European Economic Review 19 (1982) 123–143. North-Holland Publishing Company

OLIGOPOLY AND WELFARE

A Unified Presentation with Applications to Trade and Development*

Avinash DIXIT

Princeton University, Princeton, NJ 08544, USA

Nicholas STERN

University of Warwick, Coventry CV4 7AL, UK

We develop a unified sequence of models to examine the determinants of price, output, profitability and concentration for different kinds of oligopolistic market. We relate various magnitudes of significance to judges of welfare and to policy makers, such as consumer surplus, profit, or added benefits of employment or trade surplus, to observable magnitudes such as the size and concentration of the market. It emerges that various appropriately formulated Herfindahl indices are useful in several of these relations. We attempt to present results in a way that is useful for empirical investigations and pay particular attention to applications in trade and development.

1. Introduction

Models of oligopoly that permit empirical evaluation of welfare have made substantial progress from Bain's important work on structure–performance relationships. This paper offers some further developments of the more modern theory and applies the framework to some relatively new questions.

The earliest models expressed a measure of performance, such as the profit rate, as a function of a measure of structure, such as a concentration ratio. The form of the functional relation was governed by the oligopolists' conduct. In fact the studies were largely *ad hoc*, the choice of the measure being dictated by data availability, and the specification of the functional form by the linear regression format. A further problem was the implicit assumption of exogeneity of the structure. Since conduct involves setting quantities or prices for outputs, it is hard to see how any theory of oligopoly could take a concentration index to be exogenous. Finally, the welfare implications of the performance measure were usually left very tenuous. Surveys and discussions of this work can be found in Weiss (1971), Scherer (1980, ch. 9) and Jacquemin and de Jong (1977, pp. 142–144).

*We are grateful to John Cable, Keith Cowling, Paul Stoneman and a referee for very helpful comments.

A significant step forward was to base the relationship on an explicit model of conduct, as in Stigler (1964) and Cowling and Waterson (1976). Such theorising pointed out preferred choices of the measures of structure and performance, usually respectively the Hirschmann–Herfindahl index and the price–cost margin. However, the issue of exogeneity was not considered, and the price–cost margin was not explicitly translated into a welfare measure. Encaoua and Jacquemin (1980) define a Lerner index of monopoly power for the industry as a function of similar indices for the individual firms. Axioms that establish a qualitative link between the industry Lerner index and the level of welfare restrict the functional form. By choosing different weighting schemes compatible with the axioms, we can express the industry Lerner index in terms of different concentration indices. However, the quantitative connection between such measures and welfare is still weak.

This last defect was partly remedied by Dansby and Willig (1979). They expressed the best conceivable local welfare improvement from an oligopolistic equilibrium in terms of a concentration index. Different assumptions about conduct led to the choice of a different index. Apart from its local nature, this approach has the shortcoming of providing only an upper bound for achievable welfare improvements. This is because the quantity changes in the direction of steepest ascent of welfare may not in fact be feasible using the available instruments.

In this paper we present a unified treatment with the same ultimate aims, namely to link performance and structure in oligopoly, but without these defects. First, our analysis is based on explicit models, so the relevant measures of structure and performance emerge naturally. Next, our welfare analysis is explicit and global, i.e., we compute the deadweight loss associated with the oligopolistic equilibrium directly instead of relying on a proxy like the price–cost margin. We do also obtain expressions in terms of this margin to facilitate estimation. Thirdly, although the relation between the two endogenous variables measuring structure and performance can be instructive, we keep in mind its limitations, and carry out comparative statics in terms of variables which appear more naturally as exogenous like demand and technology parameters or policy variables like tax rates.

We wish to emphasize two special aims. First, we have tried to cast the theory in terms of observables as far as possible, in order to improve its applicability. To this end, we have often used special functional forms and approximations. Secondly, and with greater novelty, we have attempted to extend the applications of the approach beyond its traditional confines of industrial organisation in a closed advanced economy. It is apparent that there are significant elements of oligopoly in contexts of international trade and developing countries. In extractive and manufacturing industries, and even in agriculture, there is often significant concentration due to the presence of transnational firms, governmental coordination of trading activities, or their being channelled through a small number of trading houses. The available literature studies the repercussions of international competition on domestic structure–performance relationships, e.g. Hitiris (1978) and Pagoulatas and Sorenson (1976), or on regulatory policy, e.g. White (1974) or Katrak (1979), but not oligopoly in international trade itself.

In such contexts, there are special features we have to take into account when evaluating welfare. In the usual treatment of a modern industry, it is implicitly assumed that the marginal cost in the market is a good measure of the social opportunity cost of resources. This is often less true in international trade due to inappropriate exchange rates, and in less developed countries due to domestic factor market distortions. Methods have evolved for evaluating shadow prices in such contexts where they differ from market prices. We show how the expressions for the deadweight loss can be modified to incorporate the necessary corrections.

A second issue concerns the appropriate model of conduct. Most conventional structure–conduct-performance studies are based on a generalised Cournot model with conjectural variations. This is a surprisingly rich model, and arguably the best available tool for studying industrial oligopoly. However, in some situations in developing countries, it may be more appropriate to think of large foreign firms as being in a position of leadership relative to a competitive fringe of small domestic firms. Therefore we examine a model of price leadership in section 3, following the conjectural variations model developed in the next section. Section 4 contains concluding remarks.

For the most part, we confine the formal model to one of quantity or price setting. Other variables such as advertising, or research and development expenditures, can be incorporated at the cost of greater algebraic complexity. We indicate some implications for these dimensions that are suggested by our simple basic model.

