OUTPUT AND WELFARE IMPLICATIONS OF MONOPOLISTIC

THIRD-DEGREE PRICE DISCRIMINATION

Richard Schmalensee*

A number of Joan Robinson's classic results on third-degree monopolistic price discrimination are generalized and extended. The relation between demand function curvature and the impact of monopolistic discrimination on total output is clarified in the general N-market case. It is shown that unless total output is increased sufficiently, monopolistic discrimination produces a net (Marshallian) efficiency loss. Qualifications necessary if discrimination allows new markets to be served are discussed.

Under pure Pigouvian third-degree price discrimination, a monopolist maximizes profits by charging different prices to different markets or classes of customers, with no (second-degree discriminatory) bulk discounts or other nonlinear pricing allowed. The standard comparison of such conduct with that of a single-price monopoly remains that presented by Joan Robinson (Book V) almost a half-century ago. Using an algebraic approach, this note generalizes and extends some of her main results.

Robinson (pp. 190-2) shows geometrically that if a single-price monopoly selling in two markets under constant costs is allowed to discriminate between them, total output is unchanged if both markets have linear demand curves.[1] This result is easily extended to the N-market case below. If demand curves are not linear, she argues (pp. 192-5) that a comparison of their "adjusted concavities" at the nondiscriminating monopoly price determines whether total output rises or falls. Her formal argument depends critically on the assumption that the discriminating monopoly's prices are nearly equal, however, and Melvin Greenhut and H. Ohta show by (non-pathological) example that her proposed test does not work when those prices differ substantially.[2] The general relation between curvature of demand functions and total output changes due to monopolistic discrimination is analyzed below.

Robinson's discussion of the welfare implications of discrimination (ch. 16) is very brief and informal, and it emphasizes equity as much as efficiency. Perhaps because of this, much of the subsequent literature seems to equate the efficiency effects of discrimination with its impact on total output. Basil Yamey uses a rather special example to argue that this equation is invalid; he asserts that in general the usual

Marshallian welfare measure falls unless output increases. This assertion
is confimed below for the general N-market case with arbitrary demand
function curvatures. As an immediate corollary, it follows that if all
demands are linear, prohibiting a monopoly from practicing third-degree
discrimination produces a net welfare gain.[3]

The intuition behind these last results was presented by Pigou
(pp. 284-5, 288-9) and cited by Robinson (p. 206). For any fixed total
output of the monopolized product, efficiency requires that all buyers
have the same marginal valuation of additional units. (If all buyers are
households, they must have the same marginal rate of substitution between
the good involved and any numeraire good.) Selling the same product at
different prices to different buyers induces different marginal valuations
and produces what Robinson (p. 206) terms "a maldistribution of re-
sources as between different uses." Only an increase in total output above
the single-price monopoly level can serve to offset this distributional
inefficiency. Thus, unless total output increases, monopolistic third-
degree price discrimination produces a net efficiency loss.

I. Preliminaries

Consider a monopolist selling in N distinguishable markets (or to
N distinguishable customer classes). Let q_i be unit sales in market i,
let Q be the sum of the q_i, and let p_i be the price charged in market i. For
simplicity, it is generally assumed to be optimal for the monopolist to
make positive sales in all N markets, whether or not discrimination is
possible. (See footnote 1, above.) Following the relevant literature
(for reasons discussed below), q_i is assumed to depend only on p_i for
i = 1,..., N, and unit cost, c, is assumed constant.

The monopoly's total profit can be written as

$$(1) \qquad \Pi = \sum_{i=1}^{N} (p_i - c)q_i(p_i) \equiv \sum_{i=1}^{N} \pi_i(p_i),$$

where $q_i(p_i)$ is the demand function for market i, and $\pi_i(p_i)$ is net profit

generated in that market, for i = 1, ..., N. It is assumed that the π_i are

smooth, strictly concave functions. (We basically need smooth and declining

marginal revenue curves.) If discrimination is impossible, profits are

maximized by charging all buyers p*, the unique solution to

$$(2) \qquad \sum_{i=1}^{N} \pi_i'(p^*) = \sum_{i=1}^{N} \left[(p^* - c) \, q_i'(p^*) + q_i(p^*) \right] = 0$$

On the other hand, if pure third-degree discrimination is possible, the N

optimal prices are found as the unique solutions to

$$(3) \qquad \pi_i'(p_i^*) = \left[(p_i^* - c) \, q_i'(p_i^*) + q_i(p_i^*) \right] = 0 \qquad i = 1, ..., N.$$

Following Robinson's terminology, let the _strong_ markets be those in

which p_i^* exceeds p*, and let S be the set of the corresponding indices.

