Some Comments on the Theory of Monopoly [1]

Lucien Foldes

The familiar pair of equations (of perfect competition) is deduced by the present writer from the first principle: equilibrium is attained when the existing contracts can neither be varied without recontract with the consent of the existing parties, nor by recontract within the field of competition. The advantage of this general method is that it is applicable to the particular cases of imperfect competition, where the conceptions of *demand and supply at a price* are no longer appropriate.

.... the influence of monopoly is well traced by Cournot in a discussion masterly, but limited by a particular condition, which may be called *uniformity of price not (it is submitted) abstractedly necessary in cases of imperfect competition.* Going beyond Cournot, not without trembling, the present inquiry finds that, where the field of competition is sensibly imperfect, an indefinite number of final settlements are possible .... and in the absence of imposed conditions, the said final settlements are not *in the demand-curve,* but on the contract curve. (Edgeworth, F.Y., *Mathematical Psychics,* pp. 31 and 47-48.)

1

According to the main proposition of the theory of monopoly, a seller who cannot discriminate fixes a price which equates marginal revenue with marginal cost, and meets the demand at that price. Since the price exceeds marginal cost, there is (subject to certain conditions) a loss of welfare, which in the absence of entry can be corrected only by public intervention. The theory of discriminating monopoly, apart from the limiting case of perfect discrimination, is similar though more complicated: the equation of marginal revenue and cost must be applied to determine simultaneously the prices in each of several 'horizontal' or 'vertical' sub-markets.[2]

The present paper reconsiders this theory, paying attention in particular to two questions: (i) whether and in what circumstances the Cournot solution for simple monopoly is an equilibrium and whether it is the only equilibrium; (ii) whether even in the absence of entry there

may not be market forces which 'automatically' eliminate or limit losses of welfare due to simple monopoly.

The traditional analysis does not deal adequately with these questions, largely because the treatment of monopoly has been dominated by the concepts and methods of competitive theory. This problem of method, which affects the theory of value as a whole, is considered briefly in the next Section. The discussion then turns to a critical review of basic definitions and assumptions in the theory of simple monopoly, pointing out various inadequacies in the usual treatment (Section III). Various mechanisms are then examined by which market forces may alter the situation of simple monopoly, and so perhaps realise additional gains from trade (Section IV). The essay concludes with some general remarks on the analysis of monopoly and the approach to problems of policy.

II

As the quotations at the head of this paper show Edgeworth pointed long ago to some of the main shortcomings of what may be called the classical procedure in the theory of value. The main feature of this procedure is its primary emphasis on the resolution of conflicting interests within a given framework of perfect competition, and on the correspondence between competitive equilibrium and optimality.

Other market forms are classified according to their relationship with the competitive model, and the methods and formulations of

2.   By horizontal discrimination is meant the usual system, called by Pigou discrimination of third degree, in which different prices are charged to each of several groups of customers. Vertical or second degree discrimination operates 'in such wise that all units with a demand price greater than x were sold at a price x, all with a demand price less than x and greater than y at a price y, and so on' (*Economics of Welfare*, p.279). This system has been studied mainly by German writers, whose contributions appear largely to have escaped notice in the English literature. The main paper is Stackelberg H.V., 'Preisdiskrimination bei willkürlicher Teilung des Marktes' *Archiv für mathematische Wirtschafts- und Sozialforschung*, Vol. 5, 1939. More recent contributions, which discuss the convergence of second degree to first degree or perfect discrimination as the number of vertical 'slices' tends to infinity, include Gutenberg E., *Grundlagen der Betriebswirtschaftslehre*, Vol. 2, p.280, 1955, and Ott A.E., 'Zum Problem der Preisdifferenzierung', *Schweizerische Zeitschrift für Volkswirtschaft und Statistik*, 1959. These models suffer from the fundamental limitation that they apply strictly only where each buyer takes no more than one unit.

competitive theory applied to them as far as possible. Whatever the objections to this approach, it must be admitted at the outset that there is for many purposes no practicable alternative, particularly in problems of general equilibrium. The fascinating possibility, suggested by Cournot and Edgeworth, of a general theory of markets in which perfect competition and monopoly appear as limiting cases is still far from realisation, despite recent work on games and the core of an economy.[3] Nevertheless Edgeworth's approach to contract and bargaining provides an important complement to classical methods, notably at the level of partial equilibrium. In an extended form it has recently been applied to the theory of social costs[4] and even to the economic analysis of political action.[5] But apart from bilateral monopoly and oligopoly, it has had relatively little influence on the theory of imperfect competition, and on occasion it has even been taken for granted that this approach is not applicable to ordinary monopoly.[6] This limitation is unjustified, as will be argued below.

