Optimal Departures From Marginal Cost Pricing

By William J. Baumol and David F. Bradford*

The need for this paper is a paradox in itself and indeed it might be subtitled: The Purloined Proposition or The Mystery of the Mislaid Maxim. For the results which it describes have appeared many times in the literature and have been reported by most eminent economists in very prominent journals. Yet these results may well come as a surprise to many readers who will consider them to be at variance with ideas which they have long accepted.

The proposition in question asserts that, generally, prices which deviate in a systematic manner from marginal costs will be required for an optimal allocation of resources, even in the absence of externalities. The reason for the difficulty into which marginal cost pricing is likely to fall is rather well known. What is not widely recognized is that there exists a highly sophisticated and well-developed body of literature indicating what should be done in such circumstances.

To see how the problem arises, consider an economy in which all industry has been nationalized and in which the central planning agency is dedicated to the maximization of social welfare. Suppose, accordingly, that it is decided to set all prices equal to marginal costs and to make up any deficits by subsidy out of the governmental treasury. If these funds are derived by excise taxes this is obviously a decision to make some prices depart from marginal costs after all. Or if it is obtained by an income tax, it is the price of labor which is forced away from its marginal cost. Any tax, except a Pigouvian poll tax—which might perhaps more felicitously be called an "inescapable tax"—will unavoidably affect some price. There is no way out of it. Any level of tax revenue which the government is determined to collect, whether as a means to make up a deficit resulting from a marginal cost pricing arrangement or for any other purpose, must in practice produce some price distortion.

Once this difficulty is recognized it becomes clear that one is dealing with a problem in the area of the second best. We are now faced with a problem involving maximization in the presence of an added constraint. Resource allocation is to be optimal under the constraint that governmental revenues suffice to make up for the deficits (surpluses)1 of the individual firms that constitute the economy.

* The authors are members of the department of economics, Princeton University. We are grateful to the National Science Foundation whose grant greatly facilitated completion of this paper. We must also express deep appreciation to our colleagues, Charles Berry, Stanley Black, William Branson, and W. Arthur Lewis, and to Ralph Turvey, whose comments and suggestions helped enormously in the development of its ideas.

1 Presumably, as a practical matter, there is an asymmetry between the problems posed by surpluses and those resulting from deficits. The real difficulty arises when we try to collect resources to cover the deficits without distorting consumers' and producers' choices. Since it is not possible in practice to levy lump sum taxes whose magnitude is independent of the decisions of those who pay them, we are forced to consider second-best solutions to the tax problem. A surplus can presumably be distributed more easily in a "lump sum" manner. One can, for example, distribute shares in future surpluses to all members of the economy on the basis of their incomes at some point in the past, or one can simply divide them equally among all individuals.

It should be emphasized also that the deficit problem need not arise only in the case of decreasing costs where
The theorems which have been developed to deal with this issue derive added interest from the fact that they are (virtually) the only concrete prescriptions for any second-best problem which have so far appeared in the literature. The discussion of this paper differs from the earlier writings in several ways. First, it deals with an important case not explicitly covered in most of the other papers in the area: quasi-optimal pricing when commodity prices throughout the economy are all adjusted as far as is possible. Second, it attempts a simplified exposition which is, of course, possible only at a cost in terms of loss of generality. Third, it brings together, explicitly, all three strands of the discussion: the welfare theoretic, the regulatory, and the public finance contributions. Finally, as far as we know, it offers the first overview of the extensive literature that has grown up in the area.

I. Sources of the Analysis

The preceding formulation of the problem suggests why the results described in this paper are not widely recognized. For they arise not out of the literature of welfare economics but from two more specialized fields—the theory of taxation and the analysis of public utility regulation.

The connection of the analysis with the theory of taxation is easy to understand. For the investigation, as it has just been described, is tantamount to a study of a system of optimal excise and income taxes. Only a small modification in our statement is needed for this purpose: instead of requiring the governmental revenue to equal the (algebraic) sum of the deficits of the firms constituting the economy, the revenue constraint can be generalized to require the tax system to bring in whatever revenues the government has decided it needs. The taxes which are optimal subject to this more general revenue requirement constraint then can be shown to follow the same necessary condition as that which replaces the marginal cost pricing rule in our previous discussion.

Similarly, there is a direct line between the more general welfare analysis and the theory of public utility regulation. For one reason or another, various public or regulated private monopolistic or partially monopolistic enterprises are operated under what amounts to a fixed profit constraint. Examples are turnpike authorities (required to cover costs and pay back invested capital), and regulated electricity, water, and telephone systems. The enterprises cited as examples are all firms that market a number of commodities. The turnpike sells easy driving between many different pairs of points, and, perhaps more important, at various times of day. Similarly, the telephone system sells ordinary telephone services for various uses over various distances and at various times, and, in addition, sells private wire, tele-type, television signal transmission, and other services related to, but not identical with, telephone services as usually conceived. In each example the demands for the different services are very obviously strongly interrelated, and they are obviously produced under conditions involving common costs.

If we assume that the net revenue allowed these enterprises is less than the amount profit maximization would yield, some degree of freedom is introduced into the pricing of the commodities produced. Of course, some amount of freedom is also introduced into the choice of method of production. We assume throughout that good will, pride of service, patriotism or the shrewdness of regulators assure that whatever output combination is chosen is produced at minimum cost.

A wide variety of output combinations and
sets of prices can then be chosen, each of which would just meet the net revenue constraint. The problem, then, is to determine which of these is optimal (second best) from the point of view of the use of resources to serve consumer desires. A similar problem obviously arises in a nationalized industry which is required to earn enough to cover its costs.

II. The Nature of the Theorem

In this section we shall describe the rules indicating whether a particular price-output combination that satisfies the profit constraint is socially optimal, i.e., whether from the point of view of the economy it yields the most effective allocation of resources permitted by the constraint. A solution to this constrained maximum problem will be called "quasi-optimal," because it is a second-best solution forced upon us by the revenue requirement.

A problem such as ours is usually framed in terms of the determination of appropriate output levels, but it can also be treated in terms of the choice of a price set for the outputs. Given the relevant demand functions, the choice of prices is tantamount to the choice of (salable) output levels. The two problems are, of course, formally identical, but, as we shall show, the pricing decision approach avoids a number of difficulties and yields a surprisingly simple optimality rule. In most conventional and familiar terms this rule asserts that Pareto optimal utilization of resources in the presence of an absolute profit constraint requires (considering substitution effects alone) that all outputs be reduced by the same proportion from the quantities that would be demanded at prices equal to the corresponding marginal costs. The rule takes an even simpler form in the event cross elasticities of demand are zero. It then requires that each price be set so that its percentage deviation from marginal cost is inversely proportionate to the item’s price elasticity of demand. According to this result, the social welfare will be served most effectively not by setting prices equal or even proportionate to marginal costs, but by causing unequal deviations in which items with elastic demands are priced at levels close to their marginal costs. The prices of items whose demands are inelastic diverge from their marginal costs by relatively wider margins.