2. Conjectural variations

2.1. The model

We derive in this sub-section the equilibrium price and output of each firm in a conjectural variations model of an industry with a homogeneous product and a small number of firms. Once equilibrium has been calculated one can work out other measures, in particular concentration and welfare loss, fairly easily. We consider in subsequent sub-sections various applications, modifications and re-interpretations of the model. In sub-section 2.2 we introduce foreign owned firms, and in sub-section 2.3 we consider the problem of a firm choosing where to locate production. If market costs and prices do not reflect social opportunity costs then calculations of welfare losses have to be modified: this problem is examined in sub-section 2.4 using shadow prices. A re-interpretation of the model is offered in sub-section 2.5 which shows how it can be used to analyse a world market for primary products produced by different countries and purchased by different firms. Some effects of technical progress are discussed in sub-section 2.6 and in subsection 2.7 we show how parameter changes, such as in costs, affect concentration in equilibrium.

The output of firm i is x_i, and total output is $X=\sum_i x_i$. Writing p for the market price, let the inverse demand function for the market be

$$p=\phi(X). \tag{1}$$

Equilibrium depends on the conduct of the firms, or the 'rules of the game'. While no unique, universally valid, solution can be given, a wide variety of plausible outcomes can be captured in a conjectural variations model.

We take firm i's belief about the responses of other firms to its output changes as given by

$$\mathrm{d}x_k/\mathrm{d}x_i=\alpha(x_k/x_i) \quad \text{for} \quad k\neq i. \tag{2}$$

Thus a 1% increase in x_i is believed to provoke an α% increase in the output of each of the other firms. The case $\alpha=0$ is that of Cournot; the extreme $\alpha=1$ is where each firm believes that the others will always try to preserve their market shares. Cases of $\alpha<0$ are also conceivable, working towards 'accommodating' behaviour on part of other firms, an example being where they adjust their outputs to keep price constant. Our model allows then collusive behaviour which may be tacit.

The assumption that α is the same for each firm's belief about each other firm is, of course, special, and is made for algebraic simplicity, which improves the applicability of the theory. It captures the case where all firms are in the oligopoly on broadly equal terms.

The coefficient α can be interpreted as a subjective probability of retaliation by other firms. For example, if α is the probability that firms will respond by preserving their market shares and $(1-\alpha)$ is the probability of Cournot behaviour, the outcome will be the same, provided there is risk neutrality in the sense that firms maximise their expected profits. A related interpretation is that $(1-\alpha)$ is the degree of credibility of a threat of output-maintenance by other firms.

Given its beliefs about other firms' actions, firm i will calculate its perceived or expected marginal revenue,

$$\begin{aligned}\mathrm{d}(px_i)/\mathrm{d}x_i&=p+x_i\phi'(X)\left\{1+\alpha\sum_{k\neq i}(x_k/x_i)\right\}\\&=p\{1-[\alpha+(1-\alpha)s_i]/\varepsilon\},\end{aligned} \tag{3}$$

where

$$\varepsilon = -(p/X)(\mathrm{d}X/\mathrm{d}p) = \text{elasticity of market demand},$$

and

$$s_i = x_i/X = \text{market share of firm } i.$$

We assume that each firm has constant unit variable (equal to marginal) cost, c_i for firm i.

To maximise its profit, firm i will set the perceived marginal revenue equal to marginal cost. This yields the relation

$$p\{1-[\alpha+(1-\alpha)s_i]/\varepsilon\} = c_i, \tag{4}$$

provided firm i decides to supply to this market at all. The only serious problem this qualification can pose arises if $\alpha = 1$. Then (4) cannot hold for all i when the c_i are unequal. When all firms try to maintain their market shares, only the least-cost ones survive. We henceforth leave that case aside, and assume $\alpha < 1$. Then we can concentrate attention on the firms which are active; suppose there are n of them.

Adding eq. (4) across the active firms, we have

$$p\{n(1-\alpha/\varepsilon)-(1-\alpha)/\varepsilon\} = \sum c_i, \tag{5}$$

where the summation on the right-hand side is understood to be over the range of i going from 1 to n. Since ε is a function of p, eq. (5) can be solved for its only unknown p. The solution is simple and explicit if ε is constant. We therefore approximate the demand curve by a constant-elasticity curve, i.e., we take

$$X = Ap^{-\varepsilon}. \tag{6}$$

This conforms with the practice of regarding the elasticity of market demand as an exogenous parameter in much of industrial economics.

Then (5) can be solved for the price. Writing $\bar{c} = \sum c_i/n$, we have

$$p = \bar{c}/\{1-\alpha/\varepsilon-(1-\alpha)/(n\varepsilon)\}. \tag{7}$$

In the Cournot case, this becomes $p = \bar{c}/\{1-1/(n\varepsilon)\}$, i.e., the average firm's cost marked up by a factor corresponding to the perceived demand elasticity facing the average firm. If the other parameters and variables are known or observed, (7) allows us to calculate α.

Once the equilibrium price is known, it is easy to compute the other magnitudes. On dividing (4) by (5), we can solve for firm i's share

$$s_i = (\varepsilon - \alpha)/(1-\alpha) - (c_i/\bar{c})\{(\varepsilon - \alpha)/(1-\alpha) - 1/n\}. \tag{8}$$

Then, knowing the market size X, we can find firm i's sales $x_i = Xs_i$.

Eqs. (7) and (8) define the market price and the firm's market shares in terms of the basic parameters of the problem: demand elasticity, costs, the number of firms, and the conjecture parameter. This permits some comparative statics: for example, market share decreases as the marginal cost increases, with $s_i \gtrless 1/n$ according as $c_i \lessgtr \bar{c}$.