Similarly, let W be the set of indices of the _weak_ markets, in which

$p^* > p_i^*$, and let I be the set of indices of the _intermediate_ markets, if

any, for which $p^* = p_i^*$. Because unit cost is constant and individual

market demands are determined only by own prices, it is immediate that

iϵS (iϵW) if and only if $\pi_i'(p^*)$ is positive (negative). Under such regular

and separable conditions, gradient methods work well, and a related method

is used in what follows. It may be possible to permit some cross-effects

and still use the basic approach employed here, but I have so far found no economically meaningful way of doing this.

If income effects are assumed small and distributional effects are neglected, we can employ the standard aggregate Marshallian welfare indicator, consumers' surplus plus producer's (excess) profits:

$$(4) \qquad W = \sum_{i=1}^{N} \left\{ \int_{p_i}^{\infty} q_i(v)dv + \pi_i(p_i) \right\}.$$

Let us consider a set of smooth functions, $p_i(t)$, for $i = 1, \ldots, N$ and $t \geq 0$, such that $p_i(0) = p^*$ and p_i^* is the limit of $p_i(t)$ as t increases. We can then compare single-price and discriminating monopoly output, for instance, by examining (dQ/dt) along such a transition path. It is useful to impose two restrictions on the $p_i(t)$. First, dp_i/dt must have the sign of π_i'. Thus for $i \varepsilon S$ ($i \varepsilon W$), p_i is monotonically increasing (decreasing). Second, the sum of the π_i' must be zero for all t. (This sum is zero at both ends of the transition by (2) and (3), above.) This second restriction implies

$$(5) \qquad \sum_{i=1}^{N} \left[\pi_i'' \right] p_i' = \sum_{i=1}^{N} \left[2q_i' + (p_i - c) q_i'' \right] p_i' = 0.$$

An examination of Figure 1 establishes that such functions exist when $N = 2$, and the relevant aspects of the geometry there are also present when N exceeds two. Because cross-price effects have been assumed away, the iso-profit curves always have positive slope when $p_w^* < p_w < p^*$ and $p^* < p_s < p_s^*$. Point N is the nondiscriminating monopoly optimum, while

D is the discriminating monopoly equilibrium. ZZ' is the locus of points such that $(\pi_s' + \pi_w')$ is zero; its negative slope follows from the strict concavity of the π_i. Clearly, any pair of functions, $p_s(t)$ and $p_w(t)$, that move prices from N to D along ZZ' satisfy both restrictions in the preceding paragraph.[4]

2. Results

From the definition of Q and condition (5), we have immediately

$$(6) \qquad dQ/dt = \sum_{i=1}^{N} q_i' p_i' = (-1/2) \sum_{i=1}^{N} (p_i - c) q_i'' p_i'.$$

It follows directly from the second equality that dQ/dt is zero if demand curves are all linear, so that for any N, single-price and discriminating monopolies would produce the same total output.

Robinson's (pp. 193-5) "adjusted concavity" test rests on the assumption that the p_i^* are sufficiently close to p* that, essentially, one need only sign dQ/dt at t=0 in order to determine the effect of monopoly discrimination on total output. In the more natural case in which the P_i^* differ noticeably, so that discrimination suggests itself with some force, it should be clear that this sort of first-order local test can fail, essentially because dQ/dt can change sign. The "adjusted concavity" test would thus seem to have little real value.

