

# The Concept of Monopoly and the Measurement of Monopoly Power<sup>1</sup>

## I

MONOPOLY, says the dictionary, is the exclusive right of a person, corporation or state to sell a particular commodity. Economic science, investigating the economic aspects of this legal right, found that they all resolved themselves into the implications of the power of the monopolist—as distinguished from a seller in a competitive market—arbitrarily to decide the price of the commodity, leaving it to the buyers to decide how much they will buy at that price, or, alternatively, to decide the quantity he will sell, by so fixing the price as to induce buyers to purchase just this quantity. Technically this is expressed by saying that the monopolist is confronted with a falling demand curve for his product or that the elasticity of demand for his product is less than infinity, while the seller in a purely<sup>2</sup> competitive market has a horizontal demand curve or the elasticity of demand for his product is equal to infinity.

The monopolist is normally assumed to tend to fix the price at the level at which he makes the greatest profit or “monopoly revenue.” This monopoly revenue constitutes a levy upon the consumers that the monopolist is able to appropriate for himself purely in virtue of his restrictive powers *qua* monopolist, and it is the consumers’ objection to paying this levy that lies at the base of popular feeling against the monopolist.

In addition to this it is claimed that monopoly is harmful in a more objective sense. A levy which involves a mere transference from buyer to monopolist cannot be said to be harmful from a social point of view unless it can be shown that the monopolist is less deserving of the levy than the people who have to pay it; either because he is in general a less deserving kind of person, or because the transference will increase the evils of inequality of incomes. But the levy is not a mere transference. The method of raising it, namely, by increasing the price of the monopolised commodity, causes buyers to divert their expenditure to other, less satisfactory, purchases. This constitutes a loss to the consumer which is not balanced by any gain reaped by the monopolist, so that there is a net social loss.

The nature of the loss here loosely expressed seems to have defied attempts at more exact exposition, the difficulties encountered on these attempts having

<sup>1</sup> The great advances made in the subject of this article since the major part of it was written—particularly in the work of Mr. Chamberlin and Mrs. Robinson—have rendered many parts of it out of date. In preparing it for publication, while cutting out some of these parts, I have been so much under the influence of this recent work that I cannot say how much of what is here published is really my own.—A. P. LERNER.

<sup>2</sup> “Pure” competition is different from “perfect” competition. The former implies perfection of competition only in respect of the complete absence of monopoly and abstracts from other aspects of perfection in competition. This useful distinction is suggested by Chamberlin. See his *Theory of Monopolistic Competition*, p. 6.

even induced some to declare that this commonsense view of a social loss is an illusion, while more careful sceptics prefer to say that nothing "scientific" can be said about it. The account given above clearly will not do as a general and accurate description of the nature of the social loss. Where a consumer spends as much as before on the monopolised commodity when the price is raised, he cannot be said to divert expenditure to other and less satisfactory channels, and where he spends more<sup>1</sup> upon the commodity than at the lower competitive price it might even be argued that there is a net social gain in so far as the consumer is induced to spend more on the commodity which is more urgently needed and less on other commodities! There seems little to choose between this argument and the counter-argument, that as long as the elasticity is greater than zero some consumer (or unit of consumption) is induced to change the direction of his expenditure so that he suffers the uncompensated inconvenience which constitutes the net social loss. Does this mean that if a man's demand is completely inelastic (so that the increased price brings no diminution in the amount of the monopolised commodity consumed and the whole of the levy is sacrificed ultimately in the form of other commodities) the expenditure of the income, as diminished by the amount of the levy, is not interfered with by the existence of the monopoly?—i.e. that if he had paid the levy in cash and prices were not affected he would have reduced his consumption of other commodities in the same way? Or is it more reasonable to suppose that a rise in a particular price will always tend to diminish purchases of the dearer commodity, where a cash levy (prices remaining unchanged) would diminish all expenditures in the same proportion so that if the same amount of the monopolised commodity is bought at the higher price, a cash diminution in income of the size of the levy would

<sup>1</sup> Where as much or more is spent on a commodity when the price is raised the elasticity of demand is equal to or less than unity. This may appear incompatible with the condition of monopolistic equilibrium that elasticity of demand shall be greater than unity (as long as marginal cost is positive).

There is, however, no incompatibility, for the two elasticities of demand are different things. The elasticity that has to be greater than unity for monopolistic equilibrium is the elasticity at the *point* on the demand curve corresponding to the position of monopolistic equilibrium. The elasticity that is equal to or less than unity when the amount spent on the commodity remains unchanged or increases as the price is raised, is the elasticity over the *arc* of the demand curve from the point of competitive equilibrium to the point of monopolistic equilibrium. The arc elasticity in this sense will normally be less than the point elasticity, as will appear from the diagram. If  $tT$  is the demand curve (here drawn a straight line),  $P'$  the point of competitive equilibrium, and  $P$  the point of monopolistic equilibrium, then the *point* elasticity at the monopoly

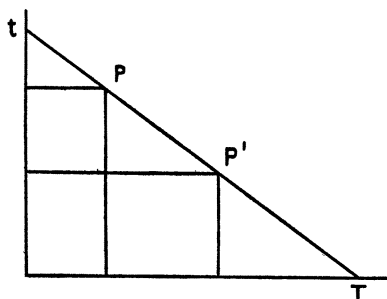


FIGURE 1

equilibrium will be  $\frac{PT}{Pt}$  while the *arc* elasticity will be  $\frac{P'T}{Pt}$ , which is smaller. The arc elasticity must

be smaller unless the demand curve is so concave (upwards) that it shows a constant or increasing point elasticity as price is lowered. The point elasticity at the competitive position will, of course,

be  $\frac{P'T}{Pt}$ . For the explanation of this definition of "arc elasticity," see my note on "The Diagram-

matical Representation of Elasticity of Demand," in No. 1 of the REVIEW.

have *increased* the demand for that commodity? The problems do not seem to be amenable to treatment on these lines.<sup>1</sup>

The commonsense attitude is, however, not easily balked. Another attempt was made to deal with the problem by Marshall, by means of the apparatus of consumers' surplus. If it is assumed that the marginal utility of money is unchanged, or that the change is so small that it may legitimately be neglected, it can be shown that the money value of the consumers' surplus lost is greater than the monopoly revenue gained, so that we have a theoretical measure of the net social loss due to the monopoly. There are, of course, many important weaknesses in this treatment, and some ways of applying it are completely wrong. The marginal utility of money can be considered unchanged only if we are considering a small change in the price of only one commodity. This makes it impossible to add the consumers' surplus obtained by an individual from different goods. Quite wrong is any attempt to speak of the consumers' surplus of a community and to derive it from the communal demand curve. And there are other traps to be avoided in this connection which are quite well known. But the exclusive preoccupation of teachers of economics with putting their pupils on their guard against these insufficiencies and dangers has tended to make them deny the problem with which the concept of consumers' surplus was intended to deal—the net social loss and its nature. It is not intended here to deny or even to belittle the dangers and confusions attendant on the use of the concept of consumers' surplus, but it does seem that some light can be thrown on the problem by its use.

