THE JOURNAL OF INDUSTRIAL ECONOMICS 0022-1821 $2.00
Volume XXXIII March 1985 No. 3

TACIT COLLUSION, FREE ENTRY AND WELFARE

JAMES A. BRANDER AND BARBARA J. SPENCER*

I. INTRODUCTION

THE effect of collusion under free entry has a long history of analysis in economics, yet many of the central issues remain unresolved. Of particular interest is the interaction between the number of firms in the industry and the degree of collusion in the industry. There are two relationships that might reasonably exist between the number of firms and collusion. On one hand, collusion is made easier if the number of firms is small, but on the other hand, an industry that successfully colludes will earn short run above normal profits and attract new entry until profits are driven to zero.

In this paper we model the interaction between these two forces and present, in a simple diagram, an illustration of the determination of equilibrium entry and collusion. We are also concerned with the implications of such a structure for public policy. In particular, direct anti-collusive policies and entry fees are examined.

One preliminary point to address is why collusion should occur at all if profits are going to be driven to zero in any case.[1] First of all, it is only marginal firms that must earn zero profits, so if there are asymmetries between firms, inframarginal firms might hope to gain from collusion even in the long run. Our analysis, however, focusses on the case of symmetric firms: all firms earn zero profits in the long run. Even in this case the profit that can be earned in the short run implies that the present discounted value of an increase in the degree of collusion is positive. Furthermore, once a collusive zero profit equilibrium is reached, industry participants will want to maintain collusion so as to avoid the short run losses that would result from a breakdown in collusive arrangements. (This is similar to the "transitional gains trap" described in Tullock [1975]).

Certainly a fairly large class of industries seem to be characterized by partial collusion and free entry. Providers of professional services are an important example. There is a wide variation in the ability and income of different providers of the same professional services; nevertheless, in most professions there are individuals at the margin earning zero economic profit. Moreover in many countries professionals such as lawyers and dentists are

* We would like to thank two anonymous referees for their very helpful comments, which led to improvements in the paper, and for correcting a number of small errors in the paper.

[1] For example, Dewey [1982] asserts that, although collusion has been observed in such low profit industries as barber shops or hot dog stands in midtown Manhattan, it cannot be aimed at increasing monopoly rents because free entry keeps monopoly profits equal to zero.

strongly discouraged or even prevented by their professional organizations from advertising prices. Partly because of this, prices set by different suppliers tend to be similar, particularly when the association provides "guidelines". That such prices are not competitive is apparent if some suppliers are underemployed in the sense that they would like to work more at the going price, as often seems to be the case with lawyers, for example.

The subtle and complex bargaining that underlies tacit collusion in professional service organizations and other collusive structures is a difficult process to model accurately. We use the conjectural variation model as a relatively simple representation of the possible outcomes of such collusive interaction. If one is to focus on the relation between free entry and collusion some simplification of (or abstraction from) the actual collusive process is necessary.

The principal feature of the conjectural variation (or "generalized Cournot") model is that each firm forms an expectation concerning how the rest of the industry would vary its output in response to a change in own output. This expectation is the "conjectural variation." In principle, the conjectural variation would vary as output levels vary (see, for example, Boyer and Moreaux [1983]) and could be different for different firms, but a normal simplification is to assume that the conjectural variation, λ, is fixed and the same for all firms. We adopt this simplification. In this case the conjectural variation is, for any fixed number of firms, a single structural parameter reflecting the degree of output restriction in the industry.

There are at least two interpretations of the link between behaviour and the conjectural variation parameter. One interpretation is that the conjectures about other firms' behaviour are not taken literally.[2] Instead, the value of the conjectural variation, which is associated with a particular price and industry output given the number of firms, is interpreted as a proxy for or a representation of the level of tacit (or explicit) collusion in the industry.[3] Alternatively, even with tacit collusion, the conjectural variation, λ, may be the literal expectation held by firms. If λ exceeds 1, each firm expects to be punished if it raises output, in the sense that the rest of the industry will also raise output. Tacit partial collusion can be maintained if such expectations are held. The use of the conjectural variation parameter to reflect collusion is a useful but not uncontroversial simplification.

A second simplification to which attention should be drawn is that the conjectural variation model is a static representation of the actual complicated dynamic model that would be needed to capture the real time action

[2] In a fully non-cooperative setting where the conjectural variation parameter is taken as a firm's "best guess" concerning the expected behaviour of other firms, it has been argued that firms would revise their conjectures until they coincided with the actual responses that would occur in the event of a change in own output. This is the "consistent conjectures" approach. See Bresnahan [1981], Perry [1982], Kamien and Schwartz [1983].

[3] See Seade [1980]. See also Thompson and Faith [1981] on the selection of optimal behavioural rules with tacit collusion.

and reaction structure of imperfectly competitive firms. A recent paper by Kalai and Stanford [1983] explicitly analyzes conjectural variations in a repeated game framework, showing that various different conjectural variations can be maintained as "credible" equilibria.[4]

In any case, the important point is that a particular value of λ is associated with a particular price and industry output given the number of firms. A conjectural variation of zero is associated with pure competition, while a conjectural variation equal to the number of firms gives rise to the fully collusive level of output. As shown in the paper, for any given number of firms, higher levels of λ induce higher prices and lower levels of industry output. Use of the conjectural variation model makes it possible to examine the interaction between the number of firms and the level of collusion.[5]

In section II we set out the conjectural variation model and briefly derive results concerning the effects on industry and firm output of changes in the conjectural variation parameter and of entry or exit. In section III the combinations of numbers of firms and levels of collusion that allow zero profit are characterized. Each such point is a possible long run equilibrium. The actual equilibrium is then determined by introducing a schedule reflecting how the ease of collusion depends on the number of firms. The equilibrium is illustrated in a simple diagram.