We can also relate the market shares to the price–cost margins by rewriting (4) as follows:

$$s_i = (\varepsilon/(1-\alpha))((p-c_i)/p) - \alpha/(1-\alpha). \tag{9}$$

This is often useful, but it must be remembered that there are endogenous variables on both sides of the equation.

The equilibrium having been determined, we now examine its effects, in particular its costs and benefits for various groups. To do this we need a standard of comparison, i.e., an alternative to which the equilibrium we have analysed may be compared. The most straightforward is to assume that the product is bought from the cheapest source, i.e., at cost c^* where

$$c^* = \min_i c_i, \tag{10}$$

and that price is set equal to this marginal cost c^*.

The attraction of this reference standard comes from the usual economic ideas of production efficiency and marginal-cost pricing. Thus it serves as a benchmark for the purpose of discussing the effects of concentration and changes in concentration; it captures the notion that oligopoly forces consumers to face prices which are 'too high'. However, it must be emphasised that we are not suggesting that the alternative state of affairs is achievable, or that the countries should necessarily strive to establish it. In particular, if there are any fixed costs of production, then marginal-cost pricing would have to be accompanied by lump-sum transfers to cover the firms' losses. This would be particularly difficult in the context of international trade.

The comparison of the equilibrium with the benchmark is shown in fig. 1. The consumer surplus lost by the price being p rather than c^* is the area $ABCD$. The shaded area gives the firms' profits ignoring fixed costs. The horizontal dimension of each step corresponds to the output of each firm, or a group of equal-cost firms, and the height of the step is the cost.

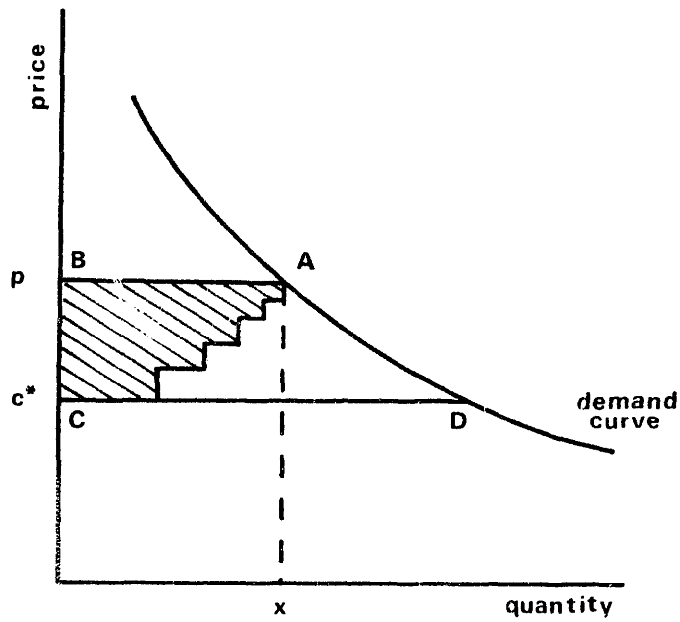

Fig. 1

The loss of consumer surplus can be evaluated as

$$\int_{c^*}^{p} Ap^{-\varepsilon}\,\mathrm{d}p=(A/(1-\varepsilon))\{p^{1-\varepsilon}-c^{*1-\varepsilon}\}$$

$$=(Ap^{1-\varepsilon}/(1-\varepsilon))\{1-(c^*/p)^{1-\varepsilon}\}$$

$$=(R/(1-\varepsilon))\{1-(c^*/p)^{1-\varepsilon}\}, \qquad (11)$$

where $R=pX$ is the value of sales in the market.

Since c^* is not readily observable, we cast this into a different form. Let s^* be the market share of the least-cost firm (or of each of a group of firms with equal least cost). This is easier to observe than c^*. The model indicates that the least-cost firm will also have the biggest market share. Thus, if it is judged that there are say k firms with nearly equal and lowest costs, then s^* can be approximated by $CR(k)/k$, where $CR(k)$ is the usual k-firm concentration ratio. Now (4) implies

$$c^*/p=1-[\alpha+(1-\alpha)s^*]\varepsilon. \qquad (12)$$

Substituting in (11), the lost consumer surplus becomes

$$\int_{c^*}^{p} Ap^{-\varepsilon}\,\mathrm{d}p=(R/(1-\varepsilon))\{1-(\alpha+(1-\alpha)s^*/\varepsilon]^{1-\varepsilon}\}. \qquad (13)$$

If $[\alpha+(1-\alpha)s^*]/\varepsilon$ is small compared to unity (for example, if behaviour is nearly non-collusive and the least-cost firm is small, or if the price-elasticity of demand is large), we can use the approximation $(1-z)^{\theta}\simeq 1-\theta z$, valid for small z. Then (13) becomes

$$\int_{c^*}^{p} Ap^{-\varepsilon}\,\mathrm{d}p \simeq (R/\varepsilon)[\alpha+(1-\alpha)s^*]. \tag{14}$$

This is the final expression for the lost consumer surplus in terms of observable magnitudes. Note that s^* is a measure of concentration: a new one but closely related to the usual concentration ratios. The added feature is that the number of firms to be chosen for inclusion in the measure is not arbitrary, but is given by the economic consideration of their being the least-cost firms.

Let us turn to the profit of the firms operating in this industry. This is evaluated as

$$\begin{aligned} \Pi &= \sum (p-c_i)x_i \\ &= R\sum ((p-c_i)/p)(x_i/X) \\ &= R\sum [(\alpha+(1-\alpha)s_i)/\varepsilon]s_i \\ &= (R/\varepsilon)[\alpha+(1-\alpha)H] \end{aligned} \tag{15}$$

where $H=\sum s_i^2$ is the Herfindahl index of concentration for the market.

