

NBER WORKING PAPER SERIES

THE SOCIAL COSTS
OF MONOPOLY AND REGULATION

Richard A. Posner*

Working Paper No. 55

CENTER FOR ECONOMIC ANALYSIS OF HUMAN BEHAVIOR AND SOCIAL INSTITUTIONS
National Bureau of Economic Research, Inc.
204 Junipero Serra Blvd., Stanford, CA 94305

September 1974

Preliminary; Not for Quotation

NBER working papers are distributed informally and in limited number for comments only. They should not be quoted without written permission.

This report has not undergone the review accorded official NBER publications; in particular, it has not yet been submitted for approval by the Board of Directors.

*The University of Chicago and NBER

Research on this paper was supported by a grant from the National Science Foundation to the NBER for its research in law and economics.

THE SOCIAL COSTS OF MONOPOLY AND REGULATION

Richard A. Posner*

I

When an industry is monopolized, price rises above and output falls below the competitive level. Those who continue to buy the product at the higher price suffer a loss, L in Figure 1, but this loss is exactly offset by the additional revenue that the monopolist obtains by charging the higher price. Other consumers, who are deflected by the higher price to substitute goods, suffer a loss, D , that is not offset by gains to the monopolist. This is the "deadweight loss" from monopoly, and in conventional analysis the only social cost of monopoly-- L being regarded merely as a transfer from consumers to owners of the monopoly seller. D , however,

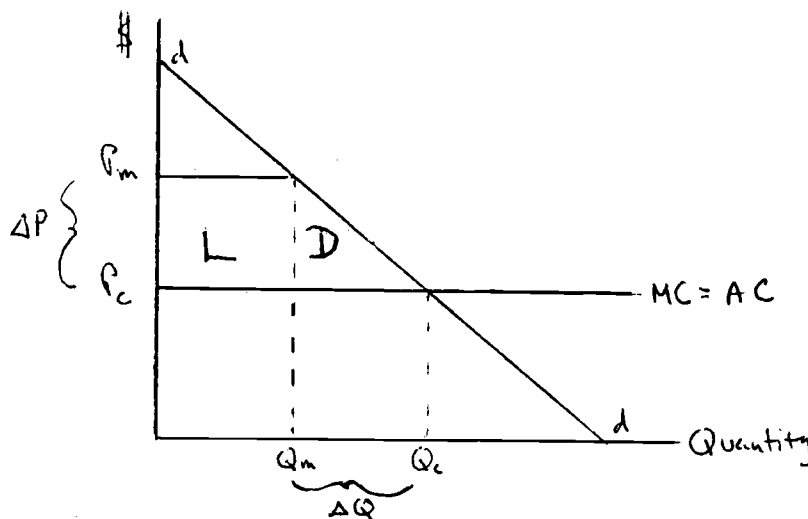


Figure 1

*Professor of Law, University of Chicago; Senior Research Associate, National Bureau of Economic Research. I am grateful to Gary S. Becker and George J. Stigler for helpful comments. Research on this paper was supported by a grant from the National Science Foundation to the National Bureau of Economic Research for research in law and economics. The paper is not an official National Bureau publication since it has not yet undergone the full critical review accorded Bureau publications, including approval by the Bureau's Board of Directors.

underestimates the social costs of monopoly. The existence of an opportunity to obtain monopoly profits will attract resources into efforts to obtain monopolies, and the opportunity costs of those resources are social costs of monopoly too. ^{1/} Theft provides an instructive analogy. The transfer of wealth from victim to thief does not involve an artificial limitation of output, ^{2/} but it does not follow that the social cost of theft is zero. The existence of the opportunity to make such transfers draws resources into thieving and in turn into protection against theft, and the opportunity costs of the resources consumed are social costs of theft. ^{3/}

This type of analysis has long been familiar in a few special contexts. For example, Plant's criticism of the patent system, made a generation ago, was based on the effect of the patent monopoly in drawing greater resources into invention than into activities that yield only competitive returns. ^{4/} And Telser's theory of resale price maintenance was in the same vein. ^{5/} However, although the tendency of monopoly

1. Gordon Tullock, *The Welfare Costs of Tariffs, Monopolies and Theft*, 5 *W. Econ. J.* 224 (1967).

2. If, however, a thief took three radios from a home and on the way out dropped one, which broke, this loss would correspond to the dead-weight loss of monopoly.

3. Gary S. Becker, *Crime and Punishment: An Economic Approach*, 76 *J. Pol. Econ.* 169, 171 n.3 (1968); Gordon Tullock, *supra* note 1.

There is another possible source of social cost from theft--the value of the stolen goods may be less to the thief than to the victim.

4. *The Economic Theory concerning Patents*, 1 *Economica* (n.s.) 30 (1934).