Edgeworth's approach, in its modern version, can be summed up in three main rules of method, which are briefly discussed in the remainder of this Section; they are not distinct principles, but are rather to be regarded as increasingly general applications of the same point of view.

First, the central problem in the analysis of a given market form is not to be formulated initially as the determination of equilibrium price, but as the determination of the equilibrium patterns of contract. Apart from perfect competition, the assumption that a contract establishes a single price applying to alternative quantities generally needs special justification. This principle has gradually found acceptance in the theory of bilateral monopoly, though some writers assumed until quite recently that one party to the bargain would name a price and leave the other to select a quantity, even though this would normally lead to an

3.   See in particular Scarf H.E., 'The Core of an N Person Game', *Econometrica* Jan. 1967.
4.   See, for example, Coase R.H., 'The Problem of Social Cost', *Journal of Land Economics* 1960, and Buchanan J.M., and Stubblebine W.C., 'Externality', *Economica* 1962. A discussion of some limitations of this analysis is given by S. Wellisz, 'On External Economies and the Government-Assisted Invisible Hand', *Economica* 1964.
5.   Buchanan J.M., and Tullock G., *The Calculus of Consent.*
6.   See, for example, Wellisz, *op.cit.*, 'On External Economies . . . . ,p. 349.

inefficient contract.[7] In unilateral monopoly, the seller's problem is still generally formulated *a priori* as one of determining the price to be charged — or, if discrimination is possible, the prices in various submarkets, or the prices at which successive units are to be sold. A more appropriate general formulation would be to consider the tariff, or schedule of alternative bargains, to be offered to each category of buyers. As will be shown below, the assumption that a 'straight' price is charged is not entirely satisfactory even in the traditional model of simple monopoly.[8]

Secondly, market forms, and perhaps some other institutions, should where possible not be postulated definitely in advance, but are to be regarded as responsive to the interests of individuals; the application of this principle is evidently a matter of degree, since some institutions must be assumed if others are to be explained. Now, it has always been realised that market forms are responsive to economic pressures. In particular, many authors have emphasised the tendency for traders to create monopolies, and some — notably Schumpeter and J.K. Galbraith — have stressed the incentives to counteract certain monopolistic influences. But this line of thought has not been thoroughly integrated with the formal micro-economic theory of markets. The analysis of imperfect competition and discriminating monopoly has tended, under the influence of the classical procedure, to be primarily a taxonomy of deviant species, in which emphasis is laid on the functioning and incidence of each variant rather than on the conditions for its emergence and survival.[9] The present paper does not deal with the emergence of monopoly, but the question whether simple monopoly creates incentives for its own destruction is considered in Section V below.

The third and most general principle is that the tendency of a system of free enterprise to realise gains from trade and thus perhaps to attain optimality, should be analysed in as general a framework as possible,

7.   For surveys, see Machlup F., and Taber M., 'Bilateral Monopoly, Successive Monopoly, and Vertical Integration', *Economica* 1960, and Schneider E., *Einführung in die Wirtschaftstheorie* Part II, 5th ed., 1958.

8.   A general mathematical treatment of certain monopoly tariffs will appear in a further publication.

.9.   An extreme example of this tendency is Triffin R., *Monopolistic Competition and General Equilibrium Theory*.

not merely in the case of perfect competition. Indeed, perfect competition itself can be expected to survive only when those gains from trade which can be realised at all can be realised by a market system in which communication among traders relates only to prices.

This statement of the principle is not intended to be precise, and a detailed discussion of its content and scope would be out of place here: but a few remarks are necessary if misunderstanding is to be avoided. Emphasis on the possibility of optimal allocation even without competition is associated in recent writing with the political standpoint of extreme liberalism. This connection is not inevitable, and the principle as stated here represents only a methodological commitment; no factual or moral opinion is implied, apart from the general judgement of relevance without which no method of analysis can be adopted. In particular, it is not argued here that a system of private enterprise tends invariably to eliminate misallocations due to monopoly or other causes, or that its ability to deal with them may not be improved by intervention, or that some activities may not be more efficiently conducted in the public sector.