We have shown how certain concentration indices can reflect important features of the solution. It is also seen that the choice of an appropriate index can vary with the question being asked. Therefore it is important to have a proper theoretical foundation which will indicate the relevant directions for empirical studies.

The simplest 'Marshallian' measure of welfare would be the sum of consumer and producer surpluses. In terms of fig. 1, this would mean subtracting the shaded area of producers' gain from the area ABCD of consumers' loss to obtain the unshaded area as the net, or dead-weight, loss from the oligopoly relative to marginal-cost pricing. Such an approach is subject to a number of objections, some of which are particularly important for our intended applications. An ever-present concern is that for distribution. If a dollar of gain to consumers is seen as more valuable than a dollar of profit, this can be handled by giving a weight less than one to producer surplus in the summation. Aggregate consumer surplus also neglects distribution of income amongst consumers. This could be dealt with, in principle, by similar weighting techniques for different consumer groups. In practice this might be rather complicated. In any case, the choice of weights is a matter of value judgement. We will therefore merely present separately the different components which might be of interest to a policy-maker. A fine enough division may not always be practicable, but separate calculations of aggregate consumer surplus and profit are easy enough. Other specific questions will be discussed as we turn to applications.

2.2. *Foreign-owned firms*

We begin with a case where some of the firms supplying to this market are owned by foreigners. This can raise several new issues. At a minimum, in its welfare evaluation the government may wish to treat the profits of domestic and foreign firms differently. It is easy to compute separate expressions for the two, and express them in terms of concentration indices, following the same steps as in the derivation of (15), but extending the range of summation only over the appropriate subset of firms. For foreign firms, writing their profits as Π_F and indicating summation as i ranges over them by $\sum_F$, we have

$$\Pi_F=(R/\varepsilon)\sum_F[\alpha+(1-\alpha)s_i]s_i.$$

Now let μ_F be the share of foreign firms in the market, and let H_F be the Herfindahl index of concentration among the foreign firms only, i.e.,

$$\mu_F=\sum_F s_i, \qquad H_F=\sum_F (s_i/\mu_F)^2.$$

Then

$$\Pi_F=(\mu_F R/\varepsilon)[\alpha+(1-\alpha)\mu_F H_F]. \tag{16}$$

In exactly analogous notation for the domestic firms, we have

$$\Pi_D=(\mu_D R/\varepsilon)[\alpha+(1-\alpha)\mu_D H_D]. \tag{17}$$

The profit of each group of firms is thus related to the Herfindahl index of concentration *within that group*. This again highlights the importance of indices of concentration, and of choosing the appropriate one.

If the policy-maker treated domestic consumer surplus and profit equally, but neglected foreign profits altogether, for example, the net welfare loss from oligopoly compared to marginal-cost pricing would be given by (14) minus (17).

2.3. *Location of production*

When some or all of the firms in this market are transnationals, they may locate production in one country to supply to another. Note that this question is distinct from that of ownership. If the market prices of factors equal their social opportunity costs, there is no concern for the level of production as such. However, when the two differ, policy-makers in the home country may have welfare objectives like employment that are not captured in the product–market deadweight loss expressions. Here we see how such concerns relate to the oligopolistic market structure.

The constant marginal cost c_i for firm i can be interpreted as the delivered cost to the market. If the firm's sources of supply are indexed by j, write c_{ij} for the constant marginal cost of supplying the market from source j. Much of the traditional theory of transnationals can be seen as explaining the reasons for differences among the c_{ij}; a recent survey is in Hood and Young (1979, ch. 2). We build on this basis to find the optimum production plan. Firm i will choose quantities x_{ij} of supplies from the various sources to

$$\text{minimise} \quad \sum_j c_{ij}x_{ij} \quad \text{subject to} \quad \sum_j x_{ij}=x_i.$$

This is obviously achieved by choosing only one non-zero x_{ij}, namely that for the source with the least cost among the c_{ij}. This smallest cost is then c_i. In other words, the decision about the source of supply is a purely technological one, independent of the oligopolistic structure. Therefore the country's cost–benefit analysis of its industrial production policy can be conducted independently of its domestic regulatory policy.

This conclusion will have to be modified if there are significant fixed costs so that decisions of whether to operate some supply sources at all are important. Marginal costs varying significantly with output will also entail modifications as will any government-imposed link between where production takes place and the right to supply to a country. However, some support for the above proposition can be found in the increasing world-wide integration of production in industries like automobiles.

2.4. Shadow prices

Like most industry analyses, ours is of a partial equilibrium nature. It is important to understand the practical significance of this in our context. Once that is done, we can devise procedures to correct for the significant effects that the earlier simple framework neglects.

Partial equilibrium analysis focuses on one industry, effectively summarising the rest of the economy in the form of the marginal costs of production in this industry. Some feedback effects, such as those on the demand curve in this industry arising from the profit generated in it, can be safely neglected so long as the industry is only a small part of the economy. However, the implicit assumption that the marginal cost of output reflects the true social opportunity cost can be questioned on several grounds. The commonest grounds are those of monopoly, external effects, taxes, and administered prices and wages associated with non-clearing markets. If there are significant distortions of this kind in the rest of the economy, then the financial marginal cost facing the industry under study will differ from the true social opportunity cost, and the difference cannot be calculated without paying attention to the rest of the economy. This 'general problem of the second-best' cannot be completely solved without constructing the model of the economy as a whole. However, there are particular aspects of practical significance that can be taken into account by some approximate devices. The two that we discuss here concern employment and the balance of payments.