Section IV is devoted to output, price, and welfare effects of government policies which reduce tacit collusion, and section V develops the second best contour, which shows how the optimal number (and scale) of firms varies with the degree of tacit collusion.

One important aspect of some industries characterized by free entry is that existing firms manage to impose entry costs on new entrants. (Consider, for example, professional service organizations in which large initial costs in the form of high opportunity costs incurred in qualifying for "certification" have a major impact on the entry decision.) Entry is still free in the sense that no arbitrary limit is set on the number of firms, and entry will take place until the profit of new entrants is driven to zero. In section VI we examine such entry costs, and also consider government imposed license fees. Section VII contains concluding remarks.

[4] There is an extensive recent literature on modelling oligopolies as repeated games in an attempt to capture the dynamic structure. Shubik [1982] is a good pedagogical source. Kalai and Stanford [1983] is the only paper (apart from a related paper by the same authors) that explicitly deals with trying to reconcile the conjectural variation and repeated game approaches. Their work is, however, confined to the 2 firm case. Incorporating free entry would be a major undertaking. Porter [1983] has a repeated game structure with uncertainty in which collusive outcomes are supported by credible finite punishment strategies, and has the attractive feature, in contrast to most of the repeated game literature, that output during cooperative phases is increasing in the number of firms. This result is broadly consistent with the model presented in this paper.

[5] The possibility of partial collusion is important. It has been argued (see Dewey [1982] and Patinkin [1947]) that such a situation would not be a true equilibrium; the collusive firms would form a profit maximizing cartel which would operate plants so as to minimize unit costs of production and use side payments to share profits. It seems to us that with large numbers of firms the organization of such a cartel would be extremely difficult.

II. THE MODEL

There are n identical firms, each with profit π, output y, variable cost function $c(y)$ and fixed cost F. Each firm maximizes expected profit π subject to its conjecture $dY/dy = \lambda$, where Y is industry output.

$$(1)\qquad \pi = yp(Y) - c(y) - F$$

The first and second order conditions for maximization of π are respectively,

$$(2)\qquad \pi_y = p - c' + yp'\lambda = 0$$

$$(3)\qquad \pi_{yy} = 2p'\lambda - c'' + yp''\lambda^2 < 0$$

where primes or subscripts denote derivatives. From (2) it is immediate that $\lambda = 1$ yields the Cournot model and $\lambda = 0$ yields pure competition. The following conditions are necessary and sufficient for local stability with identical firms. (See Seade [1980, p. 483].)

$$(4)\qquad p'\lambda - c'' < 0$$

$$(5)\qquad \alpha \equiv p'\lambda - c'' + n(p' + yp''\lambda) < 0$$

First we examine the effect of changes in n and λ on the output of the industry and on the profit of an individual firm.[6] Because firms are identical actual equilibrium changes have the property that $dY/dy = n$ in contrast to the belief held by firms that $dY/dy = \lambda$. Consequently comparative static results are related to α rather than to second order condition (3).

Totally differentiating (2) with respect to y and n yields

$$(6)\qquad \alpha\, dy + (\partial\pi_y/\partial n)\, dn = 0$$

Since $\partial\pi_y/\partial n = p'y + y\lambda p''y$ it follows that

$$(7)\qquad y_n \equiv dy/dn = -y(p' + \lambda yp'')/\alpha$$

and

$$(8)\qquad Y_n \equiv dY/dn = y + ny_n = y(1 - n(p' + \lambda yp'')/\alpha)$$
$$= y(\lambda p' - c'')/\alpha > 0$$

From (8) industry output must expand with entry if stability conditions (4) and (5) hold. The effect of entry on profit is obtained by differentiating (1):

$$(9)\qquad \pi_n = yp'Y_n + (p - c')y_n$$

$$(10)\qquad = -(p - c')(Y_n - \lambda y_n)/\lambda$$

using $yp'\lambda = -(p - c')$ from (2).

[6] We treat the integer variable n as though it were continuous in taking derivatives and for the comparative statics. This approximation is, of course, a problem only if n is very small.

From (7), (8) and (3), we have $Y_n - \lambda y_n = y\pi_{yy}/\alpha$. Therefore

$$(11) \qquad \pi_n = -(p - c')y\pi_{yy}/a\lambda < 0$$

From (11), the second order and stability conditions, each firm's profit at the Nash equilibrium declines with entry. The effect of λ on Y and π is obtained using a similar methodology. From (2)

$$(12) \qquad \alpha\, dy + (\partial\pi_y/\partial\lambda)\, d\lambda = 0 \qquad \text{where } \partial\pi_y/\partial\lambda = p'y$$

so

$$(13) \qquad y_\lambda = -p'y/\alpha < 0$$

and

$$(14) \qquad Y_\lambda = -np'y/\alpha < 0$$

Higher values of λ are associated with less output and a higher price, other things equal. As for profit, from (1)

$$(15) \qquad \pi_\lambda = yp'Y_\lambda + (p - c')y_\lambda$$

$$(16) \qquad = -(p - c')(n - \lambda)y_\lambda/\lambda \text{ using (2) and } Y_\lambda = ny_\lambda$$

π_λ is positive provided n exceeds λ and price exceeds marginal cost. The joint profit maximum ($\pi_\lambda = 0$) is achieved when $\lambda = n$.