Before presenting that test, however, Robinson (p. 193) makes some general remarks about the global consequences of demand function curvature when N = 2 that can readily be verified and extended to the case of N ≥ 2 by examination of (6). If market i is strong, the corresponding term in

the second summation in (6) has the sign of $(-q_i'')$; it is thus positive for strictly convex curves and negative for strictly concave ones. Thus if all weak markets have linear demands and all strong market demand curves are strictly convex (concave) a move from single-price to discriminating monopoly always raises (lowers) total output, no matter how much the p_i^* differ from p^* and each other. Similarly, strict concavity (convexity) of demand functions in weak markets is associated with output increases (decreases). (Recall that p_i' is always negative for $i \varepsilon W$.) If all demand functions are strictly concave or convex and if the p_i^* are not nearly equal, there is apparently no simple, general way to tell if monopolistic discrimination will raise or lower total output.[5]

All the formal analysis so far rests on the assumption that $q_i(p^*)$ and $q_i(p_i^*)$ are positive for all i. This assumption is clearly rather strong, however: some weak markets may not be served at all by a single-price monopoly even though a discriminating monopolist could profitably make sales to them. All the results above must therefore be qualified by noting the tendency of a discriminating monopoly to serve markets that would be excluded by a single-price seller. The sales made in such markets under discrimination must be added to the output increases computed above in order to assess the full effects of discrimination on total output, a point that Robinson stresses.[6] If one thinks that demand functions are as likely to be concave as convex, recognition of this effect would lead one to conclude that total output is more likely to be increased than decreased by allowing a monopoly to practice third-degree discrimination.

Next, let us go beyond Robinson's analysis and consider the

effects of allowing discrimination on the Marshallian welfare measure given by (4).[7] Differentiation of that equation and use of (6) yield

$$(7) \quad dW/dt = \sum_{i=1}^{N} \left[p_i - c \right] q_i' p_i' = \left[p^* - c \right] \left[dQ/dt \right] + \sum_{i=1}^{N} \left[p_i - p^* \right] q_i' p_i'.$$

At $t = 0$, $p_i(t) = p^*$ for all i, and the second summation is zero. It is easy to show that it is negative for all $t > 0$. If market i is intermediate, the ith term in that summation is zero for all t. But if discrimination causes anything to change, some markets must not be intermediate. If market i is strong (weak), both $\left[p_i(t) - p^* \right]$ and $p_i'(t)$ are positive (negative), and the ith term is negative as long as demand slopes down.

Integrating (7) over all non-negative t implies directly that the change in W due to discrimination is always <u>strictly</u> less than $(p^* - c)$ times the change in Q. In the linear case, we thus have a drop in Marshallian welfare. In general, unless output increases, movement from single-price to discriminating monopoly causes a fall in W, a net efficiency loss.

The first term on the right in (7) resembles the usual expression for the welfare gain from output change in a distorted market: (demand price - marginal cost) x output change. The second term reflects the efficiency cost of distributing total output inefficiently among buyers, of driving marginal valuations apart. Equation (7) indicates that the net welfare effect of allowing discrimination is the sum of an output effect of indeterminate sign and a negative distribution effect.

These two effects can be simply illustrated in the two-market

case by Figure 2. When discrimination is allowed, price in the strong

market rises from p* to p_s^*, while in the weak market it drops to p_w^*. In

the strong market, quantity falls by $(q_s^0 - q_s^1)$, while in the weak one

it rises by $(q_w^1 - q_w^0)$. The net welfare gain in the weak market is the

area a'b'e'd' = a'b'c'd' - b'c'e', while the loss in the strong market

is abcd + bce. The net gain is thus

$$\Delta W = [a'b'c'd' - abcd] - [b'c'e' + bce]$$
$$= (p^* - c)(Q^1 - Q^0) - (b'c'e' + bce).$$

The net change can thus be positive only if total output expands, only

if the increase in sales to the weak market exceeds the drop in sales to

the strong market.