Consider employment first. In a fully employed economy, an extra person in the industry in question can only be hired from the rest of the economy. The social opportunity cost of this employment in a competitive labour market is just the market wage rate. In an economy with unemployment or labour-market distortions, the two may differ. In particular, in less developed countries there are well-known systematic forces causing such differences; an exposition of the issues and methods for calculating shadow wage rates can be found in Little and Mirrlees (1974) and other manuals for cost-benefit analysis. Here we show how we can use the shadow wage to obtain the welfare effects of employment that are on a comparable basis with the product—market deadweight loss measures. We consider a closed economy for algebraic simplicity; the additional features brought in by trade can be incorporated as in sub-sections 2.2 and 2.3.

Suppose the shadow wage is a fraction ω of the market wage. Then a fraction $(1-\omega)$ of the market cost of labour to the industry is not true social cost. We should therefore exclude it from the costs, or alternatively add it to the consumer and producer surpluses, to obtain a better measure of welfare. Let labour costs be a fraction θ of marginal costs. Then the added benefit from employment is given by

$$(1-\omega)\theta\sum_i c_i x_i = (1-\omega)\theta R\sum (c_i/p)(x_i/X)$$

$$= (1-\omega)\theta R\{(1-\alpha/\varepsilon)-(1-\alpha)H/\varepsilon\}, \qquad (18)$$

using (4) for simplification.

Note that in principle the existence of oligopoly in our industry will affect the value of the shadow wage rate, a fact which will have been typically ignored by those computing that magnitude. However, this is a general equilibrium consideration that may be safely ignored so long as the industry is a small part of the economy.

The treatment of the balance of payments is similar. If the home currency is overvalued, the social opportunity cost of imports exceeds the official financial cost. We can then say that the shadow price of foreign exchange is higher than its market price. Suppose it is $(1+\lambda)$ times the latter. Then we have to recognise that a further multiple λ of any foreign exchange costs must be subtracted from the consumer and producer surplus calculations. The foreign exchange cost is all of the cost for firms who import the product, and an appropriate fraction of the cost for those who produce locally. The expression for the additional foreign exchange costs is then straightforward to calculate, and involves Herfindahl indices for the appropriate segments of the market. We omit the details to save space. Again, the problem is particularly significant in less developed countries, and Little and Mirrlees etc. indicate methods of calculating the shadow cost of foreign exchange.

2.5. *World markets for primary products*

The organisation of markets for primary products which are extracted or produced in a number of countries is a matter of great and increasing importance. Many of these can be modelled as an oligopoly where country supplies are channelled to the world market by private firms. Governments attempt to extract some of the profits through royalties or taxes. For ease of exposition, suppose that each country is linked to a distinct firm, with the index i denoting a country and its firm. Let the extraction costs per unit of output be b_i, and the tax rates $(c_i - b_i)$, so the costs to the firms are c_i. There is a world (derived) demand curve for the product, given by eq. (1).

The formal model of the oligopolistic equilibrium is exactly that presented in the main part of this section, and the solution can be read off from eqs. (4)–(9). The new feature comes from the tax revenues of the countries,

$$T_i = (c_i - b_i)x_i. \tag{19}$$

To understand the countries' incentives to levy such taxes, we differentiate this with respect to the tax rates, or equivalently, the c_i. We have

$$\partial T_i/\partial c_i = x_i + (c_i - b_i)(\partial x_i/\partial c_i). \tag{20}$$

The two terms on the right-hand side reflect the usual conflict between increased revenue on the existing tax base, and the erosion of this base, as tax rates are raised. For a more detailed analysis of the second effect, we note that $x_i = Xs_i = Ap^{-\varepsilon}s_i$ and

$$(1/X)(\partial x_i/\partial c_i) = -(\varepsilon/p)(\partial p/\partial c_i)s_i + \partial s_i/\partial c_i. \tag{21}$$

Now differentiation of (7) yields

$$\partial p/\partial c_i = p\Big/\sum_k c_k, \tag{22a}$$

while from (9) after some simplification,

$$\partial s_i/\partial c_i = -(\varepsilon/(1-\alpha))(1/p)\left(\sum_{k\neq i} c_k\Big/\sum_k c_k\right). \tag{22b}$$

These results give rise to an interesting conflict. At an initial point with no taxes, the second term in (20) is zero. The revenue gain from a small per unit tax is therefore proportional to x_i. This makes the incentive to initiate such a system of taxes greatest for the largest (i.e., least-cost) producers. For further marginal increments of the taxes, however, the base erosion effects arise. From (21) and (22a) and (22b) we see that both components of this effect are larger for the lower-cost producers. Thus the base erosion effect operates in the direction of giving the larger, i.e., lower-cost, producers an incentive to be more moderate in intercountry discussions on raising the tax rates. Both these observations find support in the markets for crude oil, metal ores, etc., although in each case one can doubtless find other arguments that support the same facts.

2.6. *Technical progress*

A proper theory of technical progress in oligopoly must be dynamic, considering how market conditions affect and in turn are affected by the research and development activities of the firms. Some recent attempts to build such a theory have appeared in the literature, for example, Dasgupta and Stiglitz (1980). Further work along those lines would be outside our present scope. However, some simple inferences from our earlier static model enable us to shed some light on the question of the rate of technical progress and its relation to concentration, and to do so in a way that is empirically testable in the familiar structure–performance framework.