Summarizing the results yields the following proposition.

Proposition 1: (i) Given the level of collusion, entry causes industry output to rise and the profit of each firm to fall. (ii) Given the number of firms, an increase in the level of collusion causes industry output and output per firm to fall and profit to rise.

The effects of entry in this model are available in Seade [1980] and Perry [1984], and Seade has some brief comments about the effect of changes in λ on profit and output. Our derivation of these results is slightly different from and more streamlined than earlier work. The results are required for the analysis of industry equilibrium and welfare effects.

III. ZERO PROFIT EQUILIBRIUM

Because changes in n and λ (the "collusion parameter") affect profit independently, there is a locus of n, λ combinations which allow zero profit for each firm. Thus $\pi(n, \lambda) = 0$ implicitly defines $\lambda = f(n)$: the zero profit locus. We assume that only points on this locus are candidates for long run equilibrium under free entry.

There are several concerns one might have about this assumption, most of which are related to the perceptions of potential entrants. In particular, if there are only a few firms in the industry, a potential entrant might reasonably expect to have a large negative impact on profit per firm, should it enter.

Thus existing profits might be positive because profits with further entry are expected to be negative. Secondly, there is the awkward issue of whether potential entrants should have the same sort of expectations as firms already in the industry, or whether they should have perfect foresight concerning the effects of entry, or some other kind of expectations. Other possibilities include Bertrand expectations, as in the contestable markets approach (see Baumol, Panzar and Willig [1982]), or Cournot expectations.

Our specification is as follows. We assume that there is a large number of firms in the industry so that the entry of any one firm has a negligible impact on profit per firm. The exact perception held by firms is then not particularly important, since the most plausible candidates approximate each other. One perception that is in the spirit of the paper is that entrants understand that if they enter they will join an implicitly collusive process governed by conjectural variation $\lambda(n)$; that is, an entrant correctly foresees that its entry will lower industry profit, but only by a negligible amount since n is large. The zero profit equilibrium then becomes a convenient approximation.[7]

In any case, differentiating $\pi(n, \lambda) = 0$ yields,

$$\pi_n \, dn + \pi_\lambda \, d\lambda = 0 \tag{17}$$

Substituting for π_n and π_λ from (10) and (16) gives

$$f'(n) = d\lambda/dn \Big|_{\pi=0} = -(Y_n - \lambda y_n)/(n - \lambda)y_\lambda \tag{18}$$

Provided $n > \lambda$, $d\lambda/dn$ must be positive since $y_\lambda < 0$ and $Y_n - \lambda y_n = y\pi_{yy}/\alpha > 0$. Greater tacit collusion is required to absorb more firms earning zero profit (provided price is below the joint profit maximizing level.) For example, in the case of linear demand, $p = a - bY$, and constant marginal cost, the equation for the zero profit locus[8] is

$$n = -\lambda + (a - c')\sqrt{\lambda/Fb} \tag{19}$$

From expression (19) it follows that, for the case of linear demand and constant marginal cost, the competitive outcome, $\lambda = 0$, is consistent only with $n = 0$. This is a manifestation of the familiar observation that, if marginal cost is below average cost, the competitive pricing rule, price = marginal cost, implies negative profits if $n > 0$. Therefore the trivial ($n = 0$) equilibrium is the only competitive zero-profit equilibrium: a non-trivial zero-profit equilibrium must be imperfectly competitive.

[7] The problems of small numbers and possible asymmetric expectations between entrants and incumbents is not, of course, unique to our model and has been much discussed, particularly in connection with the contestable markets hypothesis, but also whenever the Chamberlinian zero profit equilibrium is formalized.

[8] With linear demand, profit maximization implies $y = (a - c')/b(n + \lambda)$. The relationship between n and λ at zero profit, $n = -\lambda + (a - c')\sqrt{\lambda/Fb}$ is obtained by substituting this value of y into $\pi = 0$ and assuming c' constant.

With "U-shaped" average cost curves, on the other hand, there is some positive value of n at which $\lambda = 0$ along $\lambda = f(n)$. The number of firms at this "purely competitive" zero profit point depends on the technology and the size of the market and is not necessarily large. In general, for the zero profit equilibrium to exist with $\lambda > 0$ average cost must be downward sloping at the solution because average cost must be tangent to perceived demand, which slopes downward.

As illustrated in Figure 1, the 45° line is the locus of joint profit maximizing points ($n = \lambda$). From (18), the zero profit locus is vertical at the point $n = \lambda$. Line OA passing through the origin represents the linear example. The shape of the contour reflects the fact that in this case $f''(n)$ is strictly positive if $n > \lambda$. For completeness we continue the contour to the left of the 45° line ($n < \lambda$), but it should be noted that this region is not economically relevant. Line NB illustrates a zero profit contour for the case of "U-shaped" average cost curves. N is the number of firms at the competitive outcome.

We examine the fully symmetric case. Clearly providers of professional services do not all have the same costs and differences in ability and training lead to some differentiation in the product. Fortunately, generalization to asymmetric firms does not affect the nature of the problem, but it does substantially complicate the analysis. In the case of cost asymmetry and asymmetric product differentiation the zero profit contour is defined with reference to the marginal firm. The convenient assumptions of continuity and differentiability are harder to justify. With asymmetric conjectures and product differentiation, there is a conjecture λ_{ij} associated with each pair of firms i and j and the analysis of output and price affects would have to be carried out using some average or index of firm perceptions.

