A MODEL OF DUOPOLY SUGGESTING A THEORY OF ENTRY BARRIERS*

Avinash Dixit

NUMBER 125

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

UNIVERSITY OF WARWICK
COVENTRY

A MODEL OF DUOPOLY SUGGESTING A THEORY OF

ENTRY BARRIERS*

Avinash Dixit

NUMBER 125

February 1978.

1. Introduction

The descriptive and empirical literature in Industrial Organization abounds with references to economies of scale, product differentiation, and absolute advantages in cost or demand enjoyed by some firms over others. In particular, they feature prominently in Bain's list of barriers to entry in [1] . However, textbook models of oligopoly often abstract away from all these features, and assume a homogeneous product with equal constant costs for all firms in the industry. The real-world complexities are confined to informal remarks, suggesting vague modifications of the formal results. This happens even in such outstanding treatments as those of Fellner [3, chs. 2,3] and Scherer [8, chs. 5-8].[1]

In this paper I present a simple model of duopoly that includes all three of the features mentioned above, and concentrate on the new aspects that emerge. I hope to show that there is much to be gained from so doing, without making the analysis unduly difficult. Perhaps the most important benefit is the beginning of a unified theory of barriers to entry. By placing different forces that deter entry into a common model, we can compare their strengths and study their interactions. We can also go behind the conventional classification of entry possibilities established by Bain [1, pp.21-22], and express the various cases in terms of the underlying parameters of costs and demands. This enables us to see how entry possibilities change as a result of changes in these parameters.

The treatment of entry in the existing literature differs sharply depending on whether the prospective entrants are a fringe of small firms, or comparable in size to the existing firms.[2] In the former case, established firms are assumed to calculate their residual demand given the prospective entrants' supply curves, and then act to maximise their own profits. The entrants' supply function is then a reaction function, and established firms are acting as Stackelberg leaders. There are very sophisticated dynamic variants of this, but the basic idea remains. The analysis is very different for the case of large entrants. Here the well-known Bain-Sylos-Modigliani approach takes entry prevention as a prior constraint on the behaviour of established firms, and finds a limit price to achieve this. No account is taken of the costs of such action, and established firms are not allowed to act as leaders, allowing entry or preventing it as is more profitable for them.

In an early comment, Fisher [4] contrasted the Bain-Sylos-Modigliani model and the Cournot-Nash model of oligopoly, and concluded in a comment that:

> "The true analogy, of course, is with a Cournot 'leadership' model with all firms already in the industry playing the collective leader to the potential entrant's follower. This is another story, however."

The story was taken up independently by Osborne [6], who considered the possibility of a Stackelberg solution at an extreme of the conventional reaction function, and called this the case where limit pricing is consistent, i.e. where established firms find it more profitable than permitting entry. However, he did not make adequate allowance for scale economies, and it will be seen that they are a crucial feature of the problem.[3]

My basic model is that of duopoly, with one established firm labelled 1, and one prospective entrant labelled 2. The rest of the economy is assumed competitive, and is aggregated into a numeraire good labelled 0. The quantities of the commodities are x_i $(i=0,1,2)$; the prices of the commodities in the duopoly industry are p_i $(i=1,2)$. The firms are supposed to be quantity-setters, partly because this allows a simpler modelling of some discontinuities that are important for my purpose, and partly because pure price-setting behaviour by existing firms would be inappropriate when entry is possible. However, other aspects of what follows can be equally well illustrated assuming price-setting. My focus is on the equilibrium where firm 1 is the Stackelberg leader, and firm 2 the follower. There is some independent interest in the Nash equilibrium with scale economies, and I look at it in a brief digression.

To keep the argument simple, I shall assume that scale economies arise solely from fixed costs. However, the results carry over provided marginal costs are not falling too rapidly. I shall also consider an example with linear costs and demands, which allows some explicit solutions and comparative statics.

Finally, I shall consider some extensions. Some cases with several established firms and several potential entrants can be handled as simple generalizations of the basic model. The static nature of the model is a more serious drawback, and I shall consider one extension which allows the established firm the strategy of threatening a different output level in the event of entry, along the lines of Spence [9]. Since Spence takes entry prevention as a prior constraint, this extends his model by making the issue of entry endogenous as well.

2. The Basic Model

The demands are assumed to arise from the utility function

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

This has zero income effects on the duopoly industry, allowing us to consider it in isolation. The inverse demand functions are the partial derivatives of the function U; thus