Central to nearly all analyses of the rates of growth of products and of technical progress is the question of profitability. Consider a firm in an oligopolistic industry of the kind that was examined in the first part of this section. This firm's profits are

$$\Pi_i = (p - c_i)x_i. \tag{23}$$

The firm can undertake activities to reduce its unit cost; these might be research, exploration, or merely organisational improvement. Its incentive to do so comes from its perception of increased profits as the result. The marginal profit yield of a reduction in c_i comes from two sources. First is the direct effect, namely the change in the right-hand side of (23) with p and x_i held fixed. Secondly, we have the effect of an induced change in x_i, including the expected response of other firms, and leading to a change in the market price. It is natural, and logically consistent, to assume that the firm uses the same beliefs about the responses of other firms in making this calculation as it does in calculating its profit-maximising output choice. Then, the firm being at such a perceived profit-maximising point in equilibrium, its perception of a marginal effect on profit from an output change is of the second order of smallness. This is just an unusual use of the Wong–Viner–Samuelson envelope theorem. Thus the perceived response of profits to unit cost changes is given by

$$\partial \Pi_i/\partial c_i = -x_i. \tag{24}$$

The firm's cost-reducing activities will respond to this perceived return. Consider a simple case where the proportional rate of cost-reduction is linearly-related to the return, i.e.,

$$(1/c_i)(\mathrm{d}c_i/\mathrm{d}t) = -kx_i, \tag{25}$$

Multiplying by c_i and using (4), we have

$$-\mathrm{d}c_i/\mathrm{d}t = ks_iXp\{(1-\alpha/\varepsilon)-((1-\alpha)/\varepsilon)s_i\}.$$

Adding over i, we obtain the rate of reduction of the average of the firms' costs

$$-n(\mathrm{d}\bar{c}/\mathrm{d}t) = kR\{(1-\alpha/\varepsilon)-((1-\alpha)/\varepsilon)H\},$$

or

$$-\mathrm{d}\bar{c}/\mathrm{d}t = (kR/n)\{(1-\alpha/\varepsilon)-((1-\alpha)/\varepsilon)H\}. \tag{26}$$

This shows that, other things equal, reduction of unit costs will occur on the average more rapidly in large markets (large R), but more slowly in more concentrated markets (large H).

2.7. *Changes in concentration*

Given the importance of the Herfindahl index in describing various outcomes in this model, we ask how that index will change as costs of the firms change. We have

$$\partial H/\partial c_i = 2\sum_j s_j(\partial s_j/\partial c_i). \tag{27}$$

We already have an expression in (22b) for $\partial s_i/\partial c_i$. When $j \neq i$, we find similarly

$$\partial s_j/\partial c_i = (\varepsilon/(1-\alpha))(1/p)\left(c_j\Big/\sum_k c_k\right). \tag{28}$$

Substituting into (27) and simplifying,

$$\partial H/\partial c_i = (2\varepsilon/(1-\alpha))(1/p)\left(1\Big/\sum_k c_k\right)\left\{\sum_k c_k s_k - s_i \sum_k c_k\right\}. \tag{29}$$

Thus a rise in the cost for firm i will raise the Herfindahl index if its own share is less than the weighted average of all the firms' shares, the relative weights being simply the costs. The general intuition is clear: since the costs of firms and their market share are negatively related, a rise in the cost of a low-share firm, by further lowering its share, will raise the concentration index. What (29) does is to give a precise formulation of the relevant concept of a low share.

3. Price leadership

In this section we consider an alternative market structure, where a group of collusive firms coexists with a competitive fringe. This model may be particularly appropriate in the context of some newly industrialising less developed countries, where modern industry is dominated by large foreign firms, and a fringe of competitive domestic firms supplies a closely related product. With this application in mind, we construct a model of an industry where the goods are imperfect substitutes, although the special case of a homogeneous product is also examined. Other potential applications include the world petroleum market, where the major companies (and countries) attempt to exercise leadership and the independents follow. Related models include Encoua and Jacquemin (1980), where the dominant firms can be less than perfectly collusive, and Dickson (1978), where the domestic firms act as collusive price leaders with a fringe of competitive foreign firms.

With our main application in mind, we focus on the magnitudes of the consumer surplus and the profits of the collusive firms, and also find the effects on employment. The procedure is standard. We stipulate a system of demand functions for the set of products. The conditions of equilibrium of the competitive fringe can then be used to obtain the residual demands for the collusive firms, and their joint profit-maximising choices calculated. In our partial equilibrium framework, it is harmless to aggregate the fringe goods into one. We also assume this good to be produced under conditions of constant average cost c_0; then its competitive price p_0 equals c_0. As to the demand system, a constant-elasticity assumption is no longer appropriate. The products are sufficiently close substitutes that demands can be reduced to zero for sufficiently high prices. Therefore we use the approximation of linear demands.

There are n goods produced by the collusive group, with a vector c of their unit costs, and a vector p of prices. The demand for the fringe good is $d_0(p_0, p)$, i.e., it depends on its own price as well as the prices of the substitute goods. The demands for those goods depend on the same $(n+1)$ prices, and are written as a vector function $d(p_0, p)$. We use the following linear approximations:

$$d_0(p_0, p) = a_0 - b_{00}p_0 + bp, \tag{30}$$

$$d(p_0, p) = a + bp_0 - Bp, \tag{31}$$

where B is a symmetric positive definite n-by-n matrix reflecting substitution amongst the collusive goods, and b gives the substitution effects between the fringe good and collusive goods.

The profits of the collusive group are given by

$$\Pi(p_0, p) = (p - c)(a + bp_0 - Bp). \tag{32}$$

We suppose that p is chosen collusively to maximise (32). The first-order condition for this is

$$(a + bp_0 - Bp) - B(p - c) = 0, \tag{33}$$

which gives a profit-maximising price

$$p = \tfrac{1}{2}(p^* + c), \tag{34}$$

where

$$p^* = B^{-1}(a + bp_0). \tag{35}$$

This has a simple interpretation. From the demand function (31) and the definition (35), we see that $d(p_0, p^*) = 0$, i.e., p^* is the price vector of collusive goods which, given p_0, implies zero demand for these goods. Thus (34) tells us that the collusive profit-maximising price vector is the average of the unit cost and the price which completely chokes off demand.