In this paper, equilibrium with free entry is defined as a situation in which the (symmetric) firms earn zero profit in the long run, given their conjectures, λ. New firms, once they enter, will have the same expectations as established firms. For example, with tacit collusion via a professional organization, new

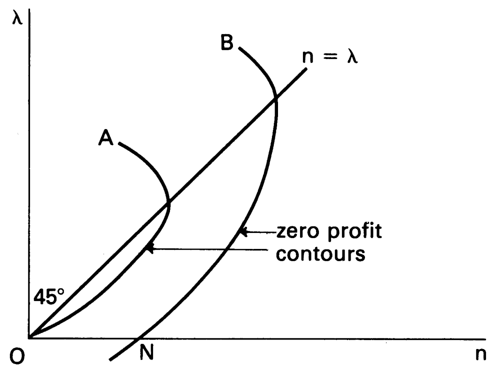

FIGURE 1

professionals joining the system follow the same pricing behaviour as will existing professionals after entry. Nevertheless, the maintenance of the tacit agreement becomes more difficult as entry occurs, which is reflected by a decrease in the value of λ.

To determine which point on the zero profit locus is the industry equilibrium with free entry, consider the way the collusion parameter, λ, is likely to vary with entry. This relationship is denoted $\lambda = g(n)$. More precisely, $g(n)$ shows the λ that would emerge if the number of firms were fixed at n while λ adjusted. In the case of monopoly, since firm output is industry output, $\lambda = 1$. In general, profit-maximizing collusion is achieved if $\lambda = n$ as entry occurs. Since a greater number of firms makes collusion more difficult, we assume that $\lambda = g(n)$ falls below the 45° line after some n ($g(n)$ may be coincident with the 45° line for some n) and eventually decreases in n, until the axis (competitive expectations) is reached. This is illustrated by $g(n)$ in Figure 2. For low values of n, $g(n)$ is in a region of positive profits to the left of the zero profit contour $\lambda = f(n)$. Industry equilibrium with free entry (if it exists) is given by the intersection of f and g. Two possible positions of the zero profit contour, f_1 and f_2, are illustrated. f_1 intersects g at A so that at the free entry equilibrium there are more firms and price is higher than at the competitive point N_1. If the zero profit contour is f_2, $g(n)$ meets the n axis ($\lambda = 0$) with firms earning positive profits. Additional entry then moves the industry from short run purely competitive equilibrium to the zero profit long run equilibrium at N_2.

The function $\lambda = g(n)$, expressing the conjectural variation as a function of the number of firms, is simply imposed exogenously. One would prefer to derive $g(n)$ from the underlying technological and strategic structure. This is, however, a substantial problem and is beyond the scope of the present paper. One could, however, look to the repeated game literature for some sort of justification. Porter [1983] develops a repeated game structure in which the

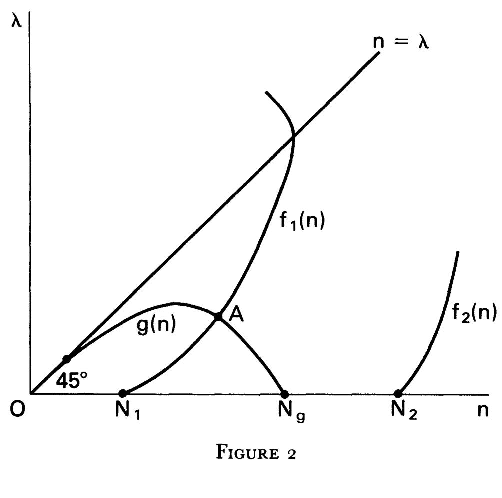

FIGURE 2

level of collusion that firms are able to maintain, as reflected by the output level, falls as the number of firms in the industry rises. This is broadly consistent with our assumption concerning $g(n)$. Our assumption is also consistent with the "free rider" problem discussed by Stigler [1974] and Kalai and Stanford [1983]. Specifically, in any implicitly collusive arrangement each firm bears some costs should it be called upon to punish deviant behaviour. With many firms, output restriction by any firm is attributable mainly to the enforcement effort of others: each firm has a tendency to free ride. This effect is stronger as the number of firms rises, making the achievable strength of collusion fall as n rises.

It could be argued that the "free rider" problem might lead to other formulations. For example, a referee suggested that, with large numbers of firms, collusion would be organized by trade associations, who in turn were financed by firms in the industry. Through this avenue the cost of collusion might show up as a part of fixed costs for each firm, hence λ would be an increasing function of F. This strikes us as a potentially fruitful modelling approach, although quite different from our model. In any case, neither Porter's argument nor the free-rider issue is formally connected to our model, so we would not wish to overstate the case. Nevertheless, we feel that these arguments do lend rough support to the assumed nature of the "ease of collusion" function.

IV. WELFARE AND THE ZERO PROFIT LOCUS

Now consider the effect of government policy on the placement of the $\lambda = g(n)$ schedule. If policy acts to make tacit collusion more difficult for any given number of firms and there is free entry, the industry moves down the zero profit locus, reducing the equilibrium number of firms. For example, one such policy might be for the government to ensure that members of professional associations such as lawyers and dentists can advertise the price of their services. This reduces the power of the association to maintain prices. This section is concerned with the welfare implications of such policies.

