value of Π_1 at B_1, denoted

$$\Pi^B = \Pi_1(B_1,0,\sigma),$$

we find

$$d\Pi^B/d\theta_1 = B_1 \tag{17a}$$

$$d\Pi^B/d\theta_2 = -2\beta_1(B_1 - M_1)/\gamma \tag{17b}$$

$$d\Pi^B/d\gamma = 2\beta_1(B_1 - M_1)B_1/\gamma\ . \tag{17c}$$

At the initial point, we assume $B_1 = Z_1$. Comparing (16a) and (17a), we then see that both are positive but the latter is larger. A greater net absolute advantage for firm 1 raises its profits in both configurations, but by a greater amount in the case where it effectively impedes firm 2's entry. Similarly, (16b) and (17b) are both negative, but the latter can be shown to have a larger absolute value. Thus a greater net absolute advantage for firm 2 lowers firm 1's profits by a greater amount when it bars entry, thus tipping the balance towards the ineffectively impeded case.

Once again, the effect of lowering γ is counterintuitive. Using the assumption of positive quantities in (10), we know that (16c) is negative, while (17c) is positive. A lower γ, making the products poorer substitutes, raises Π^s and lowers Π^B, making entry easier on both counts.

All the cross-price effects in this model go against the long tradition in the subject which regards product differentiation as a significant barrier to entry. While our formal demonstration is confined to the linear case, some rethinking is necessary. First, I would point out that the result is not unreasonable, in fact an extreme case of it should be quite evident. If γ is zero, the commodities are separate industries and firm 1's choices exert no power to prevent "entry" by firm 2. What I have done in the linear case is to show that the association of a lower cross-price effect and easier entry holds over the whole range of γ.

Second, in the descriptive literature, we often find some vagueness in the concept of product differentiation (Bain, 1956, chapter 4; Scherer, 1970, chapter 8). Most instances involve both absolute differences in demands, reflected here in the α_i, and finite cross elasticities, captured here in γ. We see now that both are relevant for entry conditions, but that they can have opposing influences.

This suggests a need for keeping the two concepts more clearly distinguished than is customary.

Some further results can be found by depicting the profits of firm 1 as a function of γ. This is done in Figure 4. Consider the profit at the von Stackelberg point Π^s. Having signed (16c) above, we see that it is a decreasing function of γ. Since at $\gamma = 0$ we have $S_1 = M_1$, it follows that at $\gamma = 0$, Π^s is just the monopoly profit Π^M. The profit at the barrier point, Π^B is increasing when $B_1 > M_1$, as (17c) shows. But B_1 is a decreasing function of γ from (12), so Π^B is increasing up to $\gamma = \gamma^{**}$, defined by $B_1 = M_1$, and decreasing thereafter. At γ^{**}, $\Pi^B = \Pi^M$. Finally, let γ^* denote the point where $\Pi^S = \Pi^B$.

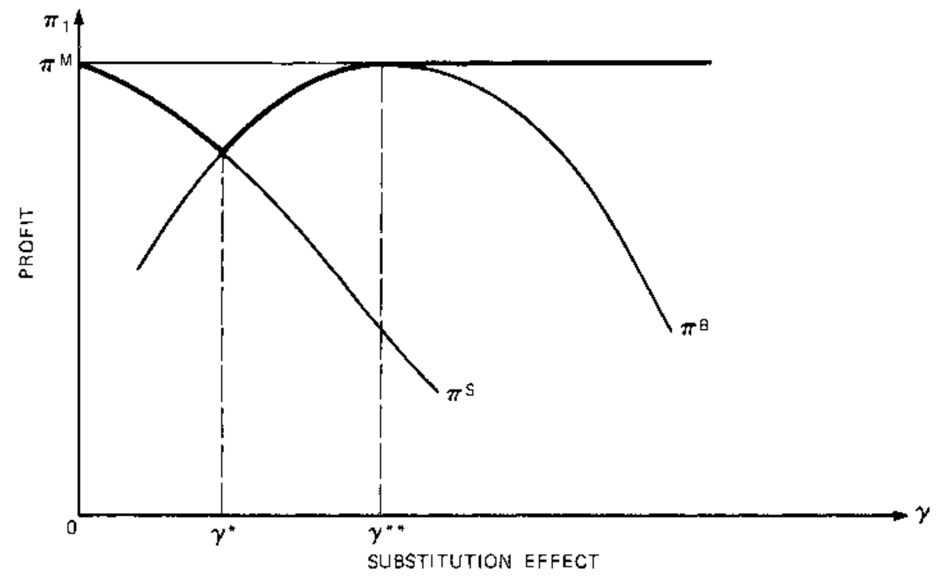
Now for $\gamma < \gamma^*$ firm 1 finds it better to allow entry. For $\gamma^* < \gamma < \gamma^{**}$ it prevents entry by producing at B_1, while for $\gamma^{**} < \gamma$ it can prevent entry while producing at the monopoly level M_1. Therefore firm 1's profit as a function of γ is the upper envelope shown as the heavy curve in Figure 4.