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

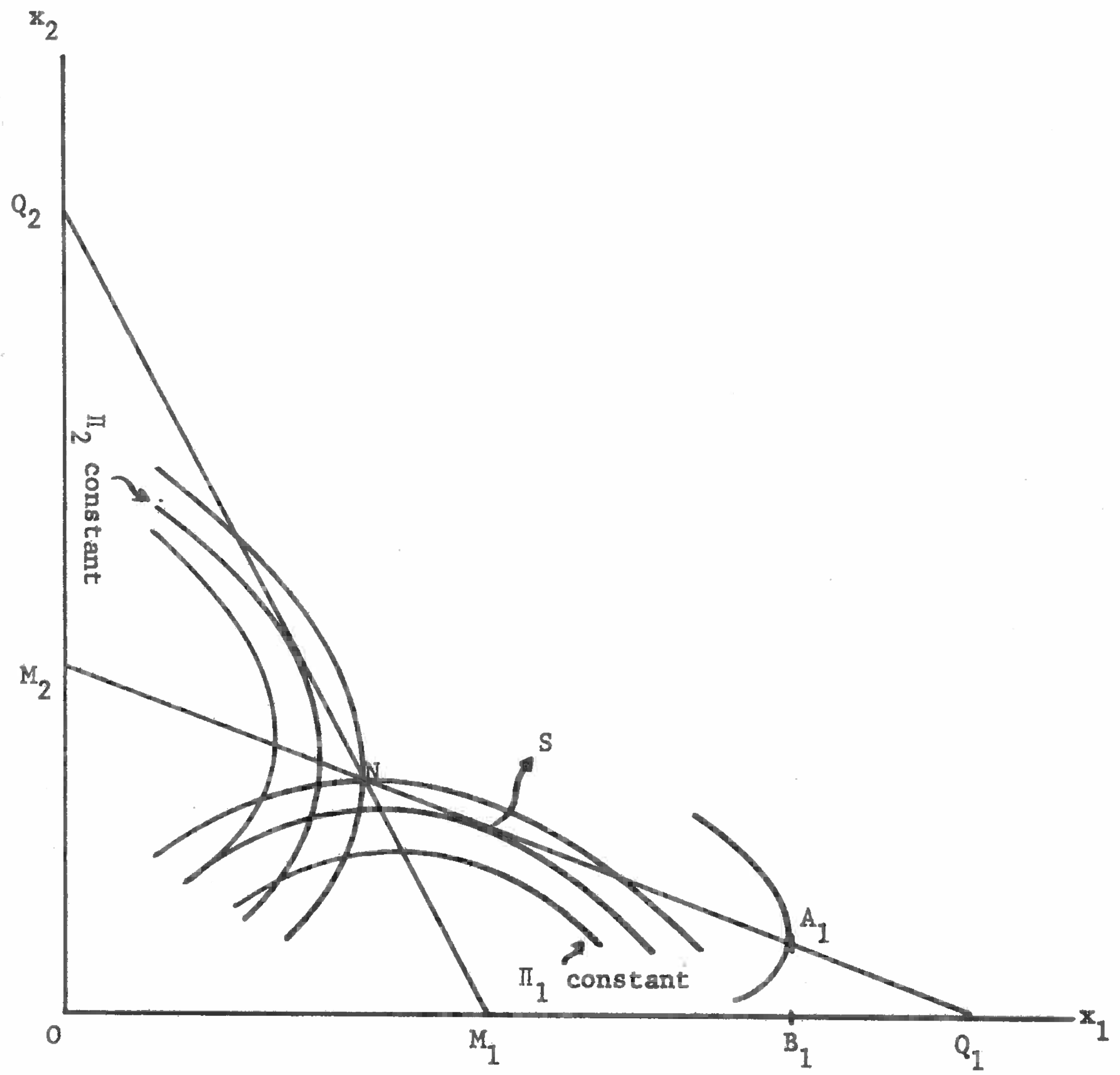
Profits of the firms are

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

To provide a familiar point of departure, neglect scale economies for a moment. Now we have the conventional Figure 1. The iso-profit curves and the reaction functions of the two firms are shown in (x_1, x_2) space. The reaction functions are downward sloping so long as the two products are substitutes in the sense that an increased quantity of one lowers the marginal revenue curve for the other. 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. 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, and the point S, where an iso-profit curve for firm 1 is tangent to the

Figure 1

Reaction functions and iso-profit curves



reaction function of firm 2, is the outcome where firm 1 is the Stackelberg leader and firm 2 is the follower.

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 iso-profit curves, so that the Stackelberg leadership equilibrium is a corner solution at Q_1. This is the case considered by Osborne [6]: 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 $P_1 = U_1(Q_1,0)$.[4] 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.

Note that in Figure 1, firm 2's profit falls steadily from its monopoly level at M_2 to zero at Q_1 in the absence of fixed costs. Now introduce scale economies in the form of fixed costs. The iso-profit 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 truly optimum response for firm 2 is no longer given by the appropriate point on A_1Q_1, as that yields negative profit. The globally optimum response is 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 including the end points of both segments. The position of the discontinuity depends on the level of 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 trivial case.