The confusion between questions of method and of substance is connected with careless interpretation in informal discussion of the usual concepts of equilibrium and optimality; a short review of alternative definitions will clarify this point. An equilibrium is defined as a state such that no individual, or group of individuals acting in agreement, could be made better off by changing his or their own action; an optimum is a state such that no individual could be made better off without another being made worse off. The relation between these states — in particular the question whether an equilibrium of a system of private enterprise is necessarily an optimum — depends entirely on the considerations taken into account in each case in defining what 'could' be done, and on the set of actors considered. In particular, three possibilities may be noted. (i) If exactly the same actors and the same actions are considered in the definitions of both equilibrium and optimality, the proposition that an equilibrium is an optimum is logically true whatever the market form; though the converse may still be false, for example, where an optimal situation leaves scope for a person to extort payments from others by threatening actions which would not directly benefit himself but would damage them. (ii) If optimality is defined as 'Pareto' or 'first best'

optimality (in the usual sense which takes into account only the constraints represented by the production and consumption sets), while the definition of equilibrium takes into account all practical constraints and all actions available in practice to all parties except government, then the proposition that an equilibrium is necessarily optimal can be false empirically. Gains from trade measured relative to Pareto optimality are necessarily realised in full only if there are no costs of information, bargaining or administration, and if negotiations never end in deadlock until all possibilities of mutual benefit are exhausted. (iii) Finally, if both definitions are framed with reference to the same practical possibilities, except that certain actions of government are taken into account in defining optimality but not in defining equilibrium, it is again clear that an equilibrium need not be optimal. Government may be able to alter institutions so as to reduce the costs of information etc. or diminish the chances of deadlock in negotiations, or may provide generally beneficial public goods, or even improve the efficiency of a monopoly by nationalisation – quite apart from intervention on distributive grounds.

It is important to distinguish carefully among these definitions in informal discussion, if false impressions are to be avoided. Consider, for example, the following passage by Buchanan and Tullock:[10]

'We shall argue that, if the costs of organising decisions voluntarily should be zero, *all* externalities would be eliminated by voluntary private behaviour of individuals regardless of the initial structure of property rights.'

When read carefully in its context, this appears to be merely a version of definition (i), and as such is not a proposition from which practical conclusions can be drawn; but as it stands it could be misleading.

The preceding remarks are concerned primarily with definitions; but they suggest that the method under discussion can take into account a wide range of considerations relevant to issues of policy, and do not lead inevitably to tendentious formulations. Indeed, a systematic development of this method would do much to make theory more relevant to the problems of monopoly policy. Action by government tends at present to be narrowly pragmatic, fostering one monopoly and condemning another in a way which owes little to theory. One reason

10. *The Calculus of Consent*, p. 47.

for this attitude is a suspicion of facile arguments for regulation based on comparisons of the Cournot solution or one of its generalisations with marginal cost pricing. More useful guidance can be given even by static theory if account is taken of the methods by which private enterprise can deal with inefficiencies in the absence of competition, of the help which public intervention can give in this process, and of the costs of communication, administration and bargaining which these processes involve.

III

Unilateral monopoly is defined by two conditions:

a) a single seller confronts numerous buyers acting independently and each taking a negligible proportion of output;[11]

b) buyers have no bargaining power.

The second condition is not easy to define precisely in general terms, but for present purposes it is enough to specify that a buyer may only accept one or other of the alternative options offered by the seller or refuse to deal, knowing that a refusal will not elicit further offers, and that the seller knows that this is the situation and knows that the buyer knows it.

Simple monopoly is unilateral monopoly in which two further conditions hold:

c) the seller quotes a 'straight' price common to all buyers, leaving them to choose the quantities which they will take at that price;

d) resale of goods among buyers is unrestrained and costless, and takes place in a perfectly competitive market.

For most purposes, it may be assumed as an alternative to (d) that direct buyers are intermediaries who resell in a common, perfectly competitive market; but for simplicity the case where goods are sold direct to final users is usually considered. Interdependence with markets in other goods is also ignored.

11. To be rigorous, it should be assumed that the buyers form a set indexed by a continuum - say, the interval $[O,n]$. For a given set R of tariffs proposed to buyers by the monopolist, let $F(i,R)$ be the total purchases by buyers with indices $I \leqslant i$; the assumption that each buyer makes negligible purchases then means that F is absolutely continuous in i for given R.

Even these elementary conditions call for some critical comment. It is usually taken that (b) is a consequence of (a), but in general this appears to be incorrect. First, it may be that each buyer spends only a negligible part of his income on the monopolised good, and therefore derives only a negligible part of his welfare from it.[12] Even assuming that this is not the case, (a) implies only that the monopolist derives a negligible *proportion* of his profit from each buyer, but in general this sum is of the same *absolute* order of magnitude as the consumer's surplus, and there is no reason in general to assume that bargaining power is related to the proportional gains from trade. In particular, even if assumption (d) applies and a straight price is normally quoted, one buyer among many might still try to bargain with the monopolist for discounts on intramarginal units of purchase, and might even obtain a special concession if he could convince the monopolist that he would not otherwise buy.