The profit-maximising output is obtained by substituting from (34) into (31) and using $p_0 = c_0$. We have

$$x = B(p^* - p) = B(p - c). \tag{36}$$

Then the maximised profits of the group are given by

$$\Pi = (p-c)'x = x'B^{-1}x. \tag{37}$$

Since the matrix B is positive definite, so is its inverse. Thus (37) is a positive definite quadratic form in the outputs x of the collusive group. Write X for the total $\sum_{\mathrm{T}} x_i$, and s for the vector of shares $s_i = x_i/X$. Note that with heterogeneous products the sum $\sum_{\mathrm{T}} x_i$ must be taken after the x_i have been converted into common units; this could be achieved by defining the units of each good so that a unit of good i has the same value as a unit of good j. Then (37) can be written as

$$\Pi = X^2(s'B^{-1}s). \tag{38}$$

A positive definite quadratic form in shares is a convex function of the shares, and therefore has many of the desirable attributes of a concentration index. We consider a particularly simple and attractive case. Suppose the substitution matrix satisfies

$$B^{-1} = \begin{pmatrix} 1 & \beta & \beta & \dots & \\ \beta & 1 & \beta & \dots & \\ \beta & \beta & 1 & \dots & \\ \dots & \dots & \dots & \dots & \\ \beta & \beta & \dots & \dots & 1 \end{pmatrix}, \tag{39}$$

where β lies between 0 and 1. Intuitively, this entails symmetry within the group of foreign goods in the sense that substitution behaviour between any pair is the same as that between any other pair. Then (38) becomes

$$\Pi = X^2[\beta + (1-\beta)H], \tag{40}$$

where $H = \sum_{\mathrm{T}} s_i^2$ is the Herfindahl index for the collusive firms.

The benefit to the country from the existence of alternatives to the fringe good is given by the difference between the consumer surplus at prices (p_0, p) and that at prices (p_0, p^*); recall that p^* is the price where there are zero purchases of the foreign good, and that fringe firms producing at constant costs and selling competitively always have $p_0 = c_0$ and zero profits. The consumer surplus function consistent with the demand functions (32) and (33) can be shown to be

$$\sum(p_0,p) = -\begin{pmatrix} a_0 \\ a \end{pmatrix}\cdot\begin{pmatrix} p_0 \\ p \end{pmatrix} + \frac{1}{2}\begin{pmatrix} p_0 \\ p \end{pmatrix}'\begin{pmatrix} b_{00} & -b' \\ -b & B \end{pmatrix}\begin{pmatrix} p_0 \\ p \end{pmatrix}. \tag{41}$$

Therefore the benefit as measured by the consumer surplus is

$$\sum(p_0,p) - \sum(p_0,p^*).$$

Substituting for p from (34) in (41), after routine manipulations, we find that the consumer surplus gain is exactly $\frac{1}{2}\Pi$, where Π is the maximised collusive profit as given by (40).

The potential gains from trade if the whole market were competitive would be given by the surplus associated with the price vector (c_0, c) as compared with that at the prices (c_0, p^*), i.e.,

$$\sum(c_0,c) - \sum(c_0,p^*).$$

It is again straightforward to show that this expression equals 2Π.

These results are in some ways extensions of the findings of Cowling and Mueller (1978) and others on the welfare costs of monopoly. They can be illustrated very simply using a diagram in the homogeneous product case. In fig. 2, we show a linear residual demand curve for the price leader, and the

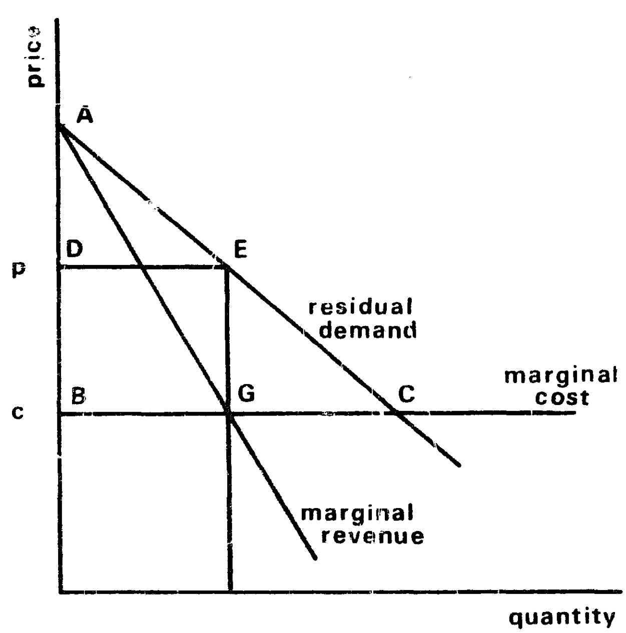

Fig. 2

corresponding marginal revenue curve. The quantity produced by the leader, BG, is equal to half of BC, the quantity that would arise if price were set equal to marginal cost. The gain of consumer surplus from having a profit-maximising price leader is the triangle ADE, which is half of this firm's profit $DEGB$, which is in turn half of the consumer surplus gain that would arise with a competitive firm instead, namely the triangle ACB.

The implications of this model for a less developed country industrialising with the involvement of transnationals can now be drawn. The presence of the transnationals permits a benefit equal to half of their profits in comparison with autarky. But we have also seen that the collusive behaviour of the oligopolists means that the gains from trade are only one quarter of their potential under an ideal competitive state, these potential gains being twice the profits. Finally we established that for certain forms of the demand functions these profits can be expressed in terms of the Herfindahl index for the transnationals. Here once again, just as in the models with homogeneous products of the previous section, a Herfindahl index is useful in summarising market conditions, reflecting both profits of transnationals and net benefits to the domestic economy. However, we should stress that the group of firms over which such an index is to be calculated differs depending on the component of net benefits being measured.