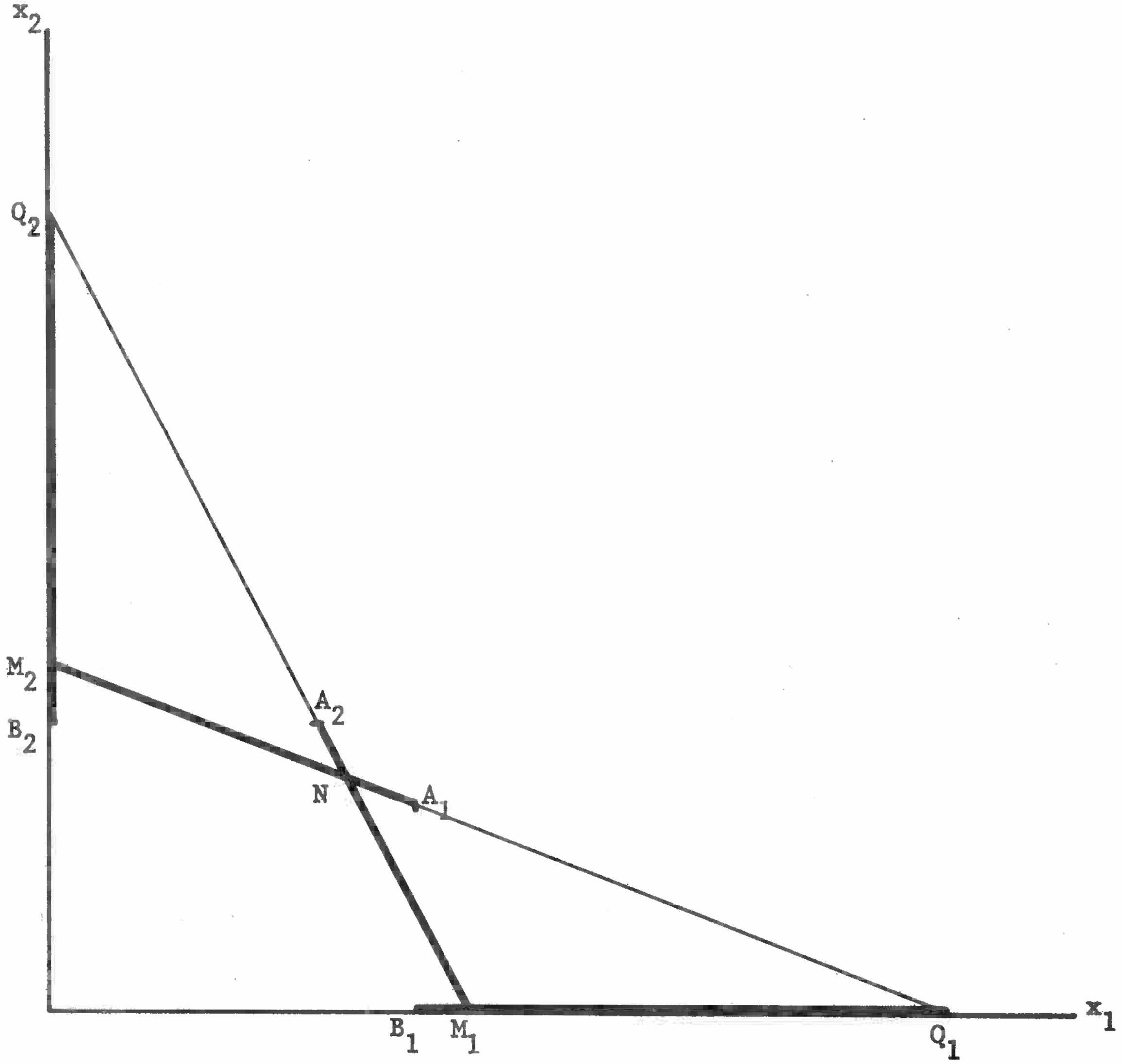
Similarly, fixed costs for firm 1 will give rise to a discontinuity in its reaction function. If there are more general scale economies, again there will be discontinuities in the reaction functions. But the nature of the discontinuities - a single downward jump to zero - will be preserved so long as 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. Figure 2 shows the case of three; the others are easy to construct.

Notice that as we gradually increase the fixed cost for firm 2, the equilibrium at M_1 is introduced <u>before</u> that at N is eliminated. This rules out the case of zero Nash equilibria. This is somewhat surprising since we are used to thinking of fixed costs as likely to produce non-existence problems. A more rigorous argument shows what is going on. The reaction functions (more properly correspondences since there are two possible responses at the critical break-points) $x_2 = \phi_2(x_1)$ and $x_1 = \phi_1(x_2)$ are both in a natural sense non-increasing. The composite $\phi_1(\phi_2(x_1))$ is in the same sense non-decreasing. Also, it has

Figure 2

Nash equilibria with fixed costs



relevant left- and right- partial continuity properties. A theorem of Roberts and Sonnenschein [7] then guarantees the existence of a fixed point, i.e. a Nash equilibrium. Unfortunately it applies only to functions of a scalar variable. Thus the proof works only for a duopoly, and with more firms non-existence becomes a possibility.[5]

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.