It is also generally assumed that condition (c) follows from (a), (b) and (d), at least if the monopolist has perfect knowledge of all costs and demands. In other words, a unilateral monopolist who is unable to restrict resale etc. cannot do better than charge a straight price — which is then fixed to equate marginal revenue and cost, so that the Cournot solution is the only possible equilibrium. This also is subject to certain exceptions, when account is taken of the possibility that the monopolist may deliberately create uncertainty about prices and quantities (in a world which is otherwise free from uncertainty and strategic considerations). Specifically, suppose that the monopolist maximises expected profit, while customers maximise expected utility and are risk averters; then the monopolist may be able to devise various schemes which raise his expected profit above the Cournot level. For example, he might conduct his dealings in two stages. At the first stage, each customer may take a consignment of fixed size in return for a lump-sum payment; at the second stage, a price is chosen at random from a suitable distribution, and the monopolist buys or sells in the free (resale) market until the chosen price prevails. It is possible to construct cases where the terms of sale at the first stage and the probability

---

12. More precisely, the buyer may choose from a continuum of goods, and the distribution function of expenditure may be absolutely continuous at the point corresponding to the monopolised good. Incidentally, this assumption does not exclude the possibility of a positive density of net benefit or consumers' surplus at that point.

distribution at the second can be so chosen that the expected profit exceeds the Cournot level. The essence of the procedure is that the monopolist takes advantage of the customers' risk aversion. Of course, the scope for practical application of such schemes is very limited — they work best when goods are produced instantaneously without cost and cannot be stored by customers, and even then the additional profit is unlikely to be large. Nevertheless, they illustrate some typical subtleties of monopoly charging.

Because of these difficulties, it is necessary for the usual theory of simple monopoly either to assume (b) and (c) directly, or to introduce some considerations of law, convention, administration or communication from which these conditions can be inferred when (a) and (d) hold. Of course, if (c) is postulated directly, the usual results can be obtained without (d); but the assumption that a price rather than a tariff is quoted can hardly be maintained in general without (d) or some equivalent condition.

IV

Attempts to realise gains from trade by modifying the simple monopoly solution can operate through various mechanisms. Three possibilities are briefly discussed here, namely:

(i)   the replacement of straight pricing by two-part tariffs combined with the limitation of resale;

(ii) the formation of unions of customers to bargain about terms of sale;

(iii) vertical integration, in particular the acquisition of shareholdings in the monopoly by groups of buyers.

The discussion which follows applies with suitable changes to monopsony. Of course, in markets for personal services the mechanism of resale cannot operate in the manner usually assumed for goods markets; but its place in the analysis can often be taken by such devices as the replacement of direct labour by purchase or sub-contracting.

(i) *Two-part tariffs and limitation of resale*
The first device to be considered leaves the unilateral character of

the market in being — at least initially — but modifies conditions (c) and (d) of Section III. The change is connected with the assumption, which is basic to the theory of simple monopoly as usually stated, that people do not necessarily honour their contracts, or alternatively that enforcement of contracts is costly. This assumption, although sometimes realistic, has no essential connection with monopoly as such and is not made in competitive theory. If it is altered to conform with the competitive case, and appropriate changes are consequently made in conditions (c) and (d), Cournot's solution for simple monopoly in a world of certainty leaves scope for mutually profitable recontract and therefore is not an equilibrium in Edgeworth's sense.

For example, both the monopolist and any one buyer can be made better off by an agreement providing for (i) sale at a (marginal) price equal to marginal cost, (ii) an undertaking by the buyer not to resell, and (iii) a lump-sum payment sufficient to compensate the monopolist for any loss of profit due to the price reduction under (i), but not exceeding the buyer's gain in surplus. If a separate lump-sum can be fixed for each buyer, such contracts for two-part tariffs allow the realisation of an optimum (at the level of partial equilibrium). Of course, if the system of simple monopoly is superseded, and two-part charges are fixed unilaterally for all customers by the monopolist, it is not true to say that everyone will benefit by the change. On the contrary, the monopolist will be able to absorb practically the whole of the consumers' surplus through the fixed charges (at least to the extent that they can be varied among individuals). The buyers will generally be worse off unless they are somehow protected, for example by a law requiring the monopolist to offer, as an alternative to the two-part tariff, a straight price at the Cournot level. But the situation will be both an optimum and an equilibrium, at least so long as the market remains unilateral.[13]