5. *Why Should Manufacturers Want Fair Trade?*, 3 *J. Law & Econ.* 86 (1960). Cf. George J. Stigler, *Price and Nonprice Competition*, in his *The Organization of Industry* 23 (1968).

rents to be transformed into costs is clearly no longer a novel insight, its implications both for the measurement of the aggregate social costs of monopoly and for a variety of other important issues relating to monopoly and public regulation (including tax policy) continue to be ignored. Studies of the costs of monopoly still focus exclusively on the deadweight loss (plus alleged consequences of monopoly that have no firm theoretical basis, such as "X-inefficiency"); analyses of public policy toward monopolies (both private and governmentally induced) continue to ignore the social costs involved in obtaining monopoly power.

The present paper is an effort to rectify this neglect. Part II presents a simple model of the social costs of monopoly, conceived as the sum of the deadweight loss and the additional loss resulting from the competition to become a monopolist. Part III uses the model to estimate the social costs of monopoly in the United States, and the social benefits of antitrust enforcement. Part IV explores the implications of the analysis for a variety of issues relating to monopoly and public regulation, such as public policy toward price discrimination and the choice between income and excise taxation.

II.

I make (and later will defend) two simplifying assumptions. The first is that becoming a monopolist is a competitive activity and hence that the expected profit from obtaining a monopoly is zero. Firms seeking monopoly expend resources until the last dollar spent increases the expected value of having a monopoly by one dollar. The second

assumption is that the costs of becoming a monopolist are constant, so that the sum of the opportunity costs of becoming a monopolist is equal to the expected value of the monopoly. ^{6/} With these assumptions, the total social costs of the monopoly depicted in Figure 1 are $D + L$.

Since $D = \frac{1}{2}\Delta P\Delta Q$ and $L = \Delta P(Q_c - \Delta Q)$, the relative sizes of D and L are given by

$$\frac{D}{L} = \frac{\Delta Q}{2(Q_c - \Delta Q)} \cdot \quad (1)$$

This ratio can also be expressed in terms of the elasticity of demand for the product in question at the competitive price (ϵ), and the percentage increase in price brought about by monopolization (p): ^{8/}

$$\frac{D}{L} = \frac{p}{2\left(\frac{1}{\epsilon} - p\right)} \cdot \quad (2)$$

It may be objected that the assumption that the monopolist's price increase is determined independently of the elasticity of demand is inconsistent with the assumption that he is a rational profit-maximizer. Profit maximization does assume that the monopolist's price is a function of the elasticity of the demand curve facing him, but a firm's demand curve may be different from the industry's demand curve, to which ϵ in equation (2) refers. For example, the demand

6. Another assumption, which does not affect the analysis, is that the monopoly is enjoyed for one period only; otherwise the optimum expenditures on obtaining a monopoly could not be compared directly with L in Figure 1.

7. This is strictly accurate only for a linear demand curve; it is only an approximation for nonlinear demand curves. The significance of this qualification is considered later.

8. $p \equiv \Delta P/P_c$. The sign of ϵ is positive in equations (2) and (3).

curve facing the individual seller in a cartelized industry will almost certainly be more elastic than the demand curve for the industry, because the cartel will not be able to eliminate all competition among its members. As a result, the actual prices charged by the cartel will be lower than the profit-maximizing level as determined with reference to the elasticity of the industry demand curve. It seems quite likely that, due to the imperfections of cartelization (especially in industries subject to the Sherman Act) and of cartel-like forms of public regulation, the prices actually charged in many monopolized industries will be lower than the optimum monopoly price for the industry.

From equation (2), it is plain that the ratio of the deadweight loss of monopoly to the loss that results from the competition to become a monopolist (what we may call the "additional loss" of monopoly) is smaller, the less elastic the demand for the product in question at the competitive price. For example, at $\epsilon = 1$, a one per cent increase in price over the competitive level will result in a deadweight loss equal to about .5 per cent of the additional loss; at $\epsilon = 2$, this figure is about one per cent; at $\epsilon = \frac{1}{2}$, it is only about .025 per cent.

C, the total social costs of monopoly, is equal to

$$D + L = pR_c - \frac{1}{2}\Delta P\Delta Q = R_c(p - \frac{1}{2}\epsilon p^2), \quad \frac{9/}{\quad} \quad (3)$$

9. R_c is total sales revenues at the competitive price.

and is higher, the larger the scale of the industry as measured by its sales revenues at the competitive output, and the less elastic the demand for the product at that output. At $\epsilon = 1$, a one per cent increase in price above the competitive level will yield a total social cost of monopoly equal to .995 per cent of the industry's revenues at the competitive price and level. At $\epsilon = 2$, the total social cost falls to .99 per cent of the industry's revenues at the competitive price and level, while at $\epsilon = \frac{1}{2}$ the percentage rises to .9975. The costs of monopoly are maximized when demand is totally inelastic at the competitive price.