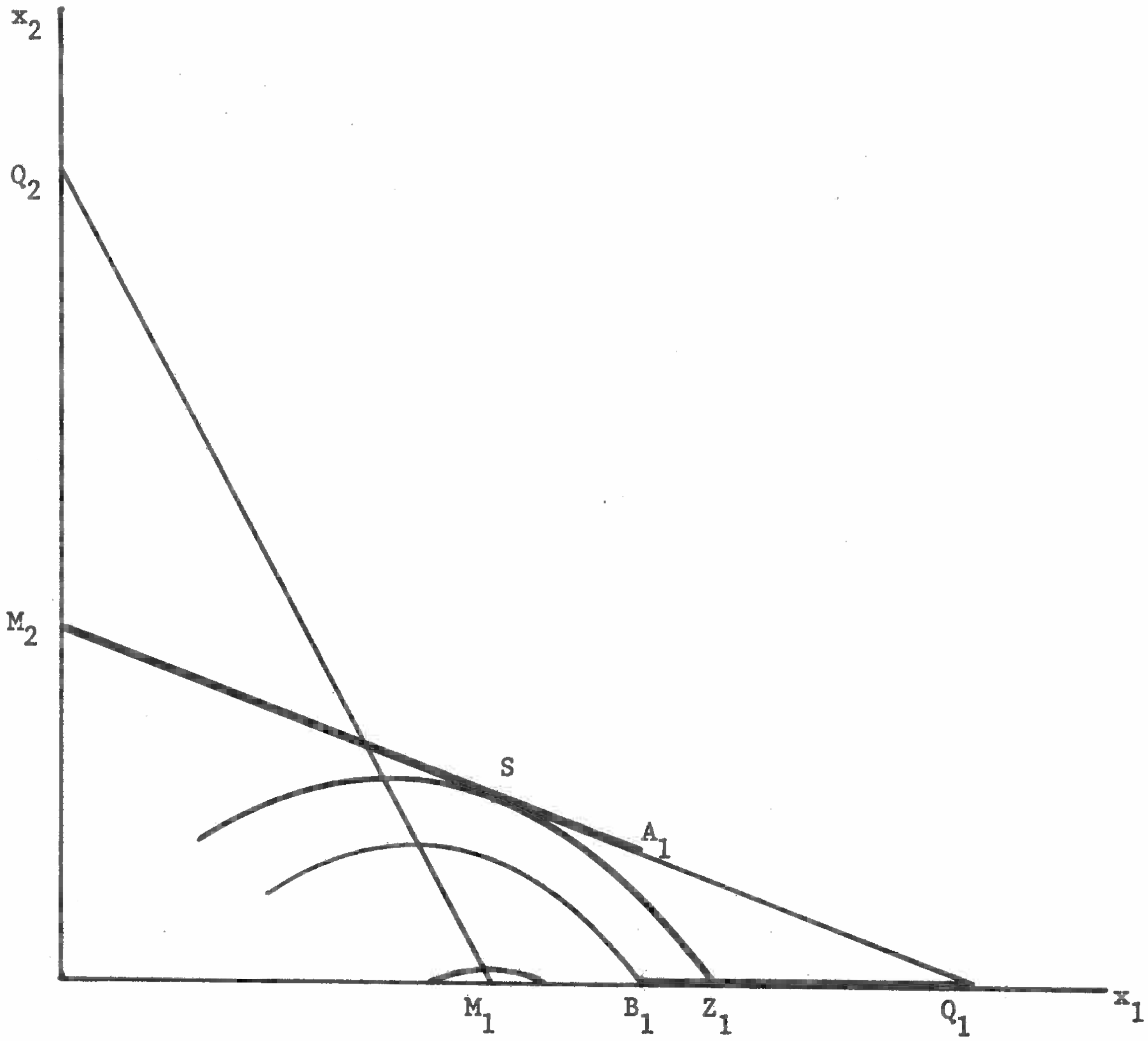
Although Nash equilibria are of interest in the general theory of oligopoly, they are not very relevant to the question of entry. The established firm can simply stick to its leadership output, leaving the prospective entrant to respond as best as it can. The resultant Stackelberg equilibrium must yield the established firm more profit than the Nash one.[6] Therefore I turn to consider how the Stackelberg point is affected by fixed costs. Firm 1 must now maximise its own profit given the discontinuous reaction function of firm 2. Fortunately this difficult problem of constrainedoptimization allows a very simple geometric solution.

Figure 3 reproduces the relevant aspects of Figure 1, with the added point Z_1, where that iso-profit 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 3. Now firm 1 can do better than the old 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.

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. But in that case the discussion of its exercise of leadership becomes vacuous anyway, so I shall assume that the problem does not arise.

Figure 3

Leadership solution with fixed costs



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 $x_1 \tilde{=} B_1$.

(3) $Z_1 < B_1$. Entry is ineffectively impeded, yielding the Stackelberg duopoly equilibrium at S.

If the problem without fixed costs yields a corner Stackelberg solution at Q_1, as in Osborne [6], 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.

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 level of it 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.

3. An Example

Suppose the utility function is quadratic,

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

yielding linear inverse demands

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

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

$$\beta_1 > 0, \quad \beta_2 > 0, \qquad \gamma^2 \leqslant \beta_1 \beta_2$$

and the commodities are substitutes if

$$\gamma > 0 .$$

They are perfect substitutes if $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2 = \gamma$; a special case which will be mentioned occasionally. An absolute advantage in demand enjoyed by one of the firms will be reflected in a higher α for it, while γ measures the extent of product differentiation, i.e. the cross-price effects.

The total costs for the two firms are

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

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

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

and the points where the conventional reaction functions when fixed costs are ignored meet the axes are

$$Q_1 = \theta_2/\gamma \quad , \quad Q_2 = \theta_1 / \gamma \tag{8}$$

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 optimum quantities say X_1 and X_2 (not shown in the figures)

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

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 Stackelberg point S with firm 1 as the leader is

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

Recall that I am assuming all these quantities to be positive. Given (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

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

The point where 1's iso-profit curve through S meets the x_1 -axis is messier to find: we have

$$Z_1 = M_1 + \{\,M_1^2 - [1 - \gamma^2/(2\beta_1\beta_2)]\;\;S_1^2\,\}^{\frac{1}{2}} \qquad (13)$$

We can now study the effects of various parameters on B_1 and M_1, thus finding how parameter changes affect entry possibilities when the cases in the balance are those of ineffectively impeded 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. A greater degree of product differentiation, i.e. lower γ, raises B_1 and so tilts the balance away from the blockaded case towards the ineffectively 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 messy. Matters are 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 are equal, we then see how they respond to parametric shifts.