13.   The use of two-part charging, both as a method of maximising profit by perfect discrimination and as a device for reconciling marginal cost pricing with solvency in public utilities with decreasing costs, has been known for some time. Both aspects are discussed in Paine C.L., 'Some Aspects of Discrimination by Public Utilities' *Economica 1937;* see also Lewis, W.A., *Overhead Costs,* Ch. 11, and Coase R.H., 'The Marginal Cost Controversy', *Economica 1946.* Paine realised (op. cit. p. 439) that the introduction of a two-part tariff might make all the monopoly's customers worse off, but wrongly inferred from this that the allocation of resources would then be worse than under simple monopoly. If

This argument is evidently subject to reservations. The enforcement of contracts forbidding resale may be costly in practice and perhaps even impossible.[14] The determination of differential lump-sum charges entails costs of information and administration which a simple monopolist could avoid. The attempt to appropriate most of the consumers' surplus might stimulate the formation of unions and convert the situation into one of bilateral monopoly, which according to circumstances might lead either to efficient contract or to deadlock.

Nevertheless it is possible to state a simple but important conclusion, which does not appear to have been noticed: the famous losses of welfare due to simple monopoly cannot exceed the costs of operating a scheme for the control of resale combined with a system of two-part tariffs. This statement assumes that the losses are measured, on the usual partial basis, with reference to a system of marginal cost pricing, and that the market remains unilateral when the two-part tariffs are introduced. (If a bargaining situation emerges the analysis is complicated by the possibility of deadlock before an efficient solution is reached; this point is briefly discussed below). The other mechanisms modifying the simple monopoly solution which are considered below define upper bounds for the loss of welfare in a similar way; the argument will not be repeated.[15]

some uniformity of charging is essential - owing perhaps to administrative or legal constraints, or to ignorance of the circumstances of particular customers - it may be advantageous to use more complicated tariffs, because a uniform fixed charge generally leads to the loss of some customers who would be willing to pay the costs of serving them. This problem is not pursued here.

14.   The difficulties are particularly serious where goods are distributed through independent intermediaries with overlapping market areas, or where they are used as components or materials in goods which compete in final markets. In such cases, all sorts of arrangements for sharing markets, vertical integration and restriction of resale prices might be needed in order to achieve the desired results.

15.   Upper bounds for welfare losses due to monopoly are estimated by Harberger A., 'Monopoly and Resource Allocation', *American Economic Review, Proceedings* 1954, and Schwartzman D., 'The Burden of Monopoly', *Journal of Political Economy* 1960. These studies suggest that the losses in practice are small. The procedure adopted is to apply to supposedly monopolistic industries the relations between welfare losses and rates of profit, or profit margins on output, derived from the theory of simple monopoly. This theory points to an increasing relation; but the use in monopolistic industries of two-part tariffs and other discriminating schemes may diminish welfare losses while increasing profitability. There is therefore no clear-cut connection between the theoretical argument given here and these estimates, except to suggest further reasons why

(ii) *Buyers' unions*

The creation of unions to deal with a monopolist may have two main objects: to acquire bargaining power by threats of collective action and thus to redistribute the gains from trade; and to create additional gains from trade by introducing more efficient contracts.[16]

The main obstacle to unions is the problem of indiscriminate benefit. Members have to bear costs of organisation and possibly losses from disruption of trade, which they will do only if some benefits can be reserved to them or losses can be inflicted on outsiders and late entrants. Where union members enjoy special terms of contract, incentives may be created for arbitrage through the resale market, which tends to destroy the union. The monopolist may take advantage of this mechanism deliberately to damage the union; and this threat is the more potent, the greater the privileges which its members enjoy. Finally, unions like other organisations may disintegrate owing to disputes over the division of their gains.[17]

In general, the market form which results from the formation of unions in simple monopoly is a complicated N-person game, which cannot be analysed in detail here. But it is necessary to consider more specifically whether, and in what circumstances, there is scope in the usual model for customers to combine in order to overthrow the Cournot solution. This depends on the assumptions, which usually are not stated very explicitly.

One difficulty occurs in connection with conditions (b) and (c), which exclude bargaining power and non-price contracts. Clearly these assumptions must somehow be modified if the problem under discussion is to have content; but merely to drop them might introduce new possibilities which have nothing to do with unions as such. It has already been mentioned that even a single individual might obtain

the upper bounds estimated for welfare losses probably exceed the true values. I am indeed indebted to Basil Yamey for this point.