If one thinks that demand curves are about as likely to be concave

as convex, and if one feels that the Marshallian measure should be taken

as seriously as it is taken in most applied welfare analysis, the fore-

going discussion might lead one to the conclusion that monopolistic third-

degree price discrimination should be outlawed.[8] As before, this must

be qualified to some extent by the possibility that such discrimination makes

it profitable to sell to markets that would not be served at all under

single-price monopoly. If discrimination makes possible a large volume

of such new sales, it can lead to an increase in welfare even if total

sales to previously served markets fail to expand.

Finally, it is worth noting explicitly that nothing here conflicts

with the "Ramsey pricing" result that a W-maximizing monopolist subject

to a lower bound on Π should practice a milder form of third-degree

discrimination.[9] That result is concerned with efficiently trading off

welfare against profit, a tradeoff not present in the context of un-regulated, profit-maximizing monopoly. In Figure 1, the point U is the unconstrainted W-maximizing point. The iso-W loci are easily proven to have negative slope when both prices are above c, as shown. It is clear that any solution to maximizing W subject to a lower bound constraint on Π must lie on the locus of tangencies UD. If the constraint is binding, pricing will involve some degree of discrimination. The point N has no special properties or attraction in this context; non-discriminatory points generally yield (W, Π) pairs that are dominated by Ramsey points on UD.

REFERENCES

William J. Baumol and David F. Bradford, "Optimal Departures from
Marginal Cost Pricing," American Economic Review, June 1970, 60,
265-83.

Raymond C. Battalio and Robert B. Ekelund, Jr., "Output Change under
Third Degree Price Discrimination," Southern Economic Journal,
October 1972, 39, 285-90.

Edgar O. Edwards, "The Analysis of Output under Discrimination,"
Econometrica, 1950, 18, 163-72.

Thomas J. Finn, "The Quantity of Output in Simple Monopoly and Discriminating
Monopoly," Southern Economic Journal, October 1974, 41, 239-43.

Melvin L. Greenhut and H. Ohta, "Joan Robinson's Criterion for Deciding
Whether Market Discrimination Reduces Output," Economic Journal,
March 1976, 86, 96-7,

John E. Kwoka, Jr., "Output under Second-Degree Price Discrimination,"
Working Paper 21, Bureau of Economics, Federal Trade Commission,
October 1979.

Merton H. Miller, "Declining Average Cost and the Theory of Railway Rates,"
Southern Economic Journal, April 1955, 21, 390-404.

A. C. Pigou, The Economics of Welfare, 4th Ed., London, 1932.

Joan Robinson, The Economics of Imperfect Competition, London, 1933.

Basil Yamey, "Monopolistic Price Discrimination and Economic Welfare,"
Journal of Law and Economics, October 1974, 17, 377-80.

FOOTNOTES

*Professor of Applied Economics, Sloan School of Management, Massachusetts Institute of Technology.

1. It is important to point out, as Robinson (pp. 188-90) does but many subsequent authors do not, that this result depends critically on the assumption that both markets are served under both regimes. In general, the profit-maximizing non-discriminatory price may be so high that no purchases are made in markets that would be profitably served under discrimination. If this occurs and demands are linear, allowing discrimination serves to increase total output by exactly the amount sold in the previously excluded markets. (See Merton H. Miller and Raymond C. Battalio and Robert B. Ekelund, Jr., for more on such cases.) On the other hand, John E. Kwoka, Jr., has recently shown by example that if such exclusion can occur, allowing a monopoly to practice second-degree discrimination (in the form of declining-block pricing) can reduce total output when demand curves cross.

2. The related local tests of Edgar O. Edwards and Thomas J. Finn share this same defect.

3. This result and that cited in the preceding sentence require that the same markets be served under both regimes (see footnote 1, above) and that distributional and income effects be neglected, so that the aggregate Marshallian surplus has welfare content.

4. With N markets, the existence problem is solved if it can be shown that for any non-negative t there exists an N-vector, $[p_i']$, such that (a) $\Sigma \pi_i'' p_i' = 0$, (b) for $i \notin I$, $p_i' \pi_i' > 0$, and (c) for $i \in I$, $p_i' = 0$. Since

the π_i'' are all negative, (a) is satisfied by a subspace of dimension (N-1) of vectors with components of different signs. Restriction (c) reduces the dimensionality by the number of markets in I, which cannot exceed (N-2) if discrimination is profitable. Restriction (b) then simply excludes half the remaining subspace, which must be at least of dimension one, as it is in Figure 1. A set of $p_i(t)$ functions can thus always be constructed by integration.