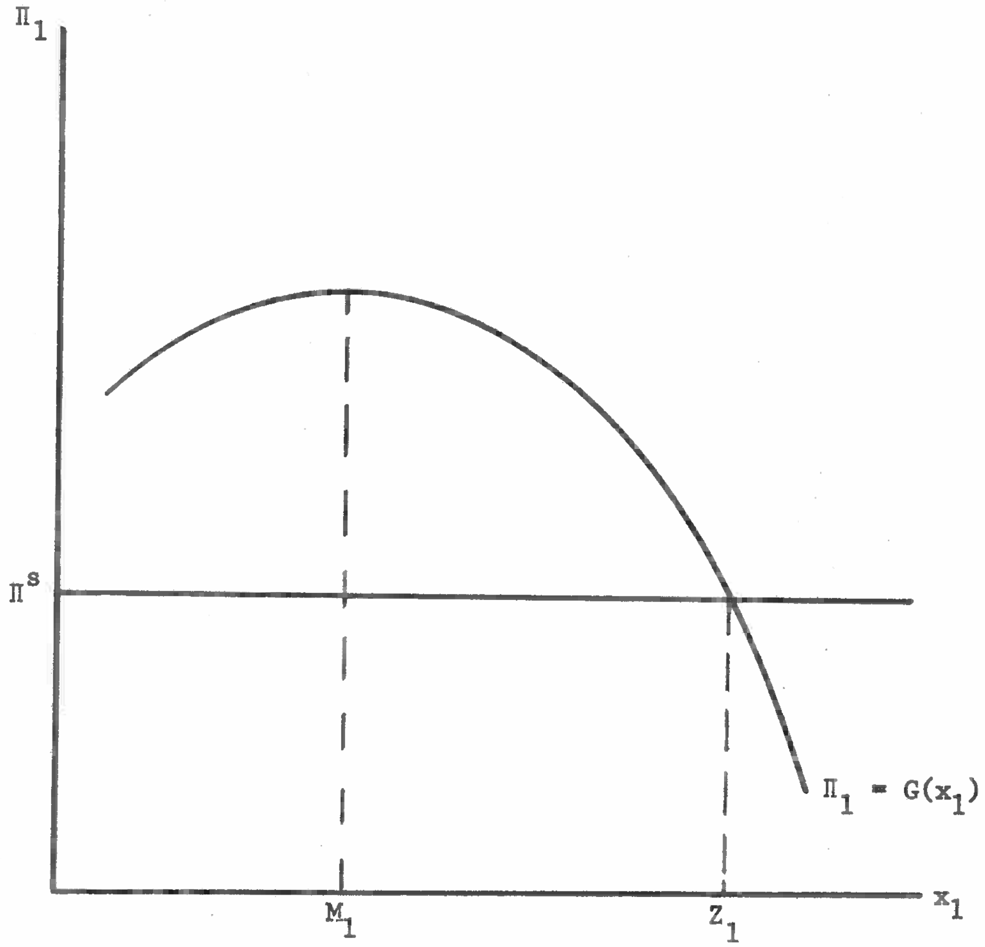
Before doing this algebraically, I shall illustrate the method in a diagram. Suppose for a moment that the output of firm 2 is held fixed at zero. The profits of firm 1 written as a function G of its output are simply

$$\Pi_1 = G(x_1) \equiv (\theta_1 - \beta_1 x_1)\, x_1 - f_1. \tag{14}$$

This has the parabolic shape shown in Figure 4, and its peak is at M_1. The level of profit at the Stackelberg point, say Π^s, is superimposed on this. The point of intersection to the right of M_1 is of course Z_1.

<u>Figure 4</u>

Profits and entry possibilities



The classification of entry possibilities can now be explained in a different way. The assumption of $x_2 = o$ in the above expression for Π_1 is valid only for $x_1 > B_1$. The feasible profit levels for firm 1 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, or the 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 firm 1 is $x_1 = M_1$ and entry is irrelevant. If $M_1 < B_1 < Z_1$, the best policy is to keep x_1 just above B_1, and entry is prevented by limit pricing. If $B_1 > Z_1$, the profit at $x_1 = B_1$ with entry prevented is not as high as that at $x_1 = Z_1$ with entry allowed.

Comparative statics can now be done by examining how the function G and the level Π^s shift as various parameters change. The idea behind the diagram is clearly valid for more general functional forms.