Assuming that partial equilibrium analysis is appropriate, consumer utility can be approximated by the form $U = u(Y) + m$ where m is expenditure on a numeraire commodity. Let $S(n, \lambda)$ be the sum of consumer surplus plus profit given n and λ. Since the marginal utility of income is constant and equal to one,

$$(20) \qquad S(n, \lambda) = u(Y) - n(c(y) + F)$$

Using $p = du/dY$, $ny_n = Y_n - y$ and $Y_\lambda = ny_\lambda$, we have

$$(21) \qquad S_n(n, \lambda) = (p(Y) - c'(y))Y_n - (c(y) + F - c'(y)y)$$

$$(22) \qquad S_\lambda(n, \lambda) = (p(Y) - c'(y))ny_\lambda$$

where S_n and S_λ are partial derivatives.

Zero profit means price equals average cost, $p = (c(y) + F)/y$, and (21) becomes

$$(23) \qquad S_n(n, \lambda) = (p - c')(Y_n - y) = (p - c')(ny_n)$$

Along the zero profit locus $\lambda = f(n)$. Therefore, using (22) and (23), we have

$$(24) \qquad dS/dn\Big|_{f(n)} = S_n + S_\lambda f'(n) = n(p - c')(dy/dn)\Big|_{f(n)}$$

where

$$dy/dn\Big|_{f(n)} = y_n + y_\lambda f'(n)$$

The effect on welfare of an increase in the number of firms along the zero profit locus depends on the sign of $(dy/dn)|_{f(n)}$, that is, on whether the output of a typical firm rises or falls along the contour. Substituting (18) for $f'(n)$ and some rearrangement using $Y_n = ny_n + y$ yields

$$(25) \qquad dy/dn\Big|_{f(n)} = -y/(n - \lambda)$$

which is negative, since $n > \lambda$. Since in this case (from (18)), $f'(n) > 0$, additional firms earning zero profit can only be accommodated with a higher value of λ (greater tacit collusion) and a lower level of output for each firm.[9] Therefore, from (24), welfare falls along the zero profit locus as n increases with λ. Greater numbers of firms do not offset the anti-competitive effect of increased collusion along $f(n)$. The industry performs better at low levels of n and λ and welfare is maximized at $\lambda = 0$. Government policies which reduce collusion in equilibrium necessarily increase welfare.

Since profit is zero along the contour the changes in welfare are just changes in consumer welfare and therefore can be related to price changes. The Nash equilibrium levels of Y and p along the zero profit contour can be represented by $Y = Y(n, f(n))$ and $p = p(Y(n, f(n)))$. The effect on price of an increase in the number of firms along $\lambda = f(n)$ is

$$(26) \qquad dp/dn\Big|_{f(n)} = p'(dY/dn)\Big|_{\pi = 0}$$

Since $dY/dn|_{f(n)} = n(dy/dn)|_{f(n)} + y$, using (25) yields

$$(27) \qquad dY/dn\Big|_{f(n)} = -\lambda y/(n - \lambda)$$

If $n > \lambda > 0$, from (27) and (26), industry output falls and price rises along the zero profit contour. These results are summarized in Proposition 2.

[9] This model captures the argument made by Machlup [1952] that collusion with free entry causes firms to move up their average cost curves, increasing the average cost of production.

Proposition 2: Provided $n > \lambda > 0$, both the output of each firm and total industry output are decreasing and price is increasing as both n and λ increase along the zero profit contour. Industry output and welfare are maximized at $\lambda = 0$.

To see the welfare change directly as a function of the price change, first substitute (27) into (26) and use first order condition (2) to obtain

$$(28) \qquad dp/dn\Big|_{f(n)} = (p - c')/(n - \lambda)$$

Then substituting (25) into (24) and using (28), we obtain

$$(29) \qquad dS/dn\Big|_{f(n)} = -Y(p - c')/(n - \lambda) = -Y\, dp/dn\Big|_{\pi = 0}$$

The change in welfare along the zero profit contour is just the marginal loss in consumer surplus from the rise in the price of Y.

V. THE SECOND BEST CONTOUR

In markets with economies of scale it has long been recognized that entry of firms into the industry may reduce welfare. (See von Weizsäcker [1980] for an analysis of the Cournot case.) An additional firm normally increases industry output, but it may also cause production to occur at a higher average cost. Perry [1984] has illustrated this in the context of the conjectural variation model. He shows that if entry reduces the output of each firm, $y_n < 0$, then welfare would be increased by reducing the number of firms below the level at the free entry equilibrium. This can be seen from our equation (24). As in Perry [1984], given the behaviour of the firms (characterized by the value of λ), the (second best) optimal number of firms, n^*, is defined by $S_n(n^*, \lambda) = 0$, where S_n is given by (21).