This result is surely not immediately acceptable through intuition. It strikes us as curious, if for no other reason, because it seems to say that ordinary price discrimination might well set relative prices at least roughly in the manner required for maximal social welfare in the presence of a profit constraint. Since the objective of the analysis can be described as the determination of the optimally discriminatory set of prices needed to obtain the required profit, some degree of resemblance is perhaps to be expected. The case studied here is, thus, in a sense the obverse of the problem of profit maximizing price discrimination, and while the two solutions bear some qualitative resemblance, it can be shown that they may in fact differ substantially in quantity.

The theorem can be reformulated and generalized in a number of ways. Instead of utilizing the producers’ and consumers’ surplus concepts with all their theoretical limitations, the analysis can be framed in terms of the Hicksian compensating variation. Or a Pareto optimality approach can be utilized, both of these procedures obviating any need for interpersonal utility
comparisons. Similarly, the theorem has been extended to cover cases involving nonzero cross-elasticities of demand, to deal with input prices and the prices of intermediate goods, etc. Of course, once this is done the result loses some of its simplicity and the preceding statement requires considerable modification. But even then, surprisingly simple versions remain possible, as we shall see.

III. *The Formal Theorem*

In previous discussions, the basic theorem has been stated in a number of different ways. The objective itself has been described alternatively as: a) maximization of the sum of consumers' and producers' surpluses; b) determination of a set of prices from which it is *not* possible to change in a way that permits the gainers to compensate the losers; and c) maximization of the level of satisfaction of any one individual, given the utility level of each other individual (Pareto quasi-optimality). Each of these maximizations is, of course, constrained by the revenue requirement.

We will discuss the following four variants of the theorem, each of which gives a set of necessary conditions for quasi-optimal pricing:

1) If prices are quasi-optimal the ratio between the marginal profit yields of unit changes in the *prices* of any two goods will be equal to the ratio between their output levels.

2) For each product, the deviation of the quasi-optimal price from marginal cost must be proportionate to the difference between the product's marginal cost and marginal revenue (i.e., its marginal social welfare cost must be proportionate to its marginal contribution to the profit requirement). This result holds only in the case where cross-elasticities are all zero.

3) For each product, the percentage deviation of quasi-optimal price from marginal cost must be inversely propor-

tionate to its price elasticity of demand. This result also holds only where cross-elasticities of demand are zero.

4) Quasi-optimal prices must yield outputs that deviate by (approximately) the same proportion from those which would result from pricing at the marginal costs corresponding to the quasi-optimal output levels. This last form of the proposition, which is the variant most frequently encountered in the literature, is more general than the second and third versions of the theorem.

In demonstrating these propositions we shall utilize a comparatively straightforward manner of proof which has not previously appeared in the literature. To facilitate the exposition we start with a partial equilibrium approach. That is, we demonstrate that socially optimal pricing by a multi-product monopolist operating under a profit constraint is described by the preceding propositions. The monopolist is assumed able to set the prices of his final good outputs, but to purchase inputs at prices which remain fixed throughout the analysis. It will be clear that his optimal policy is the same as that which would be induced by the imposition of excise taxes equal to the derived divergence between output price and the marginal costs which would obtain if the outputs were to be produced by perfect competitors. This establishes the correspondence between the monopoly regulation and the excise taxation interpretations of the theorems.

If the monopolist is taken to be very large (e.g., if substantially all of the economy is operated by the government), or correspondingly, if the commodities taxed account for a large share of total output, the assumption of fixed input

---

4 To simplify the exposition we have dispensed with some theoretical niceties in the proof presented here, especially in the definition of the objective function. These are discussed and the objective function is examined in greater detail in our forthcoming paper.
prices must clearly be relaxed. This has been recognized and dealt with in the literature (for a simple statement see A. C. Pigou (1947, pp. 105-09); a sophisticated treatment is found in M. Boiteux 1956). We do not go into detail on this problem, but we do show in a subsequent section that the same line of reasoning used in our proof can readily be extended to a general equilibrium model.

The demonstration proceeds as follows:

Let \( p_1, \ldots, p_n \) and \( x_1, \ldots, x_n \) be the prices and outputs of the \( n \) commodities produced by the monopolist and, as indicated above, take as fixed the prices of his inputs. Let \( Z(p_1, \ldots, p_n) \) be our (unspecified) measure of consumer benefit, which is to be maximized subject to the profit constraint \( \Pi(p_1, \ldots, p_n) = M \) where the profit function has now been expressed in terms of prices rather than outputs. As usual in such constrained maximization problems, \( Z \) will be maximal subject to this constraint only if

\[
\frac{\partial Z}{\partial p_i} = \lambda \frac{\partial \Pi}{\partial p_i}, \quad (i = 1, 2, \ldots, n)
\]

That is the intuitively obvious first-order requirement that the marginal welfare gain from a given price change (price reduction) must be proportionate to its marginal profit cost.

---

5 If the revenues permitted by regulation are based on total investment rather than being fixed in absolute amount, this constraint may require some modification. If (as is at least ostensibly the goal of regulation) gross profits are required to be no greater than the cost of capital, then net profits (over the cost of capital) will be zero and the constraint will hold as stated with \( M = 0 \). If, however, the profit rate is permitted to exceed the cost of capital and the firm's total capital varies with output levels we may end up with \( M > 0 \) where the total profit permitted by regulation, itself varying with output levels. Our formal results remain unchanged if in the profit constraint \( \Pi = R - C = M \) (where \( R = \) total revenue and \( C = \) total cost) we use \( C^* \) to represent \( M + C \) and rewrite the constraint as \( \Pi = R - C^* = 0 \). We can then express all of our results in terms of \( C^* \) instead of \( C \). Economic interpretations are, however, clouded by this modification.

We can make the preceding equation much more explicit by examining more specifically the nature of the consumer benefit function \( Z(p_1, \ldots, p_n) \). While we will not be able to say much about that function itself, we can be rather specific about its partial derivatives, which is all we need for our present purposes.