These shifts have to be studied using some algebra. The general expression for firm 1's profit is

$$\Pi_1 = (\theta_1 - \beta_1 x_1 - \gamma x_2)\, x_1 - f_1 \tag{15}$$

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

$$\Pi^s = \Pi_1\,(S_1, S_2, \sigma)$$

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. It is easy to calculate that

$$S_2 = (\theta_2 - \gamma S_1)/(2\beta_2)$$

Differentiating totally,

$$d\Pi^s/d\sigma = \Pi_{11}(S)\ dS_1/d\sigma + \Pi_{12}(S)\ dS_2/d\sigma + \Pi_{1\sigma}(S)$$

where Π_{11}, Π_{12} and $\Pi_{1\sigma}$ are the partial derivatives of Π_1, and (S) indicates evaluation at S. Similarly,

$$dS_2/d\sigma = \phi_{21}(S)\ dS_1/d\sigma + \phi_{2\sigma}\ (S)\ .$$

Now for each σ, S_1 is chosen to maximise the leader's profit using the reaction function, i.e. $\Pi_1(S_1, \phi_2(S_1,\sigma),\sigma)$. Therefore

$$\Pi_{11}(S) + \Pi_{12}(S)\ \phi_{21}(S) = o$$

and therefore

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

In particular,

$$d\Pi^s/d\theta_1 = S_1 \qquad (16a)$$

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

$$d\Pi^s/d\gamma = \gamma S_1^2/(2\beta_2) - S_1 S_2$$

$$= -\ S_1\ (\beta_1\theta_2 - \gamma\theta_1)/(2\beta_1\beta_2 - \gamma^2) \qquad (16c)$$

after some simplification.

For the value of Π_1 at B_1, say

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

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, $B_1 = Z_1$. 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 product differentiation is counter-intuitive. Using (10), we know that (16c) is negative while (17c) is positive. A greater degree of product differentiation (lower γ) raises Π^s and lowers Π^B, making entry easier on both counts.

All the effects of product differentiation in this model go against the long tradition in the subject of regarding it as a significant barrier to entry. While the formal demonstration is confined to the linear case, some rethinking is called for. 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 two 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 more product differentiation and easier entry holds over the whole range of γ.

If one reads the descriptive literature in the light of this result, one begins to see some vagueness in the concept of product differentiation. The examples involve not only the economic theorists' idea involving cross-elasticities that is captured here in γ, but also and even more often, an absolute advantage in demand.[7] Both are relevant in affecting entry possibilities, but there is much to be said for keeping them distinct in applied industrial economics.

It can however be argued that a lower γ lowers the demand curves facing both firms, when a more realistic comparison should be a twist of a demand curve about some fixed initial point. The net effect is not clear, but the calculation is not difficult, and the case of a twist about S follows. To leave demand prices unchanged at quantities (S_1,S_2), a change $d\gamma$ must be accompanied by changes $d\theta_1$, $d\theta_2$ such that

$$d\theta_1 - S_2\, d\gamma = o = d\theta_2 - S_1\, d\gamma .$$

Using this in (16) and (17), the appropriate total derivatives with respect to γ are

$$d\Pi^S/d\gamma = S_2S_1 - S_1 \ \gamma S_1/(2\beta_2) + \gamma S_1^2/(2\beta_2) - S_1S_2$$

$$= o ,$$

$$d\Pi^B/d\gamma = S_2 B_1 - S_1 2 \beta_1 (B_1 - M_1)/\gamma + 2 \beta_1 (B_1 - M_1) B_1/\gamma$$

$$= S_2 B_1 + 2 \beta_1 (B_1 - M_1)(B_1 - S_1)/\gamma$$

$$> o \quad \text{at} \quad B_1 = Z_1.$$

In this sense, a decrease in γ lowers the profit from impeding entry while leaving that on allowing entry unchanged. Again, entry is made easier.

In conclusion, it should be pointed out that the effect of product differentiation depends on the level of fixed costs and vice versa; in other words, the various factors affecting entry are interdependent. This point, made by Bain [1,pp.118-9], can be given more precise content in my model.

4. Some Extensions

The basic model, though it illustrates the underlying ideas in a simple diagram, is special in being static and being restricted to a duopoly. In this section I shall indicate possible generalizations. For simplicity I shall confine this to the case of linear demands and costs; the principles apply to more general functions. Also, I shall omit many details.

Consider first the case where there are n_1 established firms acting collusively towards each other and towards n_2 prospective entrants. Let x_1 and x_2 stand for industry outputs and y_1 and y_2 those of each firm of types 1 and 2 respectively. Let the demand functions be as in (5), and cost functions $C_i = f_i + v_i y_i$. For given $x_1 = n_1 y_1$, we have the demand curve facing the sub-industry of type 2 firms. Their Cournot equilibrium can be calculated using well-known methods, as the reaction function

$$x_2 = \frac{n_2}{n_2+1} \quad \frac{\theta_2 - \gamma x_1}{\beta_2} \tag{18}$$

and the sub-industry's profits are

$$\Pi_2 = n_2 \left\{ \frac{(\theta_2 - \gamma x_1)^2}{\beta_2 (n_2+1)^2} - f_2 \right\} \tag{19}$$

If n_2 is a fixed number, this will give rise to a discontinuity in (18) at the point

$$B_1(n_2) = \left[\theta_2 - (n_2 + 1)(\beta_2 f_2)^{\frac{1}{2}} \right] / \gamma \tag{20}$$

Iso $-\Pi_1$ contours can now be superimposed on the reaction function and the collusive decision of established firms analysed using the methods of Sections 2 and 3. If on the other hand n_2 can adjust to reduce Π_2 to zero, we have from (19) the equilibrium value

$$n_2 = (\theta_2 - \gamma\, x_1)/(\beta_2 f_2)^{\frac{1}{2}} - 1 \tag{21}$$

This treats n_1 as a continuous variable, but really we should take it to mean only the integer part. Substituting in (18), we see that the reaction function is decreasing, with downward jumps each time n_2 drops by one. The last jump, to zero, occurs at

$$B_1 = \left[\theta_2 - 2\ (\beta_2 f_2)^{\frac{1}{2}}\right]/\gamma$$

while some care is necessary in locating the point equivalent to S of Section

2, broadly similar ideas govern the comparison between barring and allowing entry from the point of view of the established firms.