From the consumers' surplus approach there has emerged a clarification of the rent element in monopoly revenue. It is only in the case of constant or decreasing average cost that the amount of monopoly revenue is necessarily less than the loss of consumers' surplus. The monopoly revenue will be greater if the average cost curve rises steeply enough. This gave the impression that the monopolistic restriction brought about a net social gain so that the competitive output was too great and it would be beneficial to tax industries which were "subject to diminishing returns." In correcting this view it was shown that against the monopoly revenue was to be reckoned not only the loss of consumers' surplus, but also the reduction in rents as compared with those receivable under competition. If the reductions of rent is not allowed for, the diminution of costs of the marginal units, as output is restricted, is attributed to all the *infra*-marginal costs where there has been no reduction in social costs, but only a transference of income from the receivers of rent. In the accompanying Fig. 2  $AR$  is the average revenue or demand curve (which, to avoid the quarrels over consumers' surplus, we can consider as the sum of a number of identical demand curves of similar individuals),  $MR$  is the marginal revenue curve,  $AC$  is average costs, and  $MC$  is marginal costs.  $P'$  will be the competitive point where output is  $OM'$  and price is  $M'P'$ , and  $P$ , which is perpendicularly above

<sup>1</sup> In the last few months Dr. J. R. Hicks and Mr. R. G. D. Allen have been making investigations on these lines and have demonstrated by means of the indifference curve apparatus that, with continuous indifference curves, an absolutely inelastic demand curve must be accompanied by a negatively sloping expenditure curve. This means that a change in income (prices remain unchanged) would bring about a change in *inverse direction* of the amount of the commodity bought. They have not been interested, however, in the problems dealt with in this article.

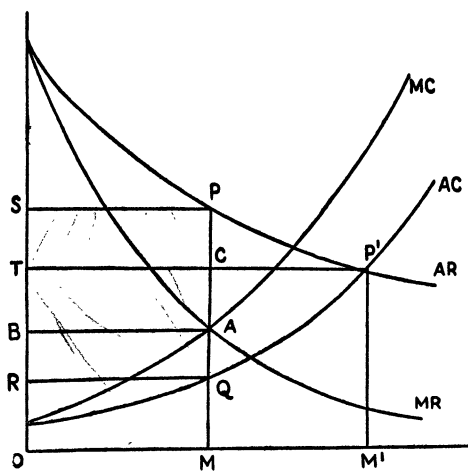


FIGURE 2

$A$ , where  $MR$  and  $MC$  cut, will be the monopoly point where output is  $OM$  and price is  $MP$ . Consumers' surplus lost is equal to  $SPPT'$ , while monopoly revenue is  $SPQR$ , which may be greater. But against this must be reckoned the loss in rents,  $RQP'T$ , so that there is a net social loss of  $PQP'$ .

One is tempted to divide the monopoly revenue  $SPQR$  into two parts,  $SPCT$  and  $RQCT$ , and to say that the former is the monopoly revenue extracted from consumers while the latter is the monopoly revenue extracted from receivers of rent or producers' surplus. It is exactly parallel to the extraction of monopoly revenue from the receivers of consumers' surplus,

but is obtained in virtue of the monopolist being confronted with a rising supply curve instead of with a falling demand curve. It is a gain obtained by a "single" buyer instead of a gain obtained by a "single" seller. The appropriate parallel name for it would be *Monopsony Revenue*.<sup>1</sup> This dichotomy of the monopoly revenue is based on a comparison of the monopoly position with the competitive position.<sup>2</sup>  $PC$  is the rise in price and  $QC$  is the fall in average cost, so that these quantities multiplied by the monopolistic output give the monopoly revenue and the monopsony revenue respectively.

It will, however, not do to compare the monopoly position with the competitive position for the purpose of making the dichotomy, for by this procedure it is made to depend upon the shape of the curves for outputs between the monopolistic output  $OM$  and the competitive output  $OM'$ , which may be a long way from it. It does not seem reasonable that the degree of monopsony or monopoly at output  $OM$  should be dependent upon what happens to demand or cost curves in the vicinity of output  $OM'$ . And apart from this the taking of the competitive output and price as a base from which everything is to be measured leads to more concrete inconsistencies. Thus we may attempt to find the amount of monopoly revenue, (in the more exact sense, that is, not including monopsony revenue) by considering what it would be if the average cost were constant at the competitive level so that there was no monopsony.  $AC$  and  $MC$  would then coincide with  $TP'$ , and the monopoly

<sup>1</sup> Joan Robinson, in *The Economics of Imperfect Competition*, introduces the word Monopsony, but does not speak of Monopsony Revenue.

<sup>2</sup> By monopoly position is meant a position in which the demand curve does not appear horizontal to all the firms in the industry. The simplest case of this is when there is only one firm which coincides with the whole industry, and that is what is shown in Fig. 2 at the monopoly position  $P$ . Monopoly is essentially a property of firms and by a monopolistic industry is meant nothing more than an industry in which firms have downward sloping demand curves. And, of course, only a firm is interested in maximising monopoly revenue. If the demand curve for the whole industry is horizontal, the industry is in a competitive condition, but that is only because in this case every firm in the industry must also have a horizontal demand curve—even if there is only one firm.

revenue would not be *SPCT* but some other larger amount, for the output could not be *OM* but some other amount. If we reverse this process, assuming that the demand curve and the *MR* curve are horizontal, we again find that the monopsony revenue is not *RQCT* but some other larger amount, and the output is not *OM* but, again, some other amount.<sup>1</sup>

The direct comparison of monopolistic with competitive equilibrium further assumes that cost conditions are the same and that demand conditions are the same. Neither of these is likely, and the combination of both is much less likely.