We need a basis on which to estimate the gain to consumers from price-quantity data. A one dollar reduction in the price of a commodity will enable a consumer who is presently consuming \( x \) units of that commodity to continue buying exactly his original bundle of goods with a budget that is smaller by \( x \) dollars than his original budget. Thus \( x \) dollars is a lower limit to the relevant compensating variation since this is the least he would be willing to pay for that price change, and he would probably pay more since the \( x \) dollar reduction in his budget accompanied by the one dollar reduction in the specified price would certainly permit him to buy his original bundle and might perhaps enable him to buy one which he prefers but which he could not previously afford. Similarly, if an individual is currently purchasing the respective amounts \( x_1, \ldots, x_n \) of our \( n \) commodities at prices \( p_1, \ldots, p_n \), the amount he would pay to confront instead prices \( p_1 + \Delta p_1, p_2, \ldots, p_n \) will be at least \( -x_1 \Delta p_1 \). John Hicks (p. 330 ff.) shows rigorously in this same general manner that the rate of consumer gain per dollar increase in the price of any commodity is precisely the negative of the amount of the commodity initially consumed, times one dollar. We may then take as our expression for the derivative of the benefit function the compensating variation corresponding to the price change

\[
\frac{\partial Z}{\partial p_i} = -x_i
\]

Hence, substitution into the previous equation (1) yields

\[
-x_i = \lambda \frac{\partial \Pi}{\partial p_i}
\]
or, dividing by the corresponding condition for any other item \( j \), we obtain

\[
\frac{1}{x_i} \frac{\partial \Pi}{\partial p_i} = \frac{1}{x_j} \frac{\partial \Pi}{\partial p_j}
\]

as the required condition for quasi-optimality.

Equation (2a), then, is the first of the four variants of the theorem described at the beginning of this section. This necessary condition for quasi-optimality requires marginal profit yields of price changes to be proportionate to output levels. It may perhaps claim the virtue that it is relatively operational. If the firm has some notion of the likely marginal profit yield of a change in any one of its product prices, management can readily check whether (2a) is approximately satisfied.

We can now proceed directly from (2) to a derivation of the other variants of the basic theorem. We use \( MR_i \), \( MC_i \), and \( E_i \) to represent, respectively, the marginal revenue, marginal cost, and price elasticity of demand of output \( i \), and for simplicity assume that all cross-elasticities of demand are zero, though this is not necessary for the more general analysis, and is certainly not required for the preceding variant of the theorem.

Since

\[
MR_i = p_i + x_i \frac{\partial p_i}{\partial x_i}
\]

then

\[
\frac{\partial \Pi}{\partial p_i} = (MR_i - MC_i) \frac{dx_i}{dp_i}
\]

\[
= (p_i + x_i \frac{\partial p_i}{\partial x_i} - MC_i) \frac{dx_i}{dp_i}
\]

Substituting this into our basic condition (2) we have

\[-x_i \frac{dp_i}{dx_i} = \lambda(p_i + x_i \frac{dp_i}{dx_i} - MC_i)\]

or adding \((p_i + x_i \frac{dp_i}{dx_i} - MC_i)\) to both sides we obtain

\[p_i - MC_i = (1 + \lambda)(p_i + x_i \frac{dp_i}{dx_i} - MC_i)\]

\[= (1 + \lambda)(MR_i - MC_i)\]

This is the second form of the quasi-optimality theorem: the difference between price and marginal cost\(^7\) should be proportionate to the difference between marginal revenue and marginal cost.

Next we may rewrite the preceding equation as

\[-\lambda(p_i - MC_i) = (1 + \lambda)x_i \frac{dp_i}{dx_i}\]

or

\[\frac{p_i - MC_i}{p_i} = \frac{1 + \lambda}{\lambda} \frac{1}{E_i}\]

since by the usual formula, \( E_i = -\frac{(p_i/x_i)}{(dx_i/dp_i)} \) which is the third form of the result.\(^8\) Note that it implies that if all elasticities in question are equal, prices

\[^{7}\text{Some care must be exercised in interpreting the various forms of the theorem described here because marginal costs are not generally constant. Let } MC_i(q) \text{ represent marginal costs corresponding to the quasi-optimal output levels while } MC_i(M) \text{ represent the marginal costs corresponding to the equilibrium outputs when all prices are set equal to marginal costs. Then in our proposition we have } \Delta p_i = p_i - MC_i(q) \text{ and they are thus not necessarily equal to the } p_i - MC_i(M). \text{ That is, the propositions do not contrast quasi-optimal prices with the prices corresponding to a marginal cost pricing equilibrium as ordinarily conceived. They refer instead to the differences between the quasi-optimal prices and the marginal costs corresponding to the quasi-optimal output levels.}\]

\[^{8}\text{It should be reemphasized that this form of the derivation deliberately ignores the supply side for the sake of expository simplicity. However, the analysis is easily extended to take supply into account, and this was done from the very beginning of the discussion, by Frank Ramsey and by A. C. Pigou in 1947. For example, Pigou states "Writing } \eta \text{ (defined as negative) for the elasticity of demand in respect of pre-tax output, } \epsilon \text{ for the corresponding elasticity of supply of the } r \text{th commodity and } t_r \text{ for the ad valorem rate of tax on } i, \text{ it can be proved that, in the conditions supposed (the optimum system of such taxes yielding a given revenue require) rates of tax . . . such that } t_r/(1/\epsilon - 1/\eta) \text{ has the same value for all values of } r^9 \text{ (pp. 107-08).}\]
should be set proportional to marginal costs.

For the final form of the proposition, write the deviation of price from marginal cost as \( \Delta p_i \), i.e., set \( p_i - MC_i = \Delta p_i \). Then equation (4) can be written after a minor rearrangement of terms as

\[
\frac{dx_i}{d \Delta p_i} = -\frac{1 + \lambda}{\lambda} x_i
\]

But the left-hand side of the preceding equation may be interpreted as an approximation to the change, \( \Delta x_i \), in the \( i \)th commodity demanded which would result from a shift from the actual prices to current marginal cost levels. We thus obtain the last version of our theorem:

\[
\Delta x_i = k x_i, \quad \text{where} \quad k = (1 + \lambda)/\lambda.
\]

The more sophisticated discussions have generally emphasized this last form of the theorem, the assertion that quasi-optimal pricing requires (after compensation for income effects) a proportionate change in all purchases from the levels that would be observed if prices were set at marginal costs.\(^9\) This interpretation has usually been preferred to the two variants that precede it because, as has been shown by M. Boiteux (1956), Paul Samuelson (1951), and others, this form of the theorem holds quite generally and does not depend on the simplifying assumptions needed to arrive at the two special forms of the theorem described before this last variant. Only our first relationship (2a), which requires the

\(^9\) In accord with the comments of fn. 6, this third form of the theorem should be read as follows: a change in prices from their quasi-optimal levels to the corresponding marginal cost levels should reduce all outputs in (approximately) the same proportions.

If we estimate demand changes on the assumption that the slope of each demand curve is constant over the relevant range, our proposition is valid for large as well as for small price changes. It now asserts that the divergence between price and marginal cost must be such that the estimated percentage changes in demand resulting from a drop in all prices to the given marginal cost levels are the same for all commodities.

marginal profit yields of commodity price changes to be proportionate to their output levels, is of the same order of generality.\(^{10}\)

IV. Intuitive Rationale

Though at first blush the rationale underlying the preceding propositions may not be entirely transparent, it is not too difficult to account for them intuitively. The last form of the theorem is perhaps the most helpful in offering us a grasp of the entire matter. We need merely think of the consequence of a deviation of price from marginal cost as a distortion of relative demand patterns. This immediately suggests the last form of the theorem for it implies that the damage to welfare resulting from departures from marginal cost pricing will be minimized if the relative quantities of the various goods sold are kept unchanged from their marginal cost pricing proportions. If we accept this plausibility argument, then the elasticity form of the theorem clearly follows. If we propose to (say) contract the demands for \( A \) and \( B \) each by precisely \( k \) percent, then if \( A \)'s demand is much more elastic than \( B \)’s, clearly \( A \)'s price must be raised by a substantially smaller percentage than that of \( B \).