Next suppose the established firms behave collusively towards potential entrants (e.g. through an institutionalized body governing entry) but in a Cournot manner among themselves. Consider the case of fixed n_2. For each given x_1 to the left of $B_1(n_2)$, the firms of type 2 will react as in (18), so the sub-industry of type 1 will face a residual demand curve

$$p_1 = \left[\alpha_1 - \frac{\gamma\, \theta_2\, n_2}{\beta_2(n_2+1)}\right] - \left[\beta_1\, \frac{\gamma^2\, n_2}{\beta_2(n_2+1)}\right] x_1 \ .$$

The Cournot equilibrium of n_1 such firms can be calculated, and the whole thing will be consistent if the resulting x_1 is less than $B_1(n_2)$. If x_1 lies to the right of $B_1(n_2)$, the residual demand curve is simply

$$p_1 = \alpha - \beta_1\, x_1$$

yielding another Cournot equilibrium. If both are possible, the governing body can choose the one with higher Π_1. The full calculation is not illuminating. In any case, this combination of internal competition and collusion when dealing with outsiders may be difficult to sustain.

Another drawback of the basic model is its static nature. A proper dynamic model will mean a very difficult differential game with non-convexities, but one important aspect deserves attention. A simple quantity-setting strategy is too restrictive as it allows no flexibility in responding to entry. We should at least allow firm 1 a two-dimensional strategy space with actual pre-entry output and threatened post-entry output as the choices, and let firm 2 respond to such a strategy. Of course the threat of a large enough post-entry output will keep firm 2 out and then need never be implemented. But such a threat has to be credible to be successful, and this credibility usually entails a cost. Spence [9] has suggested that one way to maintain credibility is to carry enough excess capacity at all times. I shall consider a simple case of Spence's model from the point of view of the theory of entry barriers.

The case is that of duopoly with linear demands and costs. Suppose the two demand functions are as in (5), but 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 $x_i \leq k_i$. The marginal cost of expanding output and capacity together is $w_i + r_i \equiv v_i$. Now firm 1 can threaten a post-entry output of k_1 while producing only $x_1 (\leq k_1)$ so long as entry does not occur. For firm 2, the relevant quantity is k_1, so it will stay out if $k_1 \geq B_1$, the amount 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. Its 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

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

This 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 $\mu_1 > M_1$.)

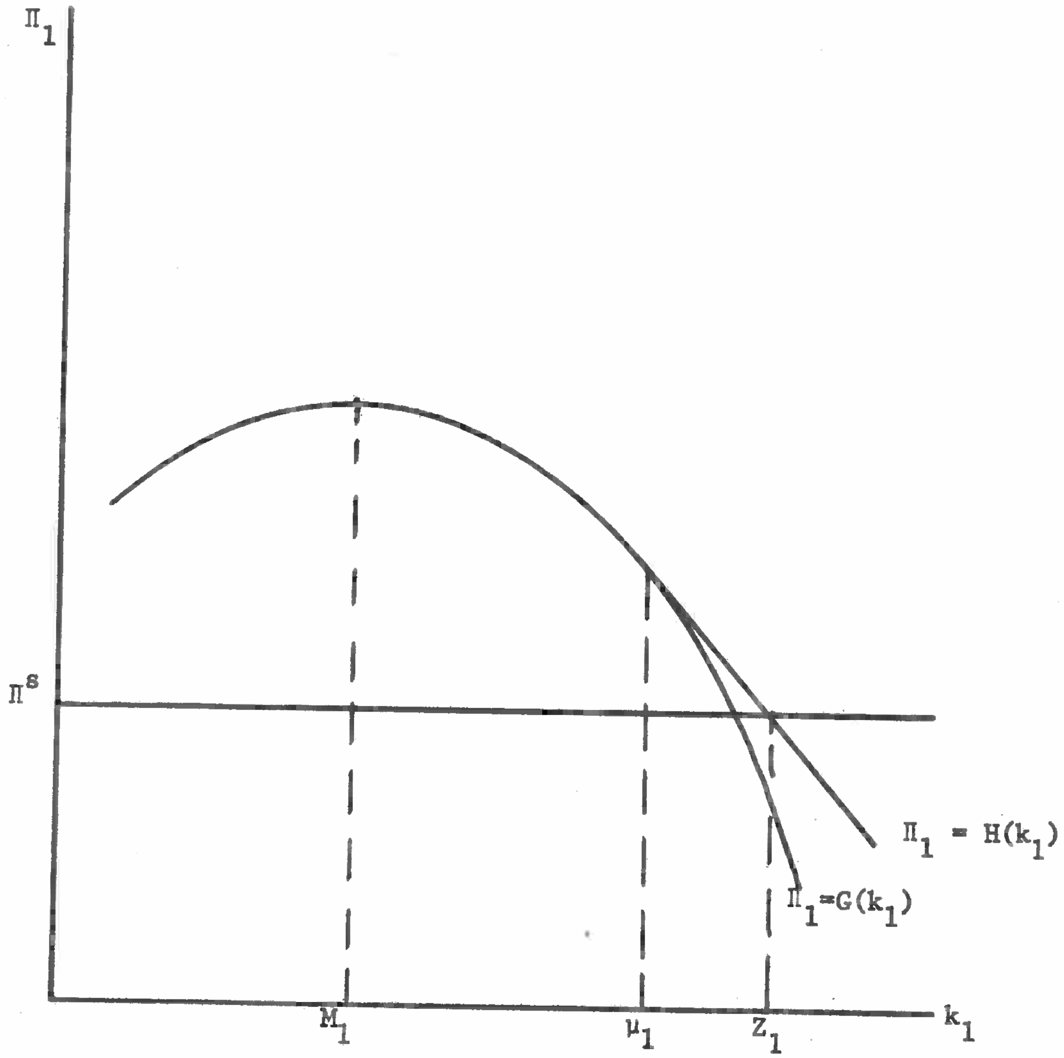
Now if k_1 is fixed at a value below μ_1, the choice of x_1 to maximize Π_1 above is at the limit of its permissible range, viz. $x_1 = k_1$. If $k_1 > \mu_1$, then x_1 is best set at μ_1 and spare capacity of $(k_1 - \mu_1)$ left. Correspondingly, after the best choice of x_1 is made, we can write Π_1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