16. Economies from joint negotiation and transport are ignored here. In principle they can be realised without the formation of a bargaining union, although it may be economical to use a single organisation for both purposes.

17. An interesting hypothesis about organisations in general, which merits further investigation, is that moderate success is best for the long-run survival: marginal enterprises are destroyed by fluctuations of fortune, outstanding ones by internal dissension.

discounts on intra-marginal units if he could convince the monopolist that he would not otherwise be a buyer; and similar discounts might be obtained by a union even if it represented only a negligible proportion of buyers. Unfortunately, there seems in general to be no simple way either to measure bargaining power or to separate the influence of size of purchase from an analysis of market strategy as a whole; and the influence of unions can generally be isolated only by considering the effect on this analysis of a change in the rules of the market game which relate to coalitions.

A second and crucial problem concerns condition (d), which stipulates costless, unrestrained resale in a perfectly competitive market. The assumptions that resale is costless and the market perfect are accepted here as data. But the restraints on dealings may change with the formation of a union — indeed, the essence of a union is the mutual promise to engage in specified dealings only with collective approval. If condition (d) is retained as it stands, the only possible union is one which has the sole right to deal with the monopolist on behalf of its members, but which cannot restrain their dealings in the resale market. Such a union could have no effect on relative bargaining power if intermediaries could buy from the monopolist and supply members at second hand. With comprehensive member-ship and complete loyalty, a union could threaten to cut off trade and so secure discounts which a single buyer could not obtain; though cohesion could hardly be maintained without some legal or social sanctions.

The usual assumption behind (d) is presumably that unions, if they can exist at all, suffer from the same disability in relation to the resale market as does the monopolist — namely that they either cannot enforce on their members contractual obligations to refrain from certain dealings, or can do so only at prohibitive cost. Such an assumption may in certain cases be acceptable for limited purposes; but in general the balance between the incentives and obstacles to organi-sation should if possible be analysed as part of the theory, not postu-lated in advance.

It may well be that the state of simple monopoly is a local, but not a global, equilibrium with respect to the formation of unions. A small union may be unable economically to control the dealings of its

members or to obtain for them any exclusive benefits which would repay the costs of organisation, and may therefore not be viable without special support from some extraneous source such as the government. But a large union, once it had come into existence and met the initial costs of organisation, might greatly reduce the cost of controlling the resale market; and it could, if it wished, impede the entry of new firms into the industry. On both counts, it would possess bargaining power; and its control of dealings at second hand would allow the use of contracts, such as two-part tariffs, which realise gains from trade not available under simple monopoly. Moreover, the monopolist might see a balance of advantage in the existence of a union if he thought he could appropriate a part of these gains; thus he might not merely refrain from attempts to destroy it, but even give support by allowing discounts to members or by refusing to supply outsiders. If such support were given even while the union was small, in anticipation of later benefits, the Cournot solution might not even be a local equilibrium. Of course, the move from simple monopoly to bargaining by unions is quite likely to be irreversible; a large union need not collapse if it loses the support of the monopolist and the government, because the initial costs of organisation have already been met, and because for most members it is usually an equilibrium strategy to conform to union policy as long as other members continue to do so.

These remarks envisage a single union bargaining with the monopolist. This is sufficient to allow the occurrence of cases where simple monopoly is not the only equilibrium; but in general the range of possibilities is very wide — including rivalry among unions, combinations between the monopolist and one group of buyers to exploit another group, 'closed shop' arrangements imposed jointly by a union and a monopolist in order to preserve gains from trade dependent on price discrimination, the same arrangement imposed on an unwilling monopolist by a union anxious to maintain its bargaining power, and various devices of inter-industry organisation.

Finally, a word about the limitations of unions as instruments of efficiency and redistribution. Even where a union manages to establish full bilateral monopoly, it must not be assumed too readily that contract is necessarily efficient. Apart from the costs of administration and negotiation, the possibility of obtaining favourable terms from the monopolist depends on threats to disrupt trade, which sometimes have

76

to be carried out. Inefficient arrangements which damage the interests of all parties often survive for long periods because terms for their elimination cannot be agreed. Sometimes they are kept in being because they allow the extent of damage to trade to be varied continuously in changing conditions, so that contracting parties can apply pressure without having to choose between complete disruption and quiescence. It must also not be assumed that a change from simple to bilateral monopoly invariably redistributes welfare in favour of buyers. If the buyers' union takes the form of an indissoluble merger, the limits of contract are widened, and the seller may be able to push the buyers closer to their no-trade indifference curves than under simple monopoly. Even if the buyers can invoke the possibility of dissolving their union in support of a claim for terms at least as favourable as in simple monopoly, they can enforce the claim only by actual dissolution, which may sacrifice their power to bargain for a long time ahead. The hope of future improvement may therefore lead them to accept worse terms in the short run.