A bit more light is shed on the matter by a simple graphic discussion that follows rather loosely an analysis first presented by Ursula Hicks (p. 167 ff) and developed

\(^{10}\) Note that application of the theorem in any of its forms does not rule out a two-part or a multi-part tariff. On the contrary, it tells us how to determine the quasi-optimal value for each of the multiplicity of prices that composes such a tariff. For example, when the consumer pays one charge for the rental of a telephone and another for each of his calls, this can be interpreted either as a two-part tariff or as a pair of prices for two distinct but related services each of which has a quasi-optimal price. This discussion is, of course, based on the observation that one cannot in reality impose a Pigouvian poll tax as one part of the multi-part tariff. Indeed, if such a lump sum tax were possible there clearly would be no need for the theorems under discussion. Marginal cost prices plus lump sum taxes could then satisfy the revenue requirement and yield an optimal allocation.
fully by William Vickrey in 1968. In Figure 1 consider the segments of two demand curves $D_aK$ and $D_bK$ through point $K$. Let us use the Marshallian measure of consumers' surplus to compare the psychic loss to consumers from a given price (tax) rise with the resulting increase in (tax) revenues. Then, if demand is in fact given by the less elastic curve $D_aK$ and price rises from $P_o$ to $P_1$, the loss of consumers' surplus will be $P_oK_aE_aP_1$. On the other hand, assuming for simplicity that marginal costs are constant at level $OC$, net revenue will be increased as a result of the price change by the quantity $P_oK_bE_bP_1$ minus $R_oRKK_b$. Thus the positive portion of the net revenue change, $P_oK_bE_bP_1$, may be considered to be offset by an approximately equivalent loss to consumers. This means that with the less elastic demand curve the rise in price will have caused a net reduction in net revenue plus consumers' surplus that is measured by the shaded area, $R_oRKE_b$.

Similarly, with the more elastic demand curve, $D_bK$, the rise in price from $P_o$ to $P_1$ will decrease the sum of the revenue net of consumers' surplus by the greater area $R_oRKE_a$. In general, then (with marginal costs constant), a given price rise will exact a larger social cost in consumers' surplus not offset by increased revenues the more elastic the demand curve. That is essentially the reason for the theorem that calls for a relatively small deviation between the price and marginal cost of any commodity whose demand is comparatively elastic.

V. A Simple General Equilibrium Model

Our interest in conveying the central results of the literature discussed in our paper to a wide readership in an intuitively convincing fashion leads us to couch our argument of Section IV in terms of partial equilibrium analysis. It is desirable, however, to sketch out the way it all works in general equilibrium terms. A simple re-interpretation of the terms used in the derivation of the theorems will show that our proof is sufficient to establish the propositions in the special general equilibrium case in which there is only one input factor, labor, and all the productive activities are in the hands of the government. Accordingly, let $p_1, \ldots, p_n$, be the prices in terms of labor hours at which the $n$ final outputs other than leisure, $x_1, \ldots, x_n$, all produced by the government, are sold to the consumer-laborers. The government is assumed to operate with a profit constraint in the form of a fixed number, $M$ (which may be zero), of units of labor. In addition, of course, there is a production constraint, which we may write as

\[ \text{Notice that in putting this constraint in the form of an equality, we have ruled out production inefficiency. Although this is a plausible requirement and makes things easier, it is not always innocuous when more is meant by "consumer benefit" than attainment of Pareto optimality—the concept utilized here. See} \]
where $C$ is the minimum labor input sufficient to produce the output vector, $x_1, \ldots, x_n$. The problem is to maximize "consumer benefit" subject to this production relation and the profit constraint,

$$\pi = p_1x_1 + \ldots + p_nx_n - C = M$$

Now let us treat the government as an ordinary firm that imagines it can buy all the labor it wants at a price of one per unit. The additional labor it would have to buy to produce an additional unit of good is then the marginal cost of that good, both in the accounting sense and in the sense of its leisure opportunity cost, all other outputs held constant.

There will be, in general, a large set of price vectors compatible with both (a) clearing of the markets for goods and factors and (b) satisfaction of the government's profit constraint. It is useful to spell this statement out in somewhat more detail. If the aggregate labor sold is $L$, we can write the set of demand equations as

$$L = L(p_1, \ldots, p_n)$$

$$x_1 = x_1(p_1, \ldots, p_n)$$

$$\ldots$$

$$x_n = x_n(p_1, \ldots, p_n),$$

where, necessarily (Walras' Law),

$$p_1x_1 + \ldots + p_nx_n = L$$

For a set of prices to be market-clearing, the government must buy exactly the specified amount of labor and produce exactly the specified quantities of goods. By (5) and (6) "feasible price vectors" are those which satisfy the profit constraint conditions

$$L - F(x_1, \ldots, x_n) = M,$$

or, equivalently,

$$p_1x_1 + \ldots + p_nx_n - F(x_1, \ldots, x_n) = M$$

The task is to find among these an optimal price vector, where we have taken this to mean a vector which consumers in the aggregate would be unwilling (after costless negotiation among themselves) to bribe the government to change. The solution to this problem is found in the propositions set forth in Section III above. The entire analysis in that section continues to hold when the profit constraint as just described is used in place of the usual partial equilibrium construct.

**VI. The Taxed and the Untaxed Sectors**

Now we are equipped with a genuine, if very simple, general equilibrium apparatus, but one which in most respects can be manipulated in partial equilibrium terms. The key to this is the form in which we have expressed the production relation (as a cost function), the fixing of the government's profit constraint in terms of labor, and the corresponding choice of labor as price numeraire (so that the price of the single input factor, the "all other prices" of partial equilibrium analysis, is fixed at unity). In this model the government can be taken to tax the output of producers of the $n$ non-leisure goods (where the tax is the difference between price and marginal cost), and hence these goods make up the "taxed sector." The untaxed sector here may reasonably be described as the production of negative labor input to the taxed goods-production sector, i.e., the production of leisure. In a rather trivial sense, in the latter sector marginal cost and price are always equal at unity.

It is illuminating at this point to reconsider "our" results in light of the following comment by Abba Lerner, who derives two special rules for quasi-optimality, each applicable to a special case:

[One rule [equalize tax elasticities] is
appropriate where the shifting is only from a taxed to an untaxed sector, while the [other] rule [maintain proportionality of price to marginal cost] is appropriate where the shifting is only to another part of the taxed sector (as would be the case if all uses of resources could be, and were, taxed).