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

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 (14). This is also shown in Figure 5 where it differs from $H(k_1)$. We see how the possibility of excess capacity has shifted 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 Stackelberg duopoly point. No excess capacity should exist at S, and its solution is as in Sections 2 and 3. Its intersection with $\Pi_1 = H(k_1)$ to the right of M_1 will be labelled 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 Spence 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, and 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 [11] analysis to make the question of entry endogenous in it.

5. Concluding Comments

In this paper I have suggested a general theoretical approach to the problem of entry of new firms comparable in size to existing ones. The approach does not take entry-prevention as a prior constraint, and 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 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 makes entry easier. This suggests that industrial economists should keep these two aspects distinct, instead of lumping them together into one vague concept of product differentiation as they usually do.

References

[1] Bain, J.S., Barriers to New Competition, (Cambridge, Mass : Harvard University Press) 1956.

[2] Dasgupta, P.S. and E. Maskin, "The existence of economic equilibria : continuity and mixed strategies", manuscript, 1977.

[3] Fellner, W., Competition Among the Few, (New York : Knopt) 1949

[4] Fisher, F.M., "New developments on the oligopoly front : Cournot and the Bain-Sylos analysis", Journal of Political Economy, 67 (4), August, 1959, pp. 410-413.

[5] Needham, D., "Entry barriers and the non-price aspects of firms' behaviour", Journal of Industrial Economics, 25 (1), September, 1976, pp. 29-43.

[6] Osborne, D.K. "On the rationality of limit pricing", Journal of Industrial Economics, 22 (1), September 1973, pp. 71-80.

[7] Roberts, J. and H. Sonnenschein, "On the existence of Cournot equilibria without concave profit functions", Journal of Economic Theory, 13 (1), August 1976, pp. 112-117.

[8] Scherer, F.M., Industrial Market Structure and Economic Performance, (Chicago : Rand-McNally) 1970.

[9] Spence, A.M., "Entry, investment and oligopolistic pricing", Bell Journal of Economics, 8 (2), Autumn 1977, 534-544.

[10] Waterson, M.J. "Osborne and limit pricing", manuscript, 1977.

[11] Williamson, O.E., "Selling expense as a barrier to entry", Quarterly Journal of Economics, 77 (1), February, 1963, pp. 112-128.

Footnotes

* This paper grew out of discussions with Michael Waterson during the Summer Research Workshop on Oligopoly Theory at Warwick University during July 1977, and benefitted greatly from his very detailed comments on a previous draft. I am also grateful to John Cubbin, Geoffrey Heal, Paul Joskow, Michael Spence, and especially, Victor Norman for their comments.

1 Scherer's formal model allows scale economies and differentiated products, but he misses all the special consequences of these.

2 See Scherer [8, ch.8] for an exposition and further references.

3 For an independent critique of Osborne's work, see Waterson [10].

4 To simplify the notation, Q_1 will denote both the point indicated, and the value of x_1 there. Similarly for other points along the axis. The quantities at N and S will be written as (N_1, N_2) and (S_1, S_2) respectively.

5 For further examination of existence of Nash equilibria, including possibilities of mixed strategy solutions, see Dasgupta and Maskin [2].

6 It is not meant (and generally not true) that the prospective entrant prefers to be the follower, but merely that given the fact that the established firm is already there and is following a quantity-setting strategy, the prospective entrant has no choice in the matter.

7 See Bain [1, ch. 4] and Needham [5].