A more reasonable procedure for the allocation of the gains as between monopoly and monopsony revenue is to take as a basis not the price which would obtain if there were neither monopoly nor monopsony, but instead of that the actual conditions of the monopoly-monopsony equilibrium. With the given demand curve pure<sup>2</sup> monopoly output could only be *OM* if the horizontal *AC* curve were coincident with *AB*, in which case the monopoly revenue would be equal to *SPAB*. With the given *AC* curve the pure monopsony output could only be *OM* if the horizontal demand curve is coincident with *AB*, in which case the monopsony revenue would be equal to *RQAB*, and *RQAB* and *SPAB* do add up to the monopoly-monopsony revenue *SPQR*.

From this it appears that the monopoly revenue per unit of output, *AP*, is the excess of price over marginal cost, so that the mark of the absence of monopoly is the equality of price or *average* receipts to *marginal* cost, and the mark of the absence of monopsony is the equality of *average* cost to *marginal* receipts.<sup>3</sup>

The test more usually accepted is the equality of average costs to price or average receipts. It is this equation which is regularly given as the definition of

<sup>1</sup> In Fig. 2, where both *AR* and *AC* are concave upwards, the output under monopoly without monopsony would be less than *OM*, and the output under monopsony without monopoly would be greater than *OM*. The outputs are given by the abscissae of the points where *TP'* is cut by *MR* and *MC* respectively. If *AR* and *AC* are convex, the outputs would move in the opposite direction. If they are straight lines, or if the convexity of one is just offset by the concavity of the other, the output will be the same as when the monopoly and monopsony are found in combination. If the elimination of monopsony changes the output in one direction, the elimination of monopoly would change output in the other direction, and *vice versa*.

<sup>2</sup> By *pure monopoly* is meant a case where one is confronted with a falling demand curve for the commodity one sells, but with a horizontal supply curve for the factors one has to buy for the production of the commodity; so that one sells as a monopolist but buys in a perfect market. Similarly, *pure monopsony* stands for perfect competition in the market where one sells, but monopsony in the market where one buys—being confronted with a horizontal demand curve but a rising supply curve. *Pure monopoly* is monopoly free from all elements of monopsony. *Pure monopsony* is monopsony free from all elements of monopoly. *Pure competition* stands for freedom from all elements of both monopoly and monopsony. The *purity* of monopoly or of monopsony has nothing to do with the *degree* of monopoly or monopsony.

<sup>3</sup> *Marginal* cost and *marginal* receipts are, of course, always equal to each other in any equilibrium, whether monopolistic or monopsonistic, or both or neither. It is, therefore, possible to express the same relationships in terms of the equality of price or average receipts to marginal receipts and the equality of average costs to marginal costs. But this procedure rules out conditions of disequilibrium together with monopoly or monopsony, so that to affirm this would be merely to say in other words that the demand or supply curve is horizontal, so that by definition there is no monopoly or monopsony. The relationships given in the text are not the merely mathematical relationships between an average and its corresponding marginal curve, but between real conditions of costs on the one hand and of receipts on the other. It will be seen below that these relationships will not always coincide with the tautologous alternatives suggested in this footnote.

"competitive" position,<sup>1</sup> and a suggestion like the one here given is likely to meet with a lecture on the impropriety of comparing averages with marginal values. It would seem, however, that the orthodox point of view is not only based upon too great a readiness to consider perfect competition as the ideal type of economic phenomena towards which all things tend, but are deterred more or less only by "frictions" (for in perfect competition all these equations become identical), but is in some measure induced by the habit of using straight lines in diagrams dealing with monopoly, and thus missing the problem. For in this case,  $AB$  of Fig. 2 would coincide with  $P'T$ , and the two dichotomies of the monopoly-cum-monopsony revenue are identical.

The point at issue is not merely a verbal one of definition—a quibble as to what it is better to call the "competitive" position. The importance of the competitive position lies in its implications of being a position which in some way or another is better than other positions. It is the position in which the "Invisible Hand" has exerted its beneficial influences to the utmost. It has become the symbol for the social optimum. Its importance for us here is in giving us a basis against which we can compare the effect of monopoly in order to see the social loss, if any, that the existence of a monopoly brings about. Is the social optimum that position at which prices are equal to average cost, or that at which price equals marginal cost and average cost equals marginal revenue?

The social optimum relative to any distribution of resources (or income) between different individuals (and we cannot here go into the problems connected with optimum distribution) will be reached only if the resources which are to be devoted to satisfying the wants of each individual are so allocated between the different things he wants, that his total satisfaction would not be increased by any transference of resources from the provision of any one of the things he gets to any other thing he wants. This would show itself in the impossibility of any individual being put in a preferred position without putting another individual in a worse position. We may adopt this as our criterion or test of the achievement of the relative optimum. If in any set of circumstances it is possible to move one individual into a preferred position without moving another individual into a worse position (i.e. such that the original position is preferred to it by the individual affected), we may say that the relative optimum is not reached; but if such a movement is impossible, we may say that the relative optimum has been attained. The conditions which must be satisfied if the optimum is attained can be formulated quite simply.

Any change in the position of any individual means a change in the quantity of goods (and services) he consumes. For any such a change to take place it is necessary that there shall be either (a) a *similar* change in the total quantity of goods produced or (b) an *opposite* change in the total quantity of goods consumed by others, or (c) some combination of (a) and (b). In the case of (a), consumption by other people need not be interfered with by the change,

<sup>1</sup> Even Mrs. Robinson defines "competitive output" and "competitive price" as that output or price at which  $AC=AR$  or price (*op. cit.*, p. 160), although she demonstrates most clearly in other parts of the book how this condition ( $AC=AR$ ) is also reached in monopolistic or imperfectly competitive equilibrium.

the whole change in the consumption by one individual being covered by changes in production. In the case of (b), there need be no change in production, any increase in the consumption of particular goods by one individual be provided by decreases in their consumption by others, and any decreases in the consumption of other goods by one individual being covered by increases in their consumption by others. In case (c) both kinds of compensating movements take place, but these can be separated and dealt with as cases of (a) and (b) so that no special treatment is necessary.

If a change in the consumption of various goods by one individual which improves his position is compensated solely by a movement of type (a), consumption by all other individuals need not be affected. This means that the effect of the movement from the previous position was to make one individual better off without making any other individual worse off. The previous position could not, therefore, have been an optimum position. One condition, then, of the optimum position is that any change in the quantity of goods consumed by

any individual which improves his position cannot be compensated by a movement of type (a).