For the purpose of illustrating this effect in our diagram and for the welfare results in section VI, we extend this analysis to show how the second best optimal number of firms varies with the degree of tacit collusion. We refer to the schedule $n^* = n^*(\lambda)$, defined implicitly by $S_n(n^*, \lambda) = 0$, as the second best contour. It relates the (second best) optimal value of n to λ. From total differentiation of $S_n(n, \lambda) = 0$ with respect to n and λ the slope of this contour is

$$(30) \qquad dn^*/d\lambda = -S_{n\lambda}/S_{nn}$$

Under most demand and cost conditions one would expect the second best optimal number of firms to increase with λ. With greater tacit collusion, the output increasing effect of a larger number of firms becomes more important. This is indeed the case if economies of scale are caused by a fixed cost, F, (with constant marginal cost) and if demand is linear. The equation for the second best welfare contour[10] with $p = a - bY$ is then

$$(31) \qquad n^* = -\lambda + (\lambda^2(a - c')^2/Fb)^{1/3}$$

with slope $dn^*/d\lambda = (2n^* - \lambda)/3\lambda$. Since with linear demand, $y_n < 0$, from (7), the second best welfare contour lies to the left of the zero profit locus and is illustrated in figure 3.

If demand is sufficiently convex ($p'' > p'/\lambda y > 0$) from (7), y_n may be positive and the second best contour may lie to the right of the zero profit contour. However, we take $y_n < 0$ as the normal case, leading to a second best contour along which firms' profits would be positive, unless lump sum entry or license fees are imposed.

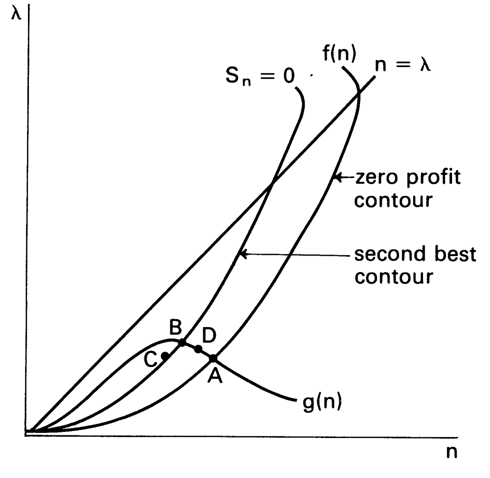

FIGURE 3

[10] With linear demand, $p = a - bY$, and constant marginal cost, from (21), the second best contour becomes,

$$S_n(n^*, \lambda) = (a - c' - bn^*y)Y_n - F = 0$$

The first order condition (2) for profit maximization implies $y = (a - c')/b(n + \lambda)$. Also from (5) $\alpha = p'(\lambda + n)$ so that from (8), $Y_n = \lambda y/(\lambda + n)$. Substituting y and Y_n into $S_n = 0$, we obtain

$$(\lambda + n^*)^3 = \lambda^2(a - c')^2/Fb$$

which implies $n^* = -\lambda + (\lambda^2(a - c')^2/Fb)^{1/3}$

$$\begin{aligned} dn^*/d\lambda &= -1 + (1/3)(\lambda^2(a - c')^2/Fb)^{-2/3}(2\lambda(a - c')^2/Fb) \\ &= -1 + 2(\lambda + n^*)^{-2}(\lambda + n^*)^3/3\lambda \\ &= (2n^* - \lambda)/3\lambda \end{aligned}$$

VI. WELFARE IMPLICATIONS OF ENTRY COSTS AND LICENSE FEES OR SUBSIDIES

It is sometimes suggested that professional associations set conditions so that it is more difficult and expensive for a professional to obtain a license than is strictly necessary based on professional standards. Long periods of study or apprenticeship before certification are particularly common entry barriers. The natural reaction of economists is that such impediments to entry are welfare reducing. However, as pointed out by von Weizsäcker [1980], in connection with a free entry Cournot model, it is conceivable that restricting entry could improve welfare. In this section such a possibility is examined using our model. We find that entry barriers imposed by the industry, and which consume real resources, definitely lower welfare, but government license fees will improve welfare if output per firm is increased.

So far we have treated F as a fixed cost: a cost which must be borne if any output at all is to be produced, but which is implicitly avoidable if no output is produced. Some license fees are well-modelled as increases in F: for example a license fee may be charged every year. True entry fees, on the other hand, such as the costs of law school for a lawyer, are more like sunk costs inasmuch as they are irreversible whether or not output is produced. However, if one is careful about the conceptual experiment being considered, these sunk costs can also be modelled as increases in F. First of all, the model is a static one period model, so the minor complication that a sunk cost must be amortized over the life of the industry rather than applied to a single period is avoided. (This is easy to deal with in any case.) The slightly more subtle point is that we do not consider the conceptual experiment of raising F for potential entrants while allowing incumbents to remain in the industry. Instead we compare two different equilibrium configurations, one in which there is a low fixed cost F for all participants and one in which there is a high F, including the entry cost or license fee, for all participants.

Since the entry cost, F, is a fixed cost from the viewpoint of an individual firm, its only direct positive effect is to reduce the equilibrium number of firms. To determine this effect it is useful to first consider the implications of a change in F for the zero profit locus implicitly defined by $\pi(n, \lambda; F) = 0$ (given by (1)). Differentiating $\pi = 0$ with respect to n and F holding λ fixed, and using $\pi_F = -1$ and $\pi_n < 0$ (from (11)) we obtain

$$(32) \qquad \partial n/\partial F \bigg|_{\tau=0} = 1/\pi_n < 0$$

The higher cost of entry shifts the zero profit locus to the left reducing n for a given value of λ. The equilibrium number of firms is determined by the intersection of this (shifted) locus with $\lambda = g(n)$. If $g'(n) < 0$, a reduction in the number of firms increases the level of the tacit collusion in equilibrium. From total differentiation of $\pi(n, g(n); F) = 0$ we obtain the effect of F on the equilibrium number of firms allowing λ to vary. If $g'(n) \leq 0$, then

$$(33) \qquad dn/dF\Big|_{\pi=0} = 1/(\pi_n + \pi_\lambda g'(n)) < 0$$

If $g'(n) < 0$, then an increase in F reduces the equilibrium number of firms, but by less than would occur if $g'(n) = 0$.