(iii) *Shareholding schemes and co-operatives*

The last possibility to be considered is that the Cournot solution would be modified, in the direction of 'ideal' output, if most of the monopoly's shares were held by its customers. It might even be that the customers would be led by self-interest to acquire shares in order to influence policy. Of course, this is only one of several methods of integration and arrangement of property rights which could modify simple monopoly; it is selected for discussion because it is of theoretical interest and has apparently not been analysed.

The possibility that share ownership by customers might eliminate monopolistic 'waste' is suggested by the theory of consumers' co-operatives. It has been argued that the interests of co-operative members, who are both proprietors and consumers, are best secured by a policy of marginal cost pricing, which maximises the sum of producers' and consumers' surpluses.[18] Unfortunately, it can be shown that further conditions are necessary for static equilibrium with marginal cost pricing, and that in general they are satisfied simul-

---

18.   Enke S., 'Consumer Co-operatives and Economic Efficiency', *American Economic Review* March 1945. An analysis which takes account of various possible conflicts of interest among members of a co-operative is given by Yamey B.S., 'The Price Policy of Co-operative Societies', *Economica*, 1950.

taneously only by accident. It is assumed for the purpose of this analysis that members' purchases are influenced only by net prices, i.e. prices after deduction of prospective rebates (usually called 'dividends'). Profits also are computed after payment of these rebates; if the firm's business is neither growing nor declining, they are precisely equal to the sums (usually called 'interest') which are distributed to members in proportion to their shareholdings.

Suppose first that only a straight price can be charged. In general, there will be a conflict of interest among members as to its level whenever shareholdings are not proportional to purchases at the ruling price; and this may lead to the formation of a coalition strong enough to change the price, possibly after some changes in the society's membership. Further, if members can vary their shareholdings independently of purchases, it is a condition of equilibrium that the scale of the firm should be such that the return on shares equals the cost of capital; it should be recalled here that co-operative shares are issued and redeemed on request by the shareholder at nominal values, without capitalisation of anticipated profit. Leaving aside the possibility of induced shifts in the demand curve, this condition cannot in general be satisfied at the point where price equals marginal cost. Even if average costs (including the cost of capital) are increasing at that point, the society is not strictly in equilibrium unless it can limit the inflow of funds; and in general some members — possibly a majority — could be expected to resist such a limitation, in an effort to increase their share of the excess earnings. Presumably this difficulty would not arise if shares *just happened* to be distributed proportionally to purchases at marginal cost price at the time of the proposed limitation; in this sense equilibrium with marginal cost pricing is possible. But it cannot be secured by a rule which *formally ties* shareholding to purchases, since this leads effectively to average cost pricing. If average costs including the cost of capital are decreasing with output, the return on shares is below the cost of capital, and a policy of marginal cost pricing leads to an outflow of funds. In this case, the greatest feasible sum of surpluses is attained with average cost pricing. This is consistent with equilibrium if shareholding is proportional to purchases, and such a distribution can if necessary be maintained by a formal rule.

Suppose now that departures from straight pricing are possible, and consider for simplicity a system of two-part tariffs with a uniform

variable charge equal to marginal cost. Then the two conditions of equilibrium stated above can be satisfied if the difference between average and marginal costs can be absorbed by standing charges (or discounts) fixed independently of customers' purchases and share-holdings. The rate of return on shares is then equated with the cost of capital, and a scheme which ties shareholding to purchases can be used to ensure that the condition of proportionality is satisfied. Of course, the two conditions mentioned may not be sufficient for equilibrium, because further problems of strategy among coalitions arise in determining the pattern of fixed charges; this question is not pursued here.

The conditions of the problem are changed in two main respects if the monopoly is a quoted joint-stock company. Voting is by shares not by heads, so that permanent control can be bought and sold. And, once the shares have been issued, their value is adjusted so that the yield equals the cost of capital, provided that this is possible with a positive value; if not, the value is zero.