This is illustrated in Fig. 3,<sup>1</sup>  $PP'$  is a section of the displacement cost curve (or productive indifference curve) of the whole community.  $I$  and  $II$  are consumption indifference curves of one individual. The indifference curves are superimposed upon the displacement cost curve, so that the point on the indifference map which represents the quantities of the commodities  $X$  (measured horizontally) and  $Y$  (measured vertically), consumed by the individual in the initial position, coincides with the point on the communal displacement cost curve which represents the total amount of the commodities ( $\bar{X}$  and  $\bar{Y}$ ) produced in the whole community in the initial position. If  $P_1$  is this position, a movement from  $P_1$  to any point above  $I$  represents a movement favourable to one individual. Compensating movements of type (a) from  $P_1$  are, however, limited to points below  $PP'$ . The shaded area in the diagram represents positions to which movements from  $P_1$  are favourable to one individual and can be compensated by movements of type (a). Thus a movement from  $P_1$  to  $P_0$  represents a diminution in the production of  $Y$  by an amount  $P_1Q$  and an increase in the production of  $X$  by an amount  $QP_0$ ;<sup>2</sup> accompanied by a similar change in one individual's con-

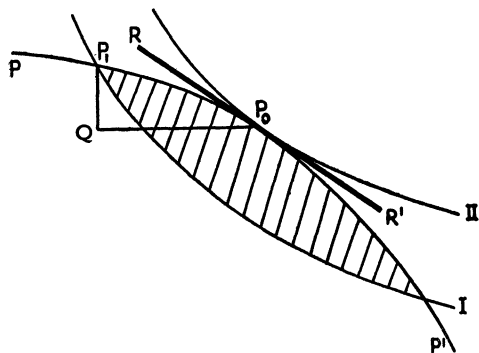


FIGURE 3

sumption by the individual in the initial position, coincides with the point on the communal displacement cost curve which represents the total amount of the commodities ( $\bar{X}$  and  $\bar{Y}$ ) produced in the whole community in the initial position. If  $P_1$  is this position, a movement from  $P_1$  to any point above  $I$  represents a movement favourable to one individual. Compensating movements of type (a) from  $P_1$  are, however, limited to points below  $PP'$ . The shaded area in the diagram represents positions to which movements from  $P_1$  are favourable to one individual and can be compensated by movements of type (a). Thus a movement from  $P_1$  to  $P_0$  represents a diminution in the production of  $Y$  by an amount  $P_1Q$  and an increase in the production of  $X$  by an amount  $QP_0$ ;<sup>2</sup> accompanied by a similar change in one individual's con-

<sup>1</sup> I am indebted to Mr. V. Edelberg for the suggestion of the application of the indifference curve apparatus to the problem in this manner.

<sup>2</sup> It is not necessary that all or any of the identical units of factors set free from the production of  $Y$  should be used in the production of  $X$ . They, or a part of them, may go to the production of a third commodity  $Z$ , as substitutes for other factors which are released to produce the additional  $X$ ; and there may be any number of such steps. This, of course, does not mean that every commodity is a *direct* displacement cost for every other commodity at the margin (in the

sumption which moves him on to the higher indifference curve  $II$ ; while the quantities of goods remaining to be consumed by other people are unaffected.

It is, of course, not necessary that any improvement should go up to the highest possible point—here  $P_0$ . A movement from  $P_1$  to any other point in the shaded area indicates an improvement, but leaves room for still further improvement.

Such a movement is possible as long as the indifference curve cuts the displacement cost curve, giving an overlapping (shaded) area. Our first condition for the optimum position can be expressed by saying that these curves must not cut.

If the curves are smooth this will mean that they are tangential as at  $P_0$ , but our condition is satisfied without the tangency of the curves, if either (or both) of the curves changes directions suddenly at the point where the curves meet or that it forms an angle. What is necessary is merely that the curves shall meet at  $P_0$  without cutting. This condition must be fulfilled for every individual in the community.

The movement of one individual to a preferred position, may, however, be covered by opposite changes in the consumption of others. This, too, can be examined in the same diagram. Let  $I$  and  $II$  represent the same indifference curves as before, but let  $PP'$  represent now not the displacement cost curve, but the indifference curve of any other individual, turned through  $180^\circ$  around the common point which shows the combinations of goods consumed by the individual. If the indifference curves cut, as they do in our diagram if  $P_1$  is the common point, there is an overlapping area, shaded in the diagram, showing the possibility of improving the position of one without worsening the position of the other. A movement from  $P_1$  to  $P_0$  improves the position of *one* individual and leaves the other at another point on the same indifference curve  $PP'$ , and, therefore, not worse off. Movements from  $P_1$  to any intermediate point in the shaded area would make both individuals better off. In order to satisfy the condition of the optimum it is therefore again necessary that there should be no gap between the curves, i.e. that they should not cut. If they are smooth, it means that they are tangential, and that the slopes of the indifference curves of both individuals were parallel in the initial position, since the turning of a curve through  $180^\circ$  does not change any slopes.

The diagrammatical treatment restricts one to the consideration of only two commodities. This does not matter for the present purpose, since the relationships described have to obtain for every pair of all the commodities in the economy. This is because the failure of the conditions to be satisfied for *any* pair of commodities shows a possibility for improvement which is incompatible with an optimum position.

If both of these conditions are satisfied, as between each individual's indifference curves and the communal displacement costs curve on the one

sense that factors can move directly from one to the other without economic loss), as would be the case if each factor had the same marginal productivity in all uses—universal substitutability of factors at the margin. It only means that there is some path, however indirect, whereby a diminution in the production of one commodity permits an increase in the production of any other commodity, leaving the quantity of the rest of the commodities unaffected. That is what is meant by drawing a displacement cost curve for any two commodities.



hand, and as between each individual's indifference curves and every other individual's (inverted) indifference curves, on the other hand, it is impossible to improve the position of any individual without worsening the position of some other individual. The optimum position, relative to the distribution of income between individuals, is attained.