The decrease in the number of firms due to a higher entry cost may increase the equilibrium level of output per firm. This is easily shown by recognizing that in zero profit equilibrium $y = y(n, \lambda; F)$ where $\lambda = g(n)$. Therefore

$$(34) \qquad dy/dF\Big|_{\pi=0} = (y_n + y_\lambda g'(n))(dn/dF)\Big|_{\pi=0}$$

which from (33) is positive if $g'(n) = 0$ and $y_n < 0$. If $g'(n) < 0$, output per firm may still rise if $y_\lambda g'(n)$ is small.

Despite this, if $g'(n) \leq 0$ industry output always falls and price rises with an increase in F. Since in equilibrium $Y = Y(n, \lambda: F)$, from (8), (14) and (33),

$$(35) \qquad dY/dF\Big|_{\pi=0} = (Y_n + Y_\lambda g'(n))(dn/dF)\Big|_{\pi=0} < 0$$

$$(36) \qquad dp/dF\Big|_{\pi=0} = p'\, dY/dF\Big|_{\pi=0} > 0$$

Now suppose that F is a real resource cost imposed by the industry. Since profit is zero, the change in welfare is then due only to changes in consumer surplus. Since price rises, consumer surplus and welfare must fall with an increase in F.[11] These results are summarized in Proposition 3.

Proposition 3: Provided $g'(n) \leq 0$, an increase in the real cost of entry, F, increases price and reduces industry output and welfare at the zero profit equilibrium.

In addition, the second best contour $n = n^*(\lambda; F)$ implicitly defined by $S_n(n^*, \lambda; F) = 0$ is dependent on a real entry cost F as in (21). Since $S_{nF} = -1$ (from (21)) and $S_{nn} < 0$ from the second order condition for a welfare optimum, for any given value of λ, n^* is decreasing in F.

$$(37) \qquad dn^*/dF = 1/S_{nn} < 0$$

This implies that the second best contour (as well as the zero profit contour) is shifted to the left by an industry policy to increase the real cost F. In the context of figure 3, although an increase in the real entry cost reduces the equilibrium number of firms, it may not bring the equilibrium closer to the second best contour.

[11] This can be shown directly using the welfare function $S(n, \lambda; F)$ as given by (20).

On the other hand, the second best contour would be unaffected if, rather than a real cost F, the government imposed a license fee, T. The license fee reduces the profit of a typical firm, but this is just offset by the additional revenue collected by the government, which is included in total welfare.

$$(38) \qquad S(n, \lambda; F, T) = u(Y) - nc(y) - n(F + T) + nT$$

The welfare function therefore reduces to the same form as before (see (20)).

However, a license fee does affect welfare through its influence on the number of firms in equilibrium. A license fee has the same effect on profit and the equilibrium number of firms as the real entry cost F.

Therefore, expressions (33) and (34) describe the marginal effects of T on n and y, respectively, along the zero profit contour, just replacing F by T. Now consider the optimal value of T, which is characterized by setting the derivative of $S(n, \lambda; F, T)$ to zero, recognizing that $\lambda = g(n)$ in equilibrium.

$$(39) \qquad dS/dT\Big|_{\pi=0} = (S_n + S_\lambda g'(n))\, dn/dT\Big|_{\pi=0} = 0$$

Noting that $c(y) + F = py - T$ and using (21), we obtain

$$(40) \qquad S_n(n, \lambda)\Big|_{\pi=0} = (p - c')ny_n + T$$

Then, from (40) and (22), (39) becomes

$$(41) \qquad dS/dT\Big|_{\pi=0} = (n(p - c')(y_n + y_\lambda g'(n)) + T)(dn/dT)\Big|_{\pi=0} = 0$$

The optimal license fee is then (using (34) with T instead of F):

$$(42) \qquad T = -n(p - c')[(dy/dT)/(dn/dT)]\Big|_{\pi=0}$$

Since $dn/dT < 0$ from (33), Proposition 4 follows.

Proposition 4: At the zero profit equilibrium, the optimal license fee, T, is positive if and only if an increase in the license fee increases the equilibrium level of output of each firm.

If there is no change in λ with entry ($g'(n) = 0$), then from (34) and (42) the optimal license fee is $T = -n(p - c')y_n$, which is positive in the normal case of $y_n < 0$. With this value of T, S_n is zero from (40). This means that if tacit collusion is unaffected by the number of firms and $y_n < 0$ then the second best optimal license fee shifts the zero profit locus to the left so that it cuts the second best contour at the constant value of λ. Assuming $g(n)$ is horizontal this is illustrated in figure 3 by the move from point A to point C. In the unusual case in which $y_n > 0$, the optimal value of T is then a subsidy. If $g'(n) = 0$, it will also bring the industry to the second best contour, which in this case lies to the right of the zero profit locus.

On the other hand, if $g'(n) < 0$, from (41) the optimal license fee is lower but still positive (by Proposition 4) if it increases firm output in equilibrium. In this case from (40), the value of S_n at zero profit is negative indicating that the optimal license fee leaves the industry to the right of the second best contour (at a point such as D in figure 3). The policy takes account of the fact that a further cutback in the number of firms (as in the move from D to B in figure 3) will increase tacit collusion sufficiently to reduce welfare. The improvement in welfare from the increase in output per firm is more than offset by the loss from the increase in tacit collusion.