Suppose again that only a straight price can be charged and that average costs are increasing. If the shares have already been issued and just happen to be distributed in proportion to purchases, a system of marginal cost pricing is an equilibrium in a limited sense : no one of a large number of shareholders acting independently could be made better off by small changes in purchases, shareholding or pricing policy. If costs are decreasing, a system of average cost pricing is an equilibrium in the same sense (if it is assumed that the firm's capital is just sufficient to sustain production at that level; otherwise a lower price might temporarily be maintained). But in neither case is there global equilibrium, because it would pay someone to buy a controlling interest at the low share values corresponding to marginal or average cost pricing, in order to revive profit-maximising simple monopoly.

If control of the resale market allows departures from straight pricing, the analysis is similar to that for a co-operative. The main difference is that a take-over is always possible as long as profits, and therefore share values, are below the level corresponding to unilateral monopoly. (This is not to say that the consumer-shareholders must necessarily maximise profit, since their special organisation may allow a degree of control of the resale market, and therefore a level of profit,

not attainable by an ordinary monopolist.) Of course, to the extent that a new owner could control resale, a take-over need not mean a change to simple monopoly or a loss of gains from trade, though it may change the distribution of these gains.

At first sight this discussion suggests that shareholding by customers can do nothing to solve the problem of simple monopoly which could not be done by action in the product market alone. Such a conclusion would be incorrect. It is true that shareholding schemes alone can achieve little when customers are numerous, but they can contribute to the success of other methods. Fear of distributive consequences may lead a monopolist or a union of buyers to destroy schemes which could realise gains from trade not available under simple monopoly; for example, a monopolist might oppose a growing union, and a union might object to the gradual introduction of two-part tariffs. Shareholding by customers reduces the conflict of interests, and so diminishes the incentive to act in this way; and in general makes it less likely that gains from trade will be lost through deadlock in bargaining.

The preceding discussion considers only the conditions for an equilibrium with customer shareholding and marginal or average cost pricing, and does not ask how such a state might be attained. If the number of customers is large, there is in general no incentive for individuals to buy shares in the monopoly, so that this solution is unlikely to be adopted without some special organisation or legislative intervention. It is also clear, without going into details, that full attainment of the equilibrium conditions discussed above is not usually possible in practice. The situation is, of course, quite different for a monopoly selling only to a small number of large customers.

V

This paper has done no more than sketch a certain approach to the theory of monopoly; but a few general conclusions can be suggested. An examination of the model of simple monopoly shows that various assumptions are made which usually are not specified or are not stated precisely, and that the Cournot solution can be regarded as a full equilibrium only in special cases. In special circumstances the monopolist can, by unilateral action, raise his expected profit above the Cournot level even if no restriction of resale is possible. More generally, the self-interest of seller and buyers may bring into play a variety of

mechanisms capable of exploiting some of the gains from trade which in the Cournot solution remain unrealised. The choice among them depends on several considerations, including the costs of information and organisation, the possibility of reaching agreement without wasteful conflict, and flexibility in changing conditions. These mechanisms tend to transform simple monopoly into a complicated N-person game, with many strategies and many possible coalitions and, perhaps, with many solutions. The analysis merges with that of bilateral monopoly and oligopoly. It cannot be claimed that an equilibrium of such a game is necessarily a Pareto optimum, or even that all gains which in some sense are obtainable in practice will automatically be realised without intervention by government.

The kind of analysis pursued here has relatively little direct bearing on the traditional methods of state regulation of monopolies. It directs attention rather to the search for legal and institutional reforms which would assist the working of 'automatic' corrective forces. Unfortunately, arrangements which allow the realisation of certain gains from trade often have features which are regarded as undesirable in other respects. For example, gains which are not available in simple monopoly may be obtainable with the aid of such 'restrictive' practices as the prohibition of resale, discriminatory charging and the closed shop. In general, there is a complicated interdependence involving restrictive contracts, bargaining power, the distribution of gains from trade and the attainment of optimality, which can be adequately evaluated only if detailed attention is paid to the circumstances of particular markets. Nothing could be less conclusive in its implications for policy than a diagram showing that price in simple monopoly exceeds marginal cost. It is a merit of the approach used in this paper that it relates the pure theory of monopoly in a natural way to various phenomena of charging, bargaining, organisation and shareholding which occur in monopolistic markets. The resulting complexity is also its most serious limitation – the range of possibilities is too vast for a useful predictive theory. For this reason the method of analysis recommended by Edgeworth as 'applicable to the particular cases of imperfect competition' must remain to some extent a theory *in posse* – an approach which organises study and suggests the possibilities to be considered in particular instances, but which does not supply an exhaustive analysis of all conceivable solutions or a definite prediction of events.