Can we make any use of such a complicated set of conditions? If it were necessary to investigate separately the slopes of the indifference curves of all individuals for all pairs of commodities in order to discover whether the conditions are satisfied, it would be most profitable to discontinue this analysis at once. But there is no need for all this. We need merely assume that some of the indifference curves are smooth at the positions representing the amounts consumed by the individuals, and that each individual, in buying goods for his own consumption, considers the price as given. Under these conditions the relative prices of each pair of goods in the market will accurately reflect the slopes of the indifference curves where these are smooth; and for those cases, where an indifference curve forms an angle, the ratio between the prices will give a line ( $RR'$  in Fig. 3) of such slope that the indifference curve will lie wholly *above* it, meeting it but not cutting it if it is superimposed on the consumption point  $P_0$ . The mere existence of a free market in consumption goods thus satisfies the second of our two conditions.

The first condition is satisfied if the price ratio on the market, represented by the slope of the line  $RR'$ , is such that the displacement curve lies wholly *below* it, meeting it at the production-consumption point  $P_0$ , but not cutting it. If the displacement cost curve is smooth and, therefore, tangential to  $RR'$ , this will mean that the price ratio is proportional to the marginal displacement costs, which condition is satisfied if *price is equal to marginal cost*.

From this analysis we see that the optimum is reached when the price reflects the alternatives given up at the margin, whether this alternative is considered in physical terms of some other commodity or whether we go direct to the satisfactions that the physical alternatives represent. The loss involved in monopoly can be seen in the divergence between price and this marginal cost. The loss involved in monopsony is of exactly the same nature, and a parallel analysis is rendered unnecessary if we translate the rising supply curve that is seen by the monopsonist into a falling demand curve by considering the purchase of  $A$  for  $B$  as the sale of  $B$  for  $A$ . This loss is avoided only if price to the consumer ( $AR$ ) is equal to marginal cost ( $MC$ ), and if the wages of labour ( $AC$ ) are equal to its marginal product ( $=MR$ ). If we prefer we may put the latter statement in the form of demand. The price of leisure demanded by labourers ( $AR$ ) (which is his wage) must be equal to the marginal cost of his leisure ( $MC$ ) (which is equal to the marginal product of the labour withdrawn).

## II

In considering the degree of monopoly in a particular field one's first inclination seems to be to hark back to the etymological meaning of the word and to see how close the situation is to the conditions which accompany a "single seller." On this line one would say that there is complete monopoly if

there is actually only one seller, and that the monopoly element diminishes as the number of sellers increases. One could construct some kind of index of the degree of monopoly, such as the inverse of the number of sellers, which would give values ranging from unity in the case of this kind of "complete" monopoly to zero in the case of an infinite number of sellers.

The most obvious of the many reasons why this will not do is that there may be a very high degree of monopoly (in any sense other than that of the formula for such an index), even where there are many sellers, if one or two sellers control a sufficiently large proportion of the total supply. For this reason one turns instead to discover how great a proportion of the total supply is controlled by one or a few individuals or organisations. The same information may also be sought more indirectly by inquiries into the size of firms.

This procedure, however, is still quite inappropriate for measuring the degree of monopoly if we are interested in its economic and social implications of control over price and social loss as discussed in the first part of this paper. This is seen most clearly when we observe that control by a single firm of 100 per cent. of the supply of a commodity for which the demand is infinitely elastic (which will always be the case if there is some equally satisfactory substitute available at a constant price) is absolutely unimportant and has no economic significance, while a "partial" monopoly of a commodity for which the demand is inelastic may be able to raise price by reducing output and is clearly a much more effective case of monopoly.

The statistical method of measuring monopoly, besides missing the main issue in this way, encounters enormous practical difficulties in which investigators can hardly hope to avoid getting entangled. The problems of allowing for changes in taste and technique, in transport and in business organisation, of dealing with firms making many products and of discovering the degree to which different firms compete with one another or mitigate the competition by Gentlemen's Agreements, trade conventions, business alliances, and so on, are just a few worth mentioning, but there is one that interests us particularly here, and that is the relatively simple one of defining the commodity.

A man may have a considerable degree of monopolistic power although he is in control of only a very small part of the supply of a commodity if he is afforded some protection from the competition of the rest of the supply by the cost of transporting other supplies to his market. Under these conditions the price of the commodity will be different in different places. The best way of dealing with this is to declare that objects having the same physical characteristics are not the same goods if they are at different places. Location is an essential and distinguishing characteristic of economic goods, and the only relationship between the prices of similar goods in different places is that which results from the possibilities of transforming the one good into the other by transporting it from the one place to the other.

And location is not the only variant of this kind, but rather the simplest species of a large genus, and is useful for a simplified exposition of the problems involved. Every specialised gradation of every particular quality of every "commodity" may be treated as "distance," and the cost of changing the quality to a particular grade as the cost of "transport." Some of these problems

are dealt with by Hotelling in his article, "Stability in Competition," *Economic Journal*, 1929, p. 41, where he gives examples ranging from the sweetness of cider to the service of churches.

To these variants must be added also all fictitious variations, such as are successfully imposed upon the minds of buyers by skilful advertising, as well as the tendencies of customers to buy from one seller rather than from another by sheer force of habit. Here the "distance" is the fictitious difference in quality or the goodwill of the customer, while the "transport costs" are the costs involved in overcoming the "goodwill" whether by reducing price or by counter-advertisement.

This splitting up of the conception of a commodity of course multiplies the number of commodities indefinitely, and seems to create monopolies in the most unexpected places. Carried to its logical extreme, every firm now becomes a monopoly, since it is impossible for more than one unit of product to be in the same place. But even without going to such extremes it becomes impossible to apply the simple measures of monopoly that we are criticising. Further difficulties are yet to arise.

While the idea of considering the same things at different places as different goods seems to have spread considerably, the full revolutionary implications of this step forward in the picturing of the equilibrial forces do not seem to have been quite realised.

*In calling the same thing at different places different commodities, we have rejected the criterion of physical similarity as a basis for the recognition or classification of commodities and have put in its place the principle of substitutability at the margin.*

If the same thing at a different place is not the same commodity it is only because the difference in its location prevents it from being substituted for, or used in the same way as, the same thing here. But this principle can be applied in the converse form too. With substitutability as the principle it is no longer necessary for different units of the same commodity to have the same physical characteristics as long as they are substitutable at the margin for the purpose that the buyer wants them. This means that if one pound of coal gives me the same heating power as four pounds of wood, that both of these items cost the same on the market, and I am indifferent as to which I have, then one pound of coal and four pounds of wood represent the same number of units of the same commodity. It means, further, that if I am indifferent as to whether I have one hundredweight of coal every week during the winter, or an overcoat to keep me warm, then a winter's coal and an overcoat are equal quantities of the same commodity. Further still, if I am indifferent as to whether I have a wireless set for £10 or whether I have the satisfaction of saving ten Chinese children from starvation, the wireless set in London is the same quantity of the same commodity as £10 worth of rice in China; while if I get the same satisfaction from a £100 motor-car here and now as I could from a Mediterranean cruise next year, which costs £100 plus the accumulated interest on the money, then the motor-car here and now and the Mediterranean cruise next year are equal quantities of the same commodity. Physical qualities, spacial and temporal position are irrelevant now that we have the ultimate criterion of substituta-

bility at the margin. If any quantity or complex of goods and services can be substituted at the margin for any other quantity of goods and services (and therefore have the same market value), then they are both equal quantities of the same commodity. It would perhaps be best to give terminological recognition to such a break with traditional usage by speaking of "units of accommodation" instead of units of commodities.