These results are summarized in Proposition 5.

Proposition 5: If $g'(n) = 0$, the optimal license fee, T, ensures that the zero profit equilibrium lies on the second best contour. If $g'(n) < 0$ and $T > 0$, then at the optimal fee, the number of firms exceeds the number on the second best contour, given λ.

One qualification to Propositions 4 and 5 raised by a referee is that license fees or certification costs might be used to exclude inefficient producers or low quality producers, leading to rather different and perhaps significant welfare affects. These possibilities are excluded by our model due to the assumption that firms are symmetric.

VII. CONCLUDING REMARKS

Freedom of entry is regarded as an important source of economic efficiency in market economies. In combination with purely competitive behaviour by firms, it leads to pareto-efficient outcomes. However, even industries with large numbers of sellers seem to appreciate the value of achieving some degree of collusion, despite free entry. As mentioned in the paper, this is perfectly rational behaviour consistent with pursuit of short run excess profits and avoidance of short run losses. Therefore it seems important to understand the positive and normative implications of the relationship between free entry and partial collusion.

The nature of this relationship can conveniently be understood within the context of the general conjectural variation model. We show that greater tacit collusion is required to absorb more firms earning zero profit, as reflected in a zero profit locus of combinations of the number of firms, n, and conjectural variation, λ. The idea that ease of collusion falls as entry occurs yields a second (n, λ) locus. Industry equilibrium occurs when both relations are satisfied. We are also able to characterize the second best optimum number of firms given the level of tacit collusion in the industry. There is a trade-off in that more firms lead to more competitiveness, but normally reduce output per firm so that firms move up their average cost curves.

The welfare analysis of impediments to entry yields some interesting results. First, not surprisingly, policies which render tacit collusion more difficult are in the public interest. Also attempts by the industry to increase the real cost of entry are welfare reducing. However, a positive government license fee, which simply taxes entrants rather than forcing them unnecessarily to use up real resources in entering, is normally welfare improving. The number of firms falls, reducing competitiveness, but output per firm rises, lowering average cost. The license fee is just a transfer and overall welfare rises.

ACCEPTED MARCH 1984

James A. Brander,
Faculty of Commerce and Business Administration,
University of British Columbia,
Vancouver, B.C. V6T 1Y8, Canada.

Barbara Spencer,
Department of Economics,
Boston College,
Chestnut Hill,
Massachusetts, 02167, U.S.A.

REFERENCES

BAUMOL, W., PANZAR, J. and WILLIG, R., 1982, *Contestable Markets and the Theory of Industry Structure* (Harcourt, Brace, Jovanovich, New York).

BOYER, M. and MOREAUX, M., 1983, "Consistent Versus Non-Consistent Conjectures in Duopoly Theory: Some Examples," *Journal of Industrial Economics* 32 (September), pp. 97–110.

BRESNAHAN, T. F., 1981, "Duopoly Models with Consistent Conjectures," *American Economic Review* 71 (December), pp. 934–45.

DEWEY, D., 1982, "Welfare and Collusion: Reply," *American Economic Review* 72 (March), pp. 276–81.

KALAI, E. and STANFORD, W., 1983, "Conjectural Variations Strategies in Dynamic Cournot Games with Fast Reactions," Northwestern University, Center for Mathematical Studies in Economics and Management Science, Discussion Paper No. 575 (September).

KAMIEN, M. and SCHWARTZ, N., 1983, "Conjectural Variations," *Canadian Journal of Economics* 16 (May), pp. 191–211.

MACHLUP, F., 1952, *The Economics of Seller's Competition: Model Analysis of Sellers' Conduct*, (John Hopkins Press, Baltimore).

PATINKIN, D., 1947, "Multiplant Firms, Cartels and Imperfect Competition," *Quarterly Journal of Economics* 61 (October), pp. 173–80.

PERRY, M. K., 1982, "Oligopoly and Consistent Conjectural Variations," *Bell Journal of Economics* 13 (Spring), pp. 197–205.

PERRY, M. K., 1984, "Scale Economies, Imperfect Competition and Public Policy," *Journal of Industrial Economics* 32 (March), pp. 313–33.

PORTER, R. H., 1983, "Optimal Cartel Trigger Price Strategies," *Journal of Economic Theory* 29 (April), pp. 313–38.

SEADE, J., 1980, "On the Effects of Entry," *Econometrica* 48 (March), pp. 479–89.

SHUBIK, M., 1982, *Game Theory in the Social Sciences: Concepts and Solutions* (The MIT Press, Cambridge, Massachusetts).

STIGLER, G. J., 1974, "Free Riders and Collective Action: An Appendix to Theories of Economic Regulation," *Bell Journal of Economics and Management Science* 5 (Autumn), pp. 359–65.

TULLOCK, G., 1975, "The Transitional Gains Trap," *Bell Journal of Economics* 6 (Autumn), pp. 671–78.

THOMPSON, E. A., and FAITH, R. L., 1981, "A Pure Theory of Strategic Behaviour and Social Institutions," *American Economic Review* 71 (June), pp. 366–80.

VON WEIZÄCKER, C. C., 1980, "A Welfare Analysis of Barriers to Entry," *Bell Journal of Economics* 11 (Autumn), pp. 399–420.