This analysis leads to the following conclusions. In the region where the established firm finds it better to allow entry, its profit would be increased if the two products were poorer substitutes. Where it prevents entry, however, its profits would be increased if the products were better substitutes. This may sound strange at first, but on reflection we see good sense behind the results. Entry is more easily prevented if your product can be claimed to be a good substitute for any prospective entrant's product. It is when entry has occurred that you can better exploit monopoly power by claiming a special niche for your product, and thus a lower cross-price elasticity between products.

It can be argued that γ is not the most appropriate parameterization of cross-price effects because a lower γ raises the demand curves facing both firms, and that a better comparison would twist each demand curve about some fixed initial point. The calculation involves changing α_i at the same time as one changes γ to maintain the chosen points fixed. The results again confirm the importance of distinguishing between absolute advantage in demand and cross-price effects.

FIGURE 4

PRODUCT DIFFERENTIATION AND PROFITS



4. The excess capacity case

■ Although the basic model illustrates the underlying ideas in a simple diagram, it is special because it is static and restricted to a duopoly. Extensions to handle several firms present difficulties largely of an algebraic nature and are left to the reader. Conceptually more interesting extensions are found by considering wider ranges of firm strategies. Of particular interest is the strategy of threatening a sufficiently large output in the event of entry while maintaining enough capacity to make that threat credible. I shall consider this strategy for a case of linear demands and costs, although a generalization on the demand side is immediate.

The two demand functions are as in (5), but the costs of firm i are

$$C_i = f_i + w_i x_i + r_i k_i, \tag{22}$$

where x_i is the output and k_i the capacity, and we require $x_i \leqslant k_i$. The marginal cost of expanding output and capacity together is $w_i + r_i \equiv v_i$. Now firm 1 can threaten a postentry output of k_1 while producing only $x_1 (\leqslant k_1)$ so long as entry does not occur. For firm 2, the relevant quantity is k_1; it will stay out if $k_1 \geqslant B_1$, the output defined in (12). There is clearly no reason for firm 2 to maintain excess capacity as there are no more potential entrants.

Now suppose for a moment that the output of firm 2 is held fixed at zero, and that firm 1 has a given capacity k_1. Firm 1's profit will be

$$\Pi_1 = (\alpha_1 - \beta_1 x_1)x_1 - f_1 - w_1 x_1 - r_1 k_1.$$

Then

$$\partial \Pi_1 / \partial x_1 = \alpha_1 - 2\beta_1 x_1 - w_1 = 2\beta_1(\mu_1 - x_1)$$

where

$$\mu_1 = (\alpha_1 - w_1)/(2\beta_1). \tag{23}$$

The quantity μ_1 is clearly firm 1's monopoly output if there is enough spare capacity so that the marginal cost of increasing output is just w_1. (If there is not enough capacity, then the marginal cost of expanding output and capacity together is relevant, and the monopoly output is M_1; clearly $\mu_1 > M_1$.)