If this way of looking at things seems paradoxical, it is only because we have not yet completely freed ourselves from the crudely materialistic conception of goods with which the Physiocrats and Adam Smith were the first to wrestle. The inadequacy of a purely physical criterion of commodities is obvious when we consider the enormous physical difference which we neglect if they do not affect the qualities in which we are interested (that is which affect our satisfactions), of which we are often completely unconscious, but which are of so much importance to Mr. Sherlock Holmes. Physically there are no two similar articles even apart from location. If two objects are considered to be items of the same good, it is only because they are "good for" the same purpose—always, ultimately, the satisfaction of a want. It is futile to say that the motor-car and the Mediterranean cruise satisfy different wants until we are able to define "similar" wants otherwise than as wants that are satisfied by physically similar objects. There is no *qualitative* criterion of wants. Wants can only be considered as similar when the person who feels them displays equal concern for their satisfaction and thus shows them to be equal in *quantity*. To follow any other course is to sacrifice the logic of the science to the irrelevant convenience of the shopkeeper.

It may be objected that this concept of commodity is so abstract and elusive as to be unusable. That is perfectly correct. But therein lies a great part of its advantages. It cannot be used like the more material conception to drown the theory in irrelevant statistics. It puts an end to attempts, here, to find a measure of monopoly in terms of the proportion of the supply of a commodity under single control and clears the way to a better understanding.

Another line of approach that suggests itself is to compare the amount of monopoly revenue with the total receipts, and to take this ratio as a measure of the degree of monopoly power. Allowance is thus made for the size of the industry or the firm. We will obtain values ranging from 0 in the case of perfect competition to 1 where the whole of receipts is monopoly revenue, and at first glance all seems well.

This procedure will, however, not do, for what we want in the measure of monopoly is not the amount of tribute individuals can obtain for themselves from the rest of the community, by being in an advantageous monopolistic position, but the divergence of the system from the social optimum that is reached in perfect competition. From this point of view the monopolist gains are not to be distinguished from rents of scarce property that he owns, or any other source of individual income. The independence of the monopolist gain from the social loss can perhaps most clearly be brought out by a consideration of how far they can vary independently. The limiting case is seen where the demand curve for the product of a monopolist coincides over considerable range with his average cost curve. Here the monopoly revenue is zero wherever

the monopolist produces within this range, yet he has control over price, and the social loss will be different according to what output the monopolist decides to produce. It clearly will not do to say that the degree of monopoly power in such a case is zero.

If the average cost curve is horizontal such a divergence cannot occur. The firm can only change output while keeping monopoly revenue zero if the demand curve is also horizontal, and that means perfect competition in either case and no social loss. But in such a case we are comparing not merely monopoly revenue with total receipts, which is the same as the ratio between average receipts minus costs and average receipts (and which is also seen in the ratio between average costs and average receipts), but also *marginal costs* with *average receipts*, and it is in divergence between these, as we have seen above, that the essence of monopoly is to be found.

In such cases (where the cost curve is horizontal) the ratio of monopoly revenue to total receipts coincides exactly with the ratio of the divergence of price from marginal cost to price, and it is this latter formula that I wish to put forward as the measure of monopoly power. If  $P$  = price and  $C$  = marginal cost, then the index of the degree of monopoly power is  $\frac{P-C}{P}$ .

It will be observed that this formula looks like the inverse of the formula for the elasticity of demand. It differs from it only in that the item marginal cost replaces the item marginal receipts. In equilibrium as normally conceived marginal costs coincide with marginal receipts so that our formula becomes identical with the inverse of the elasticity of demand. It will be best to consider this as a special case.

In this special case we can find the degree of monopoly power via the elasticity of demand. The determination of this elasticity of demand is not to be confused with that of Pigou and Schultz in finding the elasticity of demand (as part of the demand function) for a materially (physically) defined commodity on a market. What we want here is the elasticity of demand for the product of a particular firm. This is much easier to obtain, for it is only when he knows the shape of the demand curve for his product that any entrepreneur can obtain his maximum profit; and he is, therefore, always applying himself energetically to obtaining as accurate an estimate as possible of this elasticity. This does not mean that the entrepreneur will be able to fill in the elasticity of demand on a questionnaire form. He will rarely know what the term means. But his unfamiliarity with the technical jargon of economists must not be held to show an ignorance of so primary a principle for intelligent business management as the urgency of knowing the effect of price changes on sales. His behaviour in running the business for maximum profit will enable any student to deduce the (estimated) elasticity of demand from the firm's cost curve and the selling price. From the average cost curve the marginal cost curve can be derived. The marginal cost is equal to the marginal receipt, output being adjusted so as to make them equal if profit is maximised. The elasticity of demand is equal to the price divided by the

difference between price and marginal cost—it is the inverse of our formula for the measurement of the degree of monopoly power.

In finding the degree of monopoly in this special case “via the elasticity of demand” we found that the easiest way of finding the elasticity of demand was via the degree of monopoly. We may, therefore, leave out the elasticity of demand altogether and just keep to our formula all the time. In the special case both come to the same thing, but we must use the new formula and not the inverse of the elasticity of demand whenever we consider cases where the maximum monopoly revenue is not obtained in practice.

This may be accidental, as when the monopolist does not know the shape of his demand curve and his estimate of the elasticity of demand at the actual output is erroneous; or it may be intentional. The price and output may intentionally be fixed in a manner which does not give the maximum monopoly revenue:

- (a) When the monopolist is not working on purely business principles, but for social, philanthropic or conventional reasons sells *below* this price commodities which it is considered socially desirable to cheapen—as when a public authority supplies cheap transport facilities—or sells *above* this price commodities which are considered socially harmful—as may be done by a State liquor monopoly.
- (b) When the monopolist is working on purely business principles, but keeps the price and his profits lower than they might be so as to avoid political opposition or the entry of new competitors. The second could, perhaps, better be considered as a case where the demand is more elastic in the long period, taking into account the contingent competition, than in the short period, and where the monopolist takes a long period view.