What we need is a rule that is applicable whether the resources shifted by a tax remain in the taxed sector or go to the untaxed sector or are divided between the two sectors. [p. 285]

Documentation of the fact that a general rule for optimal excise taxation has long been available in the literature, and presentation of a simplified derivation of several forms of such a rule are precisely our objectives. It is easy to show that in those situations where Lerner recommends prices in proportion to marginal cost, our theorems do likewise; where Lerner recommends equalization of tax elasticities, so do our theorems; and where Lerner calls for something in between the two, our theorems provide an exact prescription.

In the case of the model which has just been described, the statement that a tax change has as its only effect a shifting of resources from taxed to untaxed sectors is clearly equivalent to the assertion that the only output affected in the taxed sector is the one whose tax rate has changed. If the effect of varying any one tax is only to shift resources into the non-taxed sector then only own elasticities of the taxed goods, and not their cross-elasticities with one another, can be non-zero. In this case, Lerner and we (most directly in variant 3 of the theorem) conclude that the percentage deviation of price of any taxed commodity from marginal cost should be inversely proportional to its own price elasticity of demand.

We turn now to the next case where the alternative use of resources induced by tax changes can be said to be entirely in the taxed sector. This is obviously equivalent to the assertion that labor is perfectly inelastically supplied over a wide range of price vectors. In this situation, Lerner and our theorems agree that prices proportional to marginal costs, i.e., relative prices equal to marginal rates of transformation between (non-leisure) commodities, is clearly optimal. This result follows easily for the case when cross-elasticities between (non-leisure) goods are all zero. Then, since a change in the price of a given good results in no change in either the purchases of other goods or of labor sold, it must lead to no change at all in labor expenditure on the good in question. That is, all goods must be demanded with equal, in fact, with unit elasticity. Our third form of the theorem states that in this case the relative deviation of price from marginal cost should be the same for all commodities.

More generally, admitting non-zero cross elasticities, the zero elasticity of labor supply means that any change in profit resulting from a price change comes purely from the change in labor cost of goods sold, since there can be no change in the

---

12 Two things should be noted. First, when Lerner speaks of "tax elasticity" he is referring to what might better be called "own tax elasticity," that is, the relative change in the tax revenue (here, profit) collected on commodity \( x_i \) when the tax on \( x_i \) (divergence between price and marginal cost of \( x_i \)) is raised by one unit. The change in tax revenue collected on other commodities, which arises via cross-elasticity of demand, is not included.

Second, it should be observed that although the term "resources" is often not definable in general equilibrium models of the type described above (since choice is made from among alternative vectors of social aggregates of final goods, without reference to the underlying mechanics of production), in this simple case we may presumably equate them with "labor."

13 Note that unless this inelasticity of supply results from some kind of artificial constraint, such as a legal maximum working day, or from a kink in the individuals' utility functions, commodity taxes will in this case not yield the same result as lump sum taxes. For if the inelasticity results from the offsetting of income and substitution effects, a doubling of prices will generally have a different effect from the subtraction of half (or any other portion) of the individuals' original labor income, leaving prices unchanged. This is obviously so because the (alternative) new price lines will not have the same slope and so they cannot be tangent to the same indifference curve at the same point.
total revenue received by the firm or government: the fixed quantity of labor offered in exchange for their products. We have, for all price vectors in question, by the inelasticity of labor supply at quantity $L_0$:

$$\sum_j p_jx_j(p) = L_0;$$

(7)

i.e., $$\sum_j p_j \frac{\partial x_j}{\partial p_i} + x_i = 0$$

for all $i$. Furthermore, as just noted, price changes can affect profits only through their effect on costs:

$$\frac{\partial \pi}{\partial p_i} = \frac{\partial}{\partial p_i} \left( \sum_j p_jx_j - C \right)$$

(8)

$$= 0 - \sum_j \frac{\partial C}{\partial x_j} \frac{\partial x_j}{\partial p_i}$$

Recall now the first form of "our" theorem which asserts that optimality requires, for all $i$,

$$-x_i = \lambda \frac{\partial \pi}{\partial p_i}$$

In the case now under consideration, this becomes by (8)

$$-x_i = -\lambda \sum_j \frac{\partial C}{\partial x_j} \frac{\partial x_j}{\partial p_i}$$

But by (7)

$$-x_i = \sum_j p_j \frac{\partial x_j}{\partial p_i}$$

Together these results imply that the first-order conditions can be written for all $i$,

$$\sum_j \left( p_j + \lambda \frac{\partial C}{\partial x_j} \right) \frac{\partial x_j}{\partial p_i} = 0$$

Since $\partial C/\partial x_j$ is the marginal cost of $x_j$, this asserts that prices bearing the proportion $-\lambda$ to marginal cost satisfy the first-order conditions.

In the more general case where shifting takes place both between sectors ($\partial C/\partial p_i \neq 0$) and within the taxed sector ($\partial x_i/\partial p_j \neq 0$, at least sometimes), an analogous algebraic chain of steps leads to the following first-order conditions, for all $i$:

(9) $$\sum_j \left( p_j + \lambda \frac{\partial C}{\partial x_j} \right) \frac{\partial x_j}{\partial p_i} = (1 + \lambda) \frac{\partial L}{\partial p_i}$$

Only by chance will these conditions be satisfied by prices proportional to marginal cost, or by prices deviating from proportionality by a simple function of own elasticity. In this we agree with Lerner. However, the theorems prescribe definite prices which do optimize, given the constraints, and there is no necessity to resort to a rough compromise between two polar rules. One rule prescribes for the polar situations as well as for intermediate ones. This does not mean that a more precise spelling out of Lerner's compromise would not yield the same (correct) conclusions: it means only that the task which remains is not to find the general rule. That is already accomplished by (9). Rather, it is only necessary to translate this rule into a form compatible with Professor Lerner's very illuminating intuitions.

VII. The Case Where All Items are Taxable

Finally, we must yet deal with the case, excluded from our discussion thus far, in which "all uses of resources" can be taxed. In view of our preceding agreement with Lerner that when tax changes result in resource shifts only within the taxed sector, prices proportional to marginal costs are optimal, our assertion that this is generally not the case when all goods are taxable may come as a surprise. The explanation is, however, relatively simple: overall marginal cost pricing will not, as a rule, yield the necessary revenue.\textsuperscript{14} This

\textsuperscript{14} We shall say that "overall marginal cost pricing" prevails when the price vector confronting consumers is a scalar multiple of the vector of first partial derivatives of the transformation function evaluated at the associated social total net output vector.
follows from a basic assumption of the entire analysis: that lump sum taxes are impossible. In terms of our general equilibrium this requires (by Walras law) that the net value of consumers' sales and purchases of all commodities, including labor, add up to zero. In that case, a proportionate tax, leading to a proportionate change in prices, will still leave the net consumer expenditures at zero and hence will yield no revenue to the government.