If k_1 is fixed at a value below μ_1, the choice of x_1 to maximize Π_1 above is at the limit of its permissible range, *viz.*, $x_1 = k_1$. If $k_1 > \mu_1$, then x_1 is best set at μ_1 and spare capacity of $(k_1 - \mu_1)$ left. Correspondingly, after the best choice of x_1 is made, we can write Π_1 as a function of k_1, say $\Pi_1 = H(k_1)$, defined as

$$\Pi_1 = H(k_1) \equiv \begin{cases} (\alpha_1 - v_1 - \beta_1 k_1)k_1 - f_1, & \text{if} \quad k_1 \leqslant \mu_1; \\ (\alpha_1 - w_1 - \beta_1 \mu_1)\mu_1 - r_1 k_1 - f_1, & \text{if} \quad k_1 > \mu_1. \end{cases} \tag{24}$$

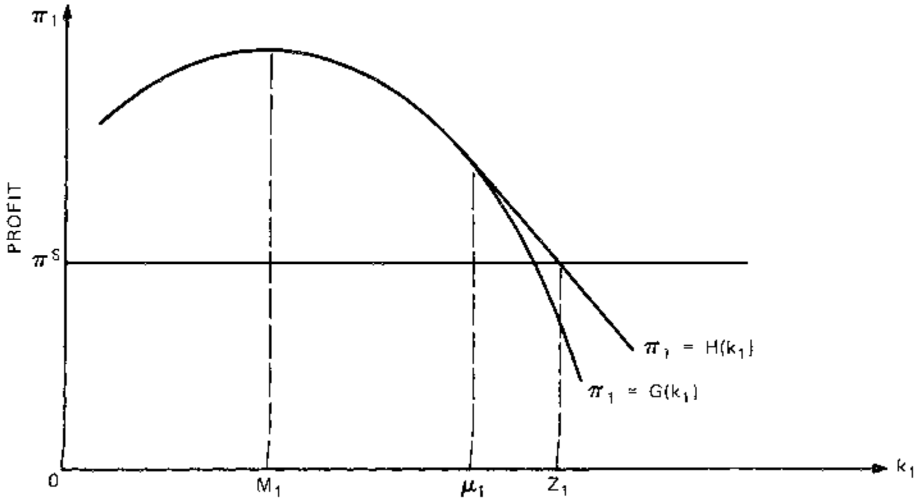
This is shown in Figure 5. If the excess-capacity strategy were not available, x_1 would have to equal k_1 and then $\Pi_1 = G(k_1)$ defined in Section 2. In our current example

$$G(k_1) = (\alpha_1 - v_1 - \beta_1 k_1)k_1 - f_1,$$

which is also shown in Figure 5 where it differs from $H(k_1)$. The figure illustrates how the possibility of excess capacity shifts up the monopoly profit function. This has obvious effects on the desirability of entry prevention. (Note that while the formula for μ_1 is special to the linear case, the idea of such a dividing level with excess capacity to its right is much more general.)

FIGURE 5

THE EXCESS CAPACITY STRATEGY



Once again we draw the level Π^s of firm 1's profit at the von Stackelberg duopoly point. No excess capacity should exist at S, and its solution is as in Sections 2 and 3. The intersection of the Π^s line with $\Pi_1 = H(k_1)$ to the right of M_1 will be labeled Z_1 and will play the same role as before.

The assumption that firm 2's output is zero is appropriate only for $k_1 > B_1$. We can then obtain various cases depending on the position of B_1. Suppose first that $Z_1 > \mu_1$ as in Figure 5. There are four cases:

(1) $B_1 < M_1$. Entry is blockaded, and $x_1 = k_1 = M_1$.

(2) $M_1 < B_1 < \mu_1$. Entry is effectively impeded by conventional limit pricing, with $x_1 = k_1 = B_1$.

(3) $\mu_1 < B_1 < Z_1$. Entry is effectively impeded by excess capacity, with $x_1 = \mu_1$, $k_1 = B_1$.

(4) $Z_1 < B_1$. Entry is ineffectively impeded, i.e., allowed to occur, and $x_1 = k_1 = S_1$, $x_2 = S_2$.