5. The approach used here does not seem to yield anything of interest when all demand functions have constant elasticities, for instance.

6. See footnote 1, above, and the references there cited.

7. As far as I know, only Yamey has formally considered this measure in the present context, and his treatment is confined to an illustrative example that does not explicitly involve third-degree discrimination.

8. It should be clear that such a conclusion would not constitute an endorsement of the Robinson-Putman Act, which cannot fairly be described as simply prohibiting the form of discrimination analyzed here.

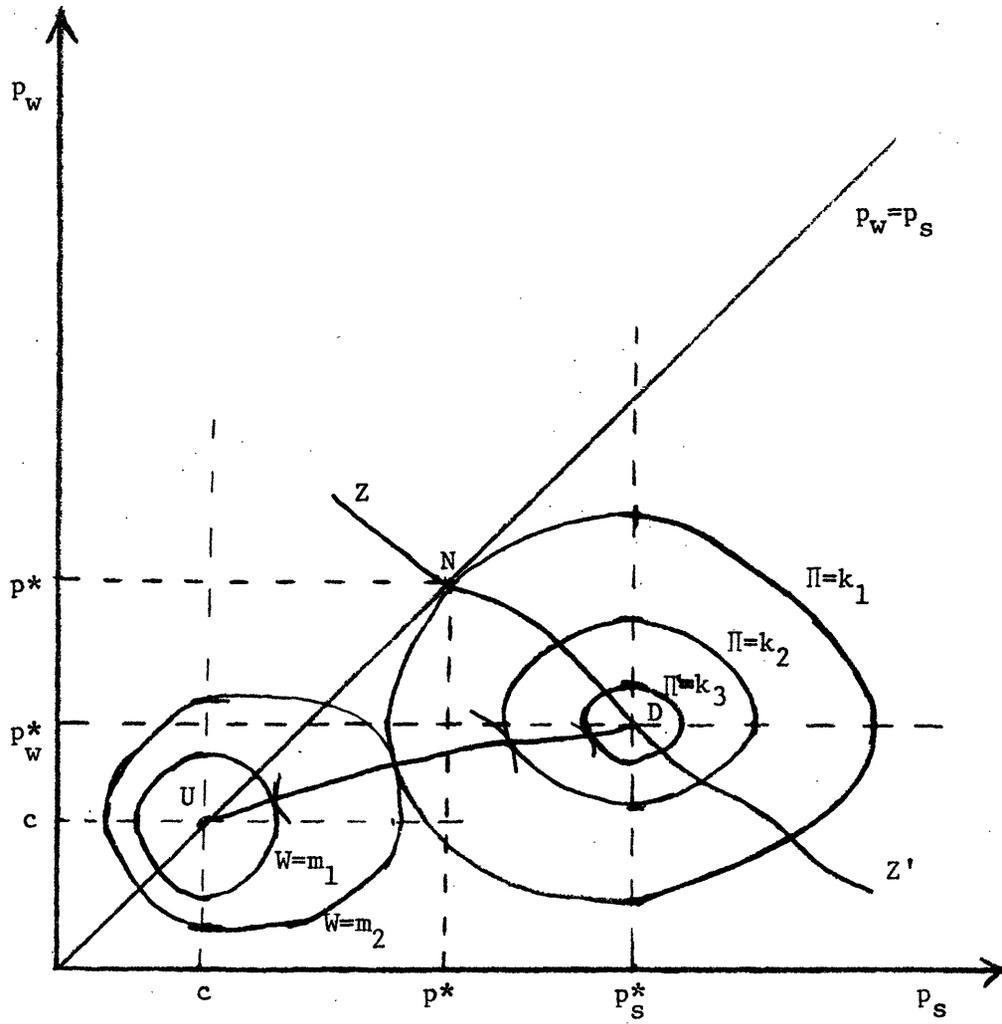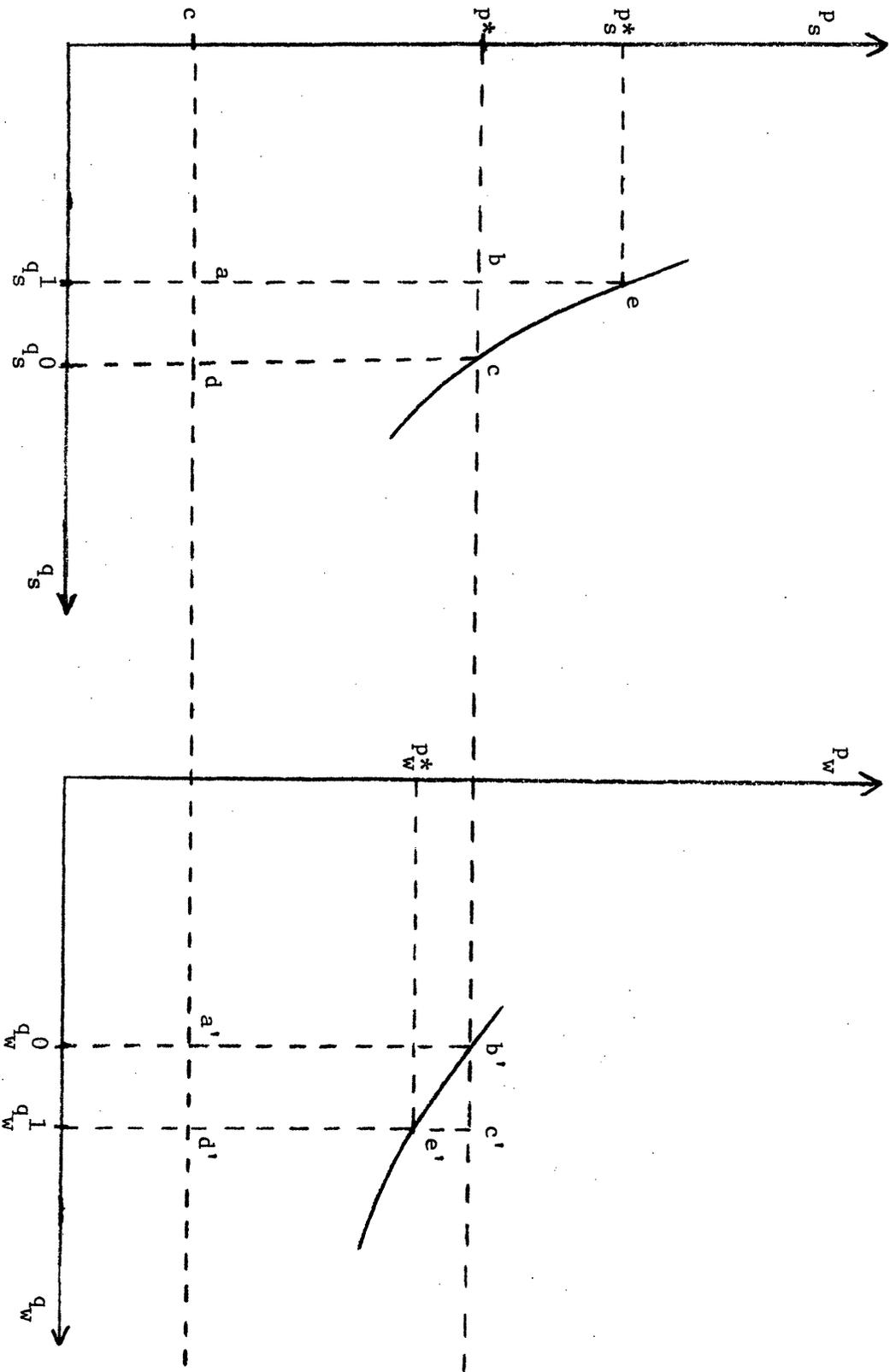9. See, for instance, William Baumol and David Bradford.

Figure 1

Figure 2