In formal terms, let \( P_i \) be the price of commodity \( i \) \((i=1, \ldots, n)\) facing the consumer, no longer expressed in labor terms, and let \( P_L \) be the price of labor. Following the sign convention that positive numbers represent net purchases of desired goods, negative numbers, net sales, let \( a_1, \ldots, a_n, a_L \) be the vector of net transactions carried out by a consumer. Our no-lump sum tax assumption can be expressed as the requirement that \( P_1 a_1 + P_2 a_2 + \ldots + P_n a_n + P_L a_L = 0 \). (This implies, of course, that the aggregate net transactions vector, summed over all consumers, satisfies the same condition, expressed above as the condition on the demand functions \( p_1 x_1 + \ldots + p_n x_n = L \), where prices were expressed in labor terms.) Obviously, here it makes absolutely no difference to the consumer whether he faces prices \( P_1, \ldots, P_n, P_L \) or prices ten times as high; only relative prices count. If the government were to collect the difference between these two price vectors (\( P \) and 10\( P \)) as a tax, it would collect precisely \( 9(\sum_i P_i a_i + P_L a_L) \), which is zero. A tax vector which is a scalar multiple of the price vector facing consumers will, under our zero-transfer assumption, always yield zero revenue.

With consumer budgets equated to zero, the vector of taxes thus cannot be proportional to consumer prices if there is to be any tax yield, and hence consumer prices cannot be proportional to marginal costs. Now we are back where we started. If the government wishes to obtain, say, a quantity \( M \) of labor for its own uses, and can only work with commodity taxes, it must get consumers to settle at an equilibrium with an aggregate net transaction vector \( x_1, \ldots, x_n, -L \), satisfying \( L - F(x_1, \ldots, x_n) = M \), i.e., such that the leisure given up for production exceeds that necessary to produce the other desired goods by \( M \).

VIII. An Analysis in Terms of Quantity Changes

Some readers may find it instructive to see a derivation of the propositions carried out explicitly in a general equilibrium context and in terms of quantity changes rather than price changes. This is easily done within our present model if it is assumed that lump sum redistributions maintain a socially optimal distribution of alternative social totals of goods and leisure. Then we can employ a Samuelsonian social utility function relating social welfare to these totals, with the property that the tangent plane to any social indifference surface is also a "budget plane" associated with prices at which that point will be sustained as a competitive equilibrium. Thus prices can be written as a function of quantities. The analysis is made more transparent if we assume that there is a certain fixed total, \( \bar{T} \), of labor time which must be allocated among production, leisure, \( R \) (for recreation), and government surplus, yielding the constraint:

\[
F(x_1, \ldots, x_n) + R + M = \bar{T}
\]

where \( F(x_1, \ldots, x_n) \) is obviously the (labor) cost function for the output bundle \((x_1, \ldots, x_n)\).

The problem is to maximize over output-leisure bundles the social utility function, \( U(x_1, \ldots, x_n, R) \), subject to this constraint, and subject to the constraint that exactly the required government surplus is obtained as profit:
\[ \sum_{i=1}^{n} p_i x_i - F(x_1, \ldots, x_n) = M, \]

where the \( p \)'s are equilibrium labor-prices associated with the chosen output-leisure bundle. Associating the Lagrange multiplier \( \lambda \) with the first constraint, and the multiplier, \( \mu \), with the second, we can write out the Lagrangian expression:

\[
V(x_1, \ldots, x_n, R, \lambda, \mu) = U(x_1, \ldots, x_n, R) + \lambda [L - F(x_1, \ldots, x_n) - R - M] + \mu \left[ M - \sum_{i=1}^{n} p_i x_i + F(x_1, \ldots, x_n) \right]
\]

Taking partial derivatives we obtain the first-order conditions:

\[
\frac{\partial U}{\partial x_i} - \lambda \frac{\partial F}{\partial x_i} - \mu \left( \frac{\partial}{\partial x_i} \left( \sum_{j=1}^{n} p_j x_j \right) - \frac{\partial F}{\partial x_i} \right) = 0,
\]

\[ i = 1, \ldots, n \]

\[
\frac{\partial U}{\partial R} - \lambda = 0
\]

Dividing each of the first \( n \) conditions by the \((n+1)\)st and noting that equilibrium of the consumer requires \( \partial U/\partial x_i/\partial U/\partial R = p_i/p_r = \mu \) (since labor is the numeraire, its price unity) we have

\[
p_i - \frac{\partial F}{\partial x_i} = \frac{\mu}{\lambda} \left( \frac{\partial}{\partial x_i} \left( \sum_{j=1}^{n} p_j x_j \right) - \frac{\partial F}{\partial x_i} \right).
\]

Employing, with their usual meanings, the symbols \( MC \) and \( MR \) (the latter including the effect on revenue of variations in all prices), the preceding condition implies

\[
p_i - MC_i = \frac{\mu}{\lambda} (MR_i - MC_i),
\]

\[ i = 1, \ldots, n \]

which is the second form of our theorem from which the other forms can be obtained exactly as before. Thus we have derived the results explicitly from a general equilibrium model. Note that the term "marginal revenue" is definable here because we have normalized prices to labor dimensions and have assumed behind-the-scenes income distribution, making determinate the effect on equilibrium prices of changes in any of the desired outputs. Note also that under these conditions, the second version of the theorem, proved in Section III for the case of independent demands, is shown to hold generally.

IX. Notes on the History of the Discussion

There is some point in going briefly into the antecedents of the discussion because this history makes it all the more difficult to understand why the propositions in question have achieved so little recognition by the profession. The general line of argument has appeared widely for the better part of a century. The formal theorems themselves date back more than forty years. As we will see, this work has appeared in some of our leading journals under the authorship of some of the luminaries of our profession and was clearly not limited to a backwater of the literature.

There is an informal proposition very closely related to the results under examination that has a long history going back at least to the 1870's. With the establishment of the Railway and Canal Commission in England in 1873, and of the discussion preceding the establishment of the Interstate Commerce Commission in the subsequent decade there arose a rich literature examining utility pricing in relation to the public interest. A number of authors advocated prices that vary directly with demand inelasticity ("value of service").\textsuperscript{15}

\[ \text{\textsuperscript{15} For some examples see A. T. Hadley (ch. 6); E. Porter Alexander (esp. pp. 2-5, 10-11) and W. M. Acworth (ch. 3, esp. pp. 57-60). Hadley, the American} \]
The mainstay of the early discussion was an argument only loosely related to the formal propositions in which we are interested. In brief, it was maintained that a relatively low price in elastic markets, provided it covers more than incremental cost, may well permit lower prices in other markets and hence be beneficial to everyone. For (particularly if the firm is subject to a constraint on its overall profit) the opening of a market which makes any net contribution may permit or may even require a reduction in prices elsewhere.\textsuperscript{16}

Discussion along these lines has since appeared throughout the literature of public utility economics. (See, e.g., James Bonbright, esp. chs. 5, 17–20.) The various forms of the proposition have also appeared at least implicitly in a number of the standard writings of economic theory, notably in Joan Robinson’s work and that of W. Arthur Lewis.\textsuperscript{17}

However, the formal mathematical propositions which are derived from optimality considerations and which are the subject of this paper, themselves have a substantial history. As propositions on optimal taxation they first appear in 1927 in Frank Ramsey’s pathbreaking article on taxation.\textsuperscript{18} Thus, more than a decade before the publication of Hotelling’s historic discussion of marginal cost pricing, Ramsey had in another context provided at least implicitly a solution to the optimal pricing problem for an industry in which marginal cost prices do not cover total costs.