If $Z_1 < \mu_1$, the excess capacity case does not arise, and we are back in the situation of Section 2.

We see that the strategy of excess capacity can enlarge the zone where entry is effectively impeded at the expense of the zone where it is allowed to occur. Allowing excess capacity to be held introduces a second way of barring entry that is preferable over a portion of the range.

Comparative static analyses for this model are similar in principle to those of Section 3 and are left to the reader. An obvious next step is to consider shifts in demand functions resulting from selling effort, thus extending Williamson's analysis by making entry endogenous in it (Williamson, 1963; Needham, 1976).

5. Concluding comments

■ In this paper I have suggested a general theoretical approach to the problem of entry of new firms which are comparable in size to existing ones. The approach does not take entry-prevention as a prior constraint, and it allows existing firms to choose their best strategy bearing in mind the reactions of prospective entrants. The analysis allows for fixed costs and differentiated products.

The method enables us to explain various entry possibilities in terms of underlying parameters, and to study comparative static effects. It is found that a greater absolute advantage in demand (or cost) for established firms makes entry harder, but lower cross-price effects with potential entrants' products make entry easier. This suggests that industrial organization economists should keep these two aspects distinct, instead of lumping them together into one vague concept of product differentiation as they usually do.

The approach takes only a limited account of the dynamics inherent in the problem, and an improved treatment of this aspect seems the most pressing problem of future research.[4] In particular, the question of credibility of threats needs more attention.

References

BAIN, J.S. *Barriers to New Competition*. Cambridge: Harvard University Press, 1956.

DASGUPTA, P.S. AND MASKIN, E. "The Existence of Economic Equilibria: Continuity and Mixed Strategies." Manuscript, 1977.

FELLNER, W. *Competition among the Few*. New York: Knopf, 1949.

GASKINS, D.W. "Dynamic Limit Pricing under Threat of Entry." *Journal of Economic Theory*, Vol. 3, No. 3 (September 1971), pp. 306–322.

MCMANUS, M. "Equilibrium Numbers and Size in Cournot Oligopoly." *Yorkshire Bulletin of Economic and Social Research*, Vol. 16, No. 2 (1964), pp. 68–75.

MODIGLIANI, F. "New Developments on the Oligopoly Front." *Journal of Political Economy*, Vol. 66, No. 3 (June 1958), pp. 215–232.

NEEDHAM, D. "Entry Barriers and the Nonprice Aspects of Firms' Behavior." *Journal of Industrial Economics*, Vol. 25, No. 1 (September 1976), pp. 29–43.

OSBORNE, D.K. "On the Rationality of Limit Pricing." *Journal of Industrial Economics*, Vol. 22, No. 1 (September 1973), pp. 71–80.

ROBERTS, J. AND SONNENSCHEIN, H. "On the Existence of Cournot Equilibria without Concave Profit Functions." *Journal of Economic Theory*, Vol. 13, No. 1 (August 1976), pp. 112–117.

SCHERER, F.M. *Industrial Market Structure and Economic Performance*. Chicago: Rand-McNally, 1970.

SPENCE, A.M. "Entry, Investment, and Oligopolistic Pricing." *The Bell Journal of Economics*, Vol. 8, No. 2 (Autumn 1977), pp. 534–544.

———. "Investment Strategy and Growth in a New Market." *The Bell Journal of Economics*, Vol. 10, No. 1 (Spring 1979), pp. 1–19.

WATERSON, M.J. "Osborne and Limit Pricing." Manuscript, 1977.

WENDERS, J.T. "Excess Capacity as a Barrier to Entry." *Journal of Industrial Economics*, Vol. 20, No. 1 (November 1971), pp. 14–19.

———. "Collusion and Entry." *Journal of Political Economy*, Vol. 79, No. 6 (December 1971), pp. 1258–1277.

WILLIAMSON, O.E. "Selling Expense as a Barrier to Entry." *Quarterly Journal of Economics*, Vol. 77, No. 1 (February 1963), pp. 112–128.

---

[4] See Spence (1979b) for a dynamic model of industry evolution with a related approach.