With heterogeneous products it is natural to think of advertising and persuasion by firms to turn consumer demand in their favour. The potential for this can be seen clearly in our model. We can write the transnationals' maximised profits (37) as

$$\Pi = (p-c)'B'(p-c). \tag{42}$$

This is a positive definite quadratic form in the profit margins $(p-c)$. The latter are given by the expression

$$\tfrac{1}{2}[B^{-1}(a+bp_0)-c], \tag{43}$$

from (34) and (35). The transnationals will therefore have an incentive to make a and b as large as possible, given that all products are mutual substitutes so that the matrix B^{-1} has all elements positive. Increasing a means that the price-insensitive part of the transnationals' demand is strong, while increasing b means that small price reductions for their own goods switch demand sharply away from the domestic good whose price is fixed by the cost of production.

It is again straightforward to discuss the employment implications in this model by linking them to output. Consider an extreme case where all of the transnationals' supply is imported from abroad. The loss in domestic sales from having the imports of the transnationals is

$$d_0(c_0, p^*) - d_0(c_0, p) = b'(p-c) = b'B^{-1}x. \tag{44}$$

Thus employment loss depends linearly on the profit margins, or the outputs, of the transnationals. If we assume additional symmetry in making all components of b equal, i.e., the domestic good substitutable to the same extent with any of the foreign goods, then the loss of domestic firms' sales becomes simply proportional to X, the aggregate output of the transnationals.

4. Concluding remarks

Our intention in this paper has been to provide a unified treatment of models of oligopoly and welfare which link peformance or outcomes to market structure. We attempted to express results in terms of the underlying parameters of the model, such as those for cost and demand, instead of proxies like price–cost margins, and to identify those measures of structure, e.g. concentration indices, and performance which emerge naturally in the expressions for important aspects of equilibrium such as profits, welfare losses and so on. One of our particular aims in the analysis has been to produce results which are in terms of observable variables so that the theory can readily be applied. Secondly we wished to show how the theory could be interpreted and extended to allow the analysis of oligopoly in international trade, for example the operation of transnationals in less developed countries.

In our discussion of a market for a homogeneous product with conjectural variations by participating firms we have seen how the Herfindahl index plays an important role in expressions for profit and welfare loss. We showed how imperfections in markets for factors used in the industry could be introduced into the welfare analysis. We offered a re-interpretation of the model as one of several suppliers of a primary commodity to a world market. A simple model of the relation between technical progress and concentration was developed and we examined the relation between parameter changes and concentration indices.

In section 3 on price leadership we introduced an analysis of collusive firms producing different products combined with a competitive fringe, motivated in part by a wish to provide a theoretical framework, and especially indices, for the analysis of the operation of transnational companies in less-developed countries. We showed how even in this case of heterogeneous products simple expressions for profits and welfare loss could be derived and again the Herfindahl index played an important role.

We have found that received theories of oligopoly and welfare, if suitably developed and extended, can provide a useful explanation of several features of trade in the presence of oligopoly. Further we found that many important components of a welfare analysis could be related to simple observable magnitudes and indices so that the theory should be capable of useful application to policy problems in the analysis of oligopoly with or without trade and in both developed and less-developed countries.

References

Cowling, K.G., 1976, On the theoretical specification of industrial structure-performance relationships, European Economic Review 8, 1–14.

Cowling, K.G. and D. Mueller, 1978, The social costs of monopoly power, Economic Journal 88, 724–748.

Cowling, K.G. and M. Waterson, 1976, Price–cost margins and market structure, Economica 43, 267–274.

Dansby, R.E. and R.D. Willig, 1979, Industry performance gradient indexes, American Economic Review 69, 249–260.

Dasgupta, P.S. and J.E. Stiglitz, 1980, Uncertainty, industrial structure and the speed of R&D, Bell Journal of Economics 11, 1–28.

Dickson, V., 1978, Foreign competition and the degree of monopoly — Comment and extension, Journal of Economic Studies 5, 71–75.

Encaoua, D. and A. Jacquemin, 1980, Degree of monopoly, indices of concentration and threat of entry, International Economic Review 21, 87–105.

Hitiris, T.L., 1978, Effective protection and economic performance in UK manufacturing industry, Economic Journal 88, 107–120.

Hood, N. and S. Young, 1979, The economics of multinational enterprise (Longman, London and New York).

Jacquemin, A.P. and H.W. de Jong, 1977, European industrial organization (Macmillan, London).

Katrak, H., 1980, Multinational monopolies and monopoly power regulation, Oxford Economic Papers 32, 453–456.

Little, L.M.D. and J.A. Mirrlees, 1974, Project appraisal and planning for developing countries (Heinemann, London).

Pagoulatas, E. and R.L. Sorenson, 1976, Foreign trade, concentration and profitability in open economics. European Economic Review 8, 255–267.

Scherer, F.M., 1980, Industrial market structure and economic performance, 2nd ed. (Rand-McNally, Chicago, IL).

Stigler, G.J., 1964, A theory of oligopoly, Journal of Political Economy 72, 44–61.

Weiss, L., 1971, Quantitative studies of industrial organisation in: M.D. Intriligator, ed., Frontiers of quantitative economics, Vol. 1 (North-Holland, Amsterdam).

White, L.J., 1974, Industrial organization and international trade: Some theoretical considerations, American Economic Review 64, 1013–1020.