A. C. Pigou took up the Ramsey discussion in the following year in his book on public finance, providing a very lucid and rather extensive summary of the argument. It was independently approached by Ursula Hicks in her 1947 book on the same subject. Her analysis is largely diagrammatic and is based entirely on the Marshallian consumers’ surplus, just as Ramsey’s and Pigou’s had been. However, as far as we have been able to tell, the formal anal-

\textsuperscript{16}“If the (New York to San Francisco) price is more than the additional outlay involved in doing it, as against leaving it alone, it is profitable to the railroad, and the business is moreover advantageous to the whole inland community served by the railroad. For it adds to the number of men employed along the line and . . . the more prosperous the road, the lower the local rates may be made” E. P. Alexander (p. 4) (author’s italics). The author adds (p. 5) that between the limits of “unreasonable” profits and net loss, actual profit rates “should be adjusted in proportion to value of service rendered” (inelasticity of demand).

\textsuperscript{17}See Joan Robinson (p. 207). Since she is dealing with a single product and does not utilize any explicit revenue constraint, the result she gives is not quite what is needed for our purposes but it is closely related to the theorem under discussion.

Lewis’ statement (pp. 20–21), on the other hand, is precisely on target:

The principle . . . is that those who cannot escape must make the largest contribution to indivisible cost, and those to whom the commodity does not matter much . . . get off lightly. . . . When there are escapable indivisible expenses to be covered the case for discrimination is clear. It secures an output nearer the optimum and levies the indivisible cost on those who get the greatest benefit (measured by their consumers’ surplus). . . . Moreover, it is possible in some cases that the net result may be that everyone pays a lower price. . . . If the undertaking is not merely to cover its costs . . . reducing the price to some persons with elastic demands may increase the surplus over marginal cost which they contribute. . . . (emphasis added).

\textsuperscript{18}The theorem seems to have made its greatest impact on the literature of public finance. A. C. Pigou (1928, pp. 126–28); (1947, pp. 105–09) reports Ramsey’s results, offers an intuitive explanation and translates the theorem into elasticity terms. He also argues that except in the case of linear supply and demand functions the argument holds only if the taxes are small.

Ursula Hicks (pp. 167 ff) provides a very similar result on the basis of an elementary graphic argument. Yet even Richard Musgrave in his classic volume (p. 148) devotes only a footnote to the theorem and dismisses it as being “arrived at within the framework of the old welfare economics of interpersonal utility comparison.” His characterization of the argument as part of the ability-to-pay approach is surely rather misleading. Cf. R. Bishop’s comment on Musgrave, p. 212n.
ysis did not reappear until 1951 in an article by M. Boiteux which was the first
to dispense with the notion of consumers’
 surplus. Apparently in the same year, Paul
Samuelson submitted to the U.S. Treasury
a closely related paper containing a gen-
eralization of Ramsey’s results. Samuel-
son’s approach differs from that which we
are discussing in that he employs no ex-
 plicit revenue requirement. Instead he
requires some preselected quantity of each
output to be left unused by the private
sector and made available to the gov-
ernment. His result, however, is essentially
the same as those of the other contributors
under discussion.

In 1952 there appeared a particularly
simple derivation of the theorem in its
most elementary form in a note by Alan
Manne. His discussion assumes that
cross-elasticities of demand are all zero and it
measures and aggregates consumers’ and
producers’ surplus in an unabashed Mar-
shallian manner, but his derivation com-

penses for these simplifications by its
great lucidity. Boiteux and Manne were
also the first to present the analysis in the
form of a discussion of public utility pricing.
One year later Marcus Fleming pro-
vided an independent and illuminating
derivation.

Meanwhile, Gerard Debreu had con-
tributed analytic materials which led
Boiteux to employ a Pareto optimality
approach to the matter. On this basis
Boiteux was then able to complete his
definitive analysis of the subject. In
Boiteux’ classic 1956 article the analysis
is, as a result, independent of any inter-
personal comparison. The notions of con-
sumers’ and producers’ surplus are dis-
Pensed with altogether. He is able to deal
with goods whose demands are interdepen-
dent, with the pricing of inputs and inter-
mediate goods as well as with outputs. It
has also been emphasized by a subsequent
writer (Jacques Drèze) that by his general
equilibrium approach to the matter, Boi-
teux was contributing an explicit second-
best solution, one of the few to be found
to date in any area of welfare economics.
Of course, no piece of analysis can answer
every relevant question, and Boiteux left
unsettled some matters of detail. There is
a moderately mysterious role played in his
analysis by the numéraire commodity; he
has not dealt with problems of nonnega-
tivity of outputs, etc. All this has left room
for subsequent work by R. Rees, William
Vickrey (Testimony and other unpublished
work), the present authors, and by a num-
ber of others.20

20 Among those who have also written on the subject
are Diamond and Mirmles and Bishop. Recall in this
connection Samuelson’s reference to unpublished work
by Hicks and Hotelling, on which more light is shed by
the following excerpt from a letter which the authors
received from Professor Hicks:

It was just as well that you sent me the cutting
from Samuelson [the preceding footnote], or I
should have found it hard to recollect the matter.
But with this refreshment it comes back. I think
the story is as follows.
However, there is no need to go beyond Boiteux' work for a basis for the welfare discussion contained in this paper. The generalization to optimal pricing for an entire economy requires only one small modification in Boiteux' two-sector analysis. In his model there are two sets of firms, the first composed exclusively of perfect competitors, while the second are subject to the author's rules for quasi-optimal pricing. For our purpose we need merely take all firms to fall into the second category and our interpretation follows at once. In sum, it follows for the economy as a whole that unless marginal cost pricing happens to provide returns sufficient to meet the social (governmentally determined?) revenue requirement, a quasi-optimal allocation calls for systematic deviations of prices from marginal costs throughout the economy in the manner specified by the theorems that this paper has described.

X. Concluding Comment

We conclude our survey of the relevant literature as we began—at a loss to explain how a set of results flowing from the pens of so distinguished a set of authors and appearing and reappearing in the profession's leading journals, should so long have escaped general notice. This is all the more curious in light of the importance of the subject and the elegance of the theory.