In all such cases our formula is not equal to the inverse of the elasticity of demand; but wherever there appears a divergence between the two it is our formula and not the inverse of the elasticity of demand which gives the measure of what we want. In the first case—where the monopolist’s estimate of the elasticity of demand is erroneous—the consumers will in every way be in exactly the same position as if the elasticity were what the monopolist thinks it is. If he over-estimates the elasticity of demand he will sell a larger amount at a lower price. If he thinks the elasticity is infinite—i.e. that if he produced less he would not be able to get a better price—he will make price equal to marginal cost, and the effect on consumers will be the same as if there were perfect competition.<sup>1</sup> The unused monopoly power will be there, but being unknown and unused it is, economically, as if it were not there. For practical purposes we must read monopoly power not as *potential* monopoly, but as monopoly *in force*.

If the monopolist underestimates the elasticity of demand he will sell a

<sup>1</sup> Mrs. Robinson has pointed out to me that the delusion that elasticity is infinite would persist only if MC happened to equal price already. This is the easiest case for the correction of a mistaken estimate in the process of adjustment to it. The same possibility exists with any estimated elasticity of demand as long as the marginal cost and the estimated marginal receipts do not coincide and so preclude any adjustments.

smaller quantity and at a higher price than at the point of maximum monopoly revenue. The only difference between this and the previous case is that the monopolist's error brings a loss to consumers instead of a gain. The monopolist himself, of course, loses by the error in either case. The consumer here has to pay a higher price or else do without. It is again just as if the elasticity of demand were what the monopolist thinks it is. This may sound as if the monopoly *in force* is here greater than the *potential* monopoly power, but the inverse of the elasticity of demand at the maximum revenue point does not really give the potential monopoly power. It gives just that degree of monopoly power which it is necessary to put into force in order to obtain the maximum revenue and which is in force where the maximum revenue is being obtained. The monopolist always has power in excess of this ; but as the employment of it can only bring him a loss, he normally does not use it intentionally. If he chooses to use it he can, of course, for the exercise of this power consists of diminishing the amount he produces. Potential monopoly power is only used to its maximum when the monopolist stops all production. What our formula gives is the degree of monopoly power in force.

The same arguments apply to cases where the maximum monopoly revenue is not obtained for social, philanthropic or conventional reasons or for the purpose of avoiding political opposition or contingent competition. In the last case, our procedure saves us all further investigation into the complications involved in considering the length of the period upon which the demand curve is based. The appropriate costs to be reckoned are those of the present, or rather of the immediate future, so as to enable us to measure temporary monopolies. The degree of monopoly over a long period is perhaps best expressed in an average of the short-period monopolies over the period.

The primary unit to which our measure of monopoly applies is the firm in the very shortest period. In order to get a measure of monopoly over a period we had to take an average of such coefficients of monopoly. In order to get a measure of monopoly over an industry we have to follow the same procedure and find an average of monopoly of the separate firms included in the industry. The "industry" is to be considered as a group of firms, chosen for the purpose of the special investigation. It is quite unnecessary, for this purpose, to say anything at all about the "commodity" which the "industry" produces, nor is there any need to be able to draw demand or supply curves for the industry. All the difficulties of definition of "commodity" or "industry" are completely avoided.

More strictly a simple average of the degrees of monopoly in firms may be used to indicate the degree of monopoly in an industry only in the very limited sense of the degree of monopoly *at that stage*. It is not a measure of the degree to which the application of the resources of the community to the production of the products of the "industry" diverges from the social optimum. That depends upon two other sets of conditions in addition to this *local* element of monopoly.

The first of these is the degree of monopoly in the firms (or "industries") producing the raw materials for all the previous stages in the production of the products. The restriction of productions in any stage has its effects in all the

succeeding stages. The final degree of reduction of product will depend upon the degree of monopoly in all the preceding stages. These have to be aggregated so as to give the tendency to divergence from the social optimum in the whole series of the production stages of the product ; this phenomenon may be called the transitiveness of monopoly.

Theoretically, this can be done quite simply. What we want is the divergence between the price of the product and its marginal *social* cost. If in all the previous stages price is equal to marginal cost, the marginal cost to the firm is also the marginal social cost. If in any stage there is a divergence, price being above marginal cost, that divergence is a gap in the social cost. The social cost can then be calculated by multiplying the price by a factor for each stage in production, each factor being the ratio of the marginal cost to the price in the corresponding stage. Thus, if there are five stages and in each stage the degree of monopoly is  $\frac{1}{5}$ , marginal cost over price in each stage is  $\frac{4}{5}$ , the social cost is  $(\frac{4}{5})^5$  of the price of the final product, and by our formula the "social" degree of monopoly is  $1 - (\frac{4}{5})^5$ .

Practical difficulties that arise in attempts to measure the "social" degree of monopoly, or different products may be attacked by any of the tricks of the trade of mathematical statistics. It may be necessary to assume average degrees of monopoly in separate stages and to calculate "social" degree of monopoly by the number of stages, and so on ; but it is not intended here to discuss anything but the simplest theoretical implications.

The second set of complicating considerations arise when we ask the even more ambitious question : What is the (social) degree of monopoly in the society as a whole ? From this general point of view the conditions for that optimum distribution of resources between different commodities that we designate the absence of monopoly are satisfied if prices are all *proportional* to social marginal cost. If the "social" degree of monopoly is the same for *all* final products (including leisure) there is no monopolistic alteration from the optimum at all. The absolute height of "social" degrees of monopoly becomes completely unimportant.

This is because if the "social" degree of monopoly is the same for all products it *must* be equal to zero in real terms. For from the social point of view, the marginal cost of any product is always some other product. If the "social" degree of monopoly for product *A* is positive, this means that the price of *A* is greater than the price of some other product *B* which is the alternative foregone. The price of *B* cannot then be greater than the price of *A*. If both degrees of monopoly are equal they must both be zero.

What is important is the deviations between the degrees of monopoly ; and it is this which must be measured in order to answer our question. A suitable measure for this is the standard deviation of the "social" degrees of monopoly of all final products in the society.