As a theoretical matter, the theorem we have discussed seems fundamental not only for the analysis of public finance and the regulation of utilities, but also for some of the basic precepts of welfare economics. Earlier we spoke of the result as a proposition in the theory of the second best. But it is more than this. In a world in which marginal cost pricing without excise or income taxes is normally not feasible, the solution we have usually considered to characterize the “best” is none too good, because it is simply unattainable. In that case, the systematic deviations between prices and marginal costs that the theorem calls for may truly be optimal because they constitute the best we can do within the limitations imposed by normal economic circumstances.

Appendix

Comments On Lerner's Appendix

In December 1946 (in the last week of my first visit to the United States) I stayed for a couple of days with Hotelling at Chapel Hill. I then told him that I had worked out (using a quite crude consumers' surplus method) that if a given sum had to be raised in excise taxes, the least-sacrifice way of doing it was the proportional all-round reduction in consumption, as compared with an optimum position. For if the burden of the tax is

$$\frac{1}{2} \sum q_p d_p d_i$$

this is minimised subject to $G = \Sigma q_p d_p$ if

$$d_q = \sum q_p d_p = \lambda_i$$

where $\lambda$ is the Lagrange multiplier. Not much more than this, except that I had allowed (in the spirit of the old Hotelling article) that the $q_i$ could incorporate supply as well as demand reactions.

I remember he told me I ought to publish this, but I didn't—mainly, I suppose, because I was conscious of the qualifications to which Samuelson alludes in his paper, and which, if I had set them out in my style, would have whittled away the result so near to nothing.
interchangeably, we would use the two terms “consumers’ loss” and “social damage” interchangeably as well. We would use neither without a grain of salt. The touchstone of quasi-optimality is plain and simple Pareto optimality, given the profit or tax revenue constraint. As is by now generally recognized, Pareto optimality cannot be comfortably equated with a “higher ethical good,” and we apologize if we give the impression that a quasi-optimal price structure (for the monopolist) or tax structure was “good” in any sense other than Pareto optimality.

The second assertion that occurs in several places is that our analysis assumes constant marginal costs. This is simply incorrect. The only point at which marginal costs are taken to be constant is in our Section IV, the graphical presentation, which is intended solely as an aid to intuition. Elsewhere there is no assumption of any sort about marginal costs and none of the results depend on such an assumption. Lerner seems to suggest that our assumption that the prices of inputs are fixed can be equated to the premise that marginal cost is constant. We fail to see any connection between the two. Our assumption that input prices are constant is another, perhaps mistaken, compromise in the direction of intuitive appeal.21

Incidentally, the point of our footnote 7 is not, as Lerner suggests, to acknowledge in passing that marginal costs might in fact be variable. Rather, it is intended to point out that the marginal costs in question in the various forms of the theorem are not those which prevail in some initial situation before any taxes are imposed, but those which in fact rule at the solution values. Rather than apologizing for assuming marginal costs constant, we are, on the contrary, saying that marginal costs will very likely change as taxes are varied, so one must watch out and be sure to use the correct values in testing for optimality.

The basic task undertaken by Lerner in his Appendix relevant to the discussion here is derivation of our four forms of the “mislaid maxim” from his second rule. Since all four are derivable from either the first \( \frac{\partial \pi}{\partial p_i} = \lambda x_i, \) all \( i \) or the fourth \( \Sigma_i(\partial x_i/\partial p_i)(p_i - MC_i) = \gamma x_i, \) all \( i \) it is reasonable that he should be able to do this. It is nevertheless somewhat surprising that he did not in the process notice that his second rule takes into account only “own tax elasticity” and hence implicitly assumes the zero cross elasticities which he finds so objectionable in our second and third forms. It is probably because of this that he comes to the conclusion that our first form holds only when cross-elasticities are zero, which is simply incorrect.

Lerner's handling of our first form of the theorem (which, incidentally, is the one form we did not distill from the literature although since deriving it, we have seen essentially the same variant in Diamond and Mirrlees) is altogether curious. Almost paradigmatic of the economist's approach to a welfare problem, say, how large a park to build, is the following: “Clearly our rule should be, enlarge the park until the incremental social gain from enlargement by an additional unit is just balanced by the social loss from the space and other resources thereby foregone. But that is, of course, purely tautological. What we now need to do is clothe the concepts of social gain and social loss with operational measures. . . .” In this case Lerner goes backwards: “The equalization of the ratio between marginal yield and output level turns out to be an obscure way of expressing the equalization of marginal tax yield and marginal social damage—an equalization that minimizes the total social damage for a given total tax yield” (p. 289). We were, as it happens, especially pleased with the first form of the theorem which we think is new, precisely

---

21 Analytically, it does serve to a set a price level so that the government profit constraint in pecuniary form is pinned down. In the case of more than one input factor it does, however, hide rather strong assumptions about the transformation function in a full general equilibrium treatment. In any event, the pecuniary variant of the government's budget constraint is perhaps not the most felicitous for use in such a treatment. The theorems in question are, however, in no way sensitive to these assumptions. Again we recommend perusal of the literature on this point.
because it seemed to promise more than the others a degree of operational applicability. Our feeling was that experienced tax officials (or monopoly managers) would have a much better chance of estimating at least roughly the magnitudes that enter that expression—the effect on the total tax take (profit) (including cross effects) of a change in a given tax (price) than econometricians would have of estimating the full matrix of cross elasticities, just as it seems plausible that a businessman may come close to maximizing profits even though he is pretty vague about quantities such as marginal cost and marginal revenue. If this is misplaced faith it doesn’t matter, because the first form of the theorem, being equivalent to the others (and for more forms see our forthcoming paper), is at least no harder than they to put into practice. One thing is clear, the operational meaning of the first form is easier than the others (except, perhaps, for the fourth) to explain to a non-economist.

The derivations of the second and third forms of our theorem, like those of the first and fourth, make no use of and in no way depend upon constancy of marginal cost. The second and third variants do, however, involve an assumption of zero cross elasticities of demand among the goods whose prices are in question. Lerner is quite properly disturbed by the implication of extending this assumption to all goods. The demand relationship being traced out by variations in the price vector is simply the aggregate offer curve of the economy with the initial commodity vector of each citizen specified (generally at zero) and no net transfers. Independent demands for the individual goods are impossible here except in the trivial case in which the demanded vector is completely independent of the price vector, a vector which would in turn have to be the zero net transactions vector. The implication is not that the results are incorrect but that the second and third forms are not relevant when all goods are taxable and there are no net transfers.

Again, contrary to Lerner, the fourth form of the theorem is as general as the first, and more general than the second and third (not because it doesn’t depend upon constant costs—they don’t either—but because it does not assume zero cross-elasticities). Given the parenthetical word “approximately” its generality would seem assured. But we were willing to go farther than that, and took great pains to make precise in Section III, in the algebraic development of the fourth variant, what operationally was meant by “approximately” here. In footnote 9 we spell this out still further.

REFERENCES