Another complication arises in the growingly important cases where it is found to be profitable to extend or maintain the amount sold, not by reducing price but by expenditure on advertising, salesmanship, gifts, coupons and beautiful wrappings—all of which can be subsumed under the heading of "marketing costs." In such cases what becomes of the elasticity of demand ?



In the recent cost controversy, "marketing costs" were eagerly seized upon in attempts at a conciliation between decreasing costs and competitive equilibrium.<sup>1</sup> Such arguments may be described with some justification as contriving to exhibit decreasing costs at peace with competitive equilibrium by the device of leaving out of account the marketing element in the costs which is increasing so rapidly that *total costs* are not decreasing at all; the contradiction being hidden by a separation of "productive" from "marketing" costs.

This solution of the problem cannot, however, be dismissed as mere word-jugglery. It does show the actual working of the forces involved, and it is only the terminology that is unfortunate. What we have here is not perfect competition but *monopolistic* or *imperfect* competition. Chamberlain and Robinson have developed a more satisfactory line of attack on these problems, but how are we to find the falling demand curve which will entitle us to put these cases into this category and enable us to deal with them in the same way?

In order to obtain this it is essential to separate productive from marketing costs. The marketing costs involved in selling a given quantity of product must be subtracted from the gross receipts, just as if they were all direct or indirect reductions in price, leaving a definite total and average net receipts. For each quantity produced different prices may be charged and different marketing costs incurred. For each output some combination of prices charged and marketing costs incurred will leave a maximum average (and total) residue after subtracting the average (or total) marketing costs, and this maximum is the relevant Average Net Receipts for that output. The locus of such points will be the Average Net Receipts curve for the firm, and this is the "demand curve" which we need. This average net receipts curve and the corresponding marginal net receipts curve have to be used in conjunction with the "productive" cost curves which we may call "net" cost curves.

If the average net receipts curve is negatively inclined, one proceeds just as in the simple analysis of imperfect competition where there are no selling costs. The firm equates its marginal net cost to its marginal net receipts, and the degree of monopoly is equal to average net receipts over average net receipts minus marginal net costs, and the divergence of the position from the social optimum is illustrated by the fact that production is not carried on at the minimum average cost, but the firm produces less than this optimum output, stopping at a point where the average net cost curve is tangential to the average net receipts curve. The social loss, if any, due to the expenditure of resources on advertising is *not* taken into account in the measurement of monopoly. The measure will be the same whether the marketing costs are large or small, and whether they are given to the consumer in forms corresponding to cash, or whether they have important influences on his tastes for good or for bad. The social effects of different kinds of advertising constitute a quite separate problem.

If the average net receipts curve is horizontal where the marginal net costs curve cuts it, there is no monopoly. The existence of marketing costs is quite another matter.

<sup>1</sup> As by R. Harrod in his article on "The Law of Decreasing Costs," *Economic Journal*, Dec. 1931.

But there is no reason why the average net receipts curve should not slope upwards! It may well be that a larger quantity can be sold at a higher price at the same or a smaller *average* cost of marketing, and there is no ground for considering such a combination of circumstances as in any way exceptional. We must apply the same analysis here and not be deterred if the results at first appear a little strange.

If the firm with a rising average net revenue curve has a constant cost curve, or can acquire more of the product from other firms without affecting its marketing possibilities, we have another form of the paradox of the incompatibility of equilibrium with a horizontal demand curve and a falling average cost curve below it. The marginal revenue and the marginal cost curves cannot meet until the conditions are changed. Either the receipts curves must begin to fall or the cost curves must rise.

The interesting case—the one which can remain in equilibrium in these conditions—is the case where the average costs of the firm rise after a time as

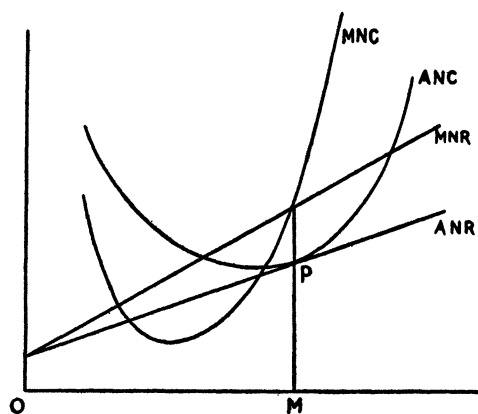


FIGURE 4

output increases, and where it cannot obtain more from other firms at the same price, either because the other firms' costs rise or because to do this would interfere with the reputation of the firm and upset its marketing possibilities.

This is shown in Fig. 4, where the firm is in equilibrium producing an output  $OM$ .

Average net receipts ( $ANR$ ) are equal to average net costs ( $ANC$ ), and marginal net receipts ( $MNR$ ) are equal to marginal net costs ( $MNC$ ). The degree of monopoly is here *negative* since marginal cost is

greater than average receipts. This may appear surprising, but it merely means that the divergence from the social optimum is in the direction opposite to that usually brought about by monopolies. Instead of the firm producing *less* than it should, it is producing *more*; the same kind of social harm is done, and it is reflected in the same way by the excess of the average cost over the minimum.

In finding an average degree of monopoly in an "industry," positive and negative monopolies may cancel out in whole or in part. Does this harm our apparatus?

I do not think it does this at all. It rather brings out the true nature of our measure as an index of *divergence* from an optimum. In any group of firms taken together to make an "industry," divergences may, and should, be expected to some extent to cancel out. For we are now considering the application of resources to this "industry" as against the rest of the economy. If of two firms within the "industry," one is producing too much and the other too little from the point of view of the economy as a whole; the industry may

not be producing either too much or too little. The maladjustment becomes a local affair which we must neglect in this larger consideration.

When our "industry" becomes the whole society, there cannot be too much or too little resources used, and as we have seen above, all the individual positive and negative monopolies must cancel out. This does not mean that society as a whole must always be in an optimum position, nor does it take any meaning away from the concept. It only means that the larger the fraction of the whole society one wishes to examine, the less legitimate is it to use particular analysis. In applying the particular mechanism to the whole economy we get the appropriate *reductio ad absurdum*. What is relevant for general analysis is not the *sum* of individual degrees of monopoly but their *deviations*. The standard deviations as suggested above may perhaps be used one day to give an estimate of the divergence of society from the social optimum of production relative to a given distribution of income.

A. P. LERNER.

*London.*