These formulas are accurate only for small changes in the price level; for larger changes, ϵ , which measures elasticity at a point, can no longer be used. And monopolization might result in large price increases. Hence the general formulas, (1) and (3), remain important. For purposes of empirical estimation, it will be helpful to derive two additional formulas: one for the case where the deadweight loss, the elasticity of demand, and the monopoly price increase are known and the elasticity of demand is assumed to be constant, and the other for the case where the monopoly price increase, the monopoly output, and the elasticity of demand at the monopoly price are known and the demand curve is assumed to be linear.

(1) For constant elasticity, since $Q_c = \alpha P_c^\epsilon$ and $Q_m = \alpha P_m^\epsilon$, $\frac{10}{}$ and therefore $\Delta Q = \alpha(P_c^\epsilon - P_m^\epsilon)$,

10. In these formulas, ϵ has its usual negative sign.

$$\frac{D}{L} = \frac{(kP_m)^\epsilon - P_m^\epsilon}{2P_m^\epsilon}, \quad (4)$$

where $k \equiv P_c/P_m$. Equation (4) can be simplified to

$$\frac{D}{L} = \frac{1}{2} (k^\epsilon - 1). \quad (5)$$

From (5), we can readily deduce that C is equal to

$$D + L = D \left(1 + \frac{2}{k^\epsilon - 1}\right) = R_m(1-k) \left(\frac{k^\epsilon + 1}{2}\right), \quad (6)$$

where R_m is total sales revenue at the monopoly price and output. ^{11/}

A question arises whether, since a demand curve of constant elasticity is nonlinear, the linear approximation of the deadweight loss used in equations (5) and (6) introduces a source of serious inaccuracy. It appears not to, at least in the simple case where $\epsilon = 1$, and therefore

$$\frac{D}{L} = \frac{\int_{Q_c}^{Q_m} P dQ - P_c \Delta Q}{(P_m - P_c)Q_m} = \frac{\ln \frac{1}{k} - 1 + k}{1 - k}. \quad (7)$$

The following table compares D/L calculated from equation (5) (with $\epsilon = -1$) and equation (7), and shows that the linear approximation overestimates the deadweight loss, but not seriously.

11. For the special case where the firm is able to charge the optimum monopoly price for the industry, so that $P_c = MC = P_m \left(1 + \frac{1}{\epsilon}\right)$, equation (5) becomes

$$\frac{D}{L} = \frac{1}{2} \left(1 + \frac{1}{\epsilon}\right)^\epsilon - 1, \quad (5')$$

and equation (6) becomes

$$C = \frac{-R_m}{\epsilon} \left(\frac{\left(1 + \frac{1}{\epsilon}\right)^\epsilon + 1}{2}\right). \quad (6')$$

Table 1

P (Monopoly price increase)	D/L (ratio of deadweight to additional loss) Equation (5)	D/L (ratio of deadweight to additional loss) Equation (7)
5%	.025	.025
10	.050	.049
15	.075	.072
20	.100	.094
50	.250	.216

The partial derivatives of D/L in equation (4) with respect to k and ϵ are

$$\frac{\partial(D/L)}{\partial k} = \frac{\epsilon k^{\epsilon-1}}{2}; \quad \frac{\partial(D/L)}{\partial \epsilon} = \frac{1}{2} k^{\epsilon} \ln k. \quad (8)$$

$\partial(D/L)/\partial k$ is negative since ϵ is negative, meaning that the ratio of the deadweight loss of monopoly to the additional loss is smaller, the smaller the monopoly price increase (k, the ratio of the competitive to the monopoly price, is larger the smaller the relative price increase). $\partial(D/L)/\partial \epsilon$ is also negative (since the natural logarithm of a fraction is negative), meaning that the ratio of the deadweight to the additional loss of monopoly is greater, the more elastic the demand is.

(2) For the case where ^{the} elasticity of demand at the monopoly price (as well as the monopoly price increase and the quantity sold at the monopoly price) is known or can be computed, and it is believed that the demand curve can be approximated by a straight line, we first express the slope of the demand curve at the monopoly price in terms of ϵ :

$$\frac{\Delta Q}{\Delta P} = \frac{\epsilon Q_m}{P_m} \cdot \frac{12}{1} \quad (9)$$

Since the slope of a linear demand curve is constant, we can use equation (9) to find ΔQ ; it is $\epsilon Q_m p k$. Equation (9) can now be transformed into

$$C = R_m(1-k)\left(1 + \frac{1}{2}\epsilon(1-k)\right), \quad (10)$$

and D/L is given by

$$\frac{D}{L} = \frac{\epsilon(1-k)}{2} \cdot \frac{13}{1} \quad (11)$$

We now have two formulas for the ratio of the deadweight to the additional loss from monopoly--equations (5) and (11)--and we ask: how different are the estimates that they produce, under various assumptions as to the monopoly price increase and the elasticity of demand? The answer is, not very, for price increases of less than 25 per cent, and even for much larger price increases if the elasticity of demand is no greater than one. These results are shown in Figure 2.

III A

The formulas developed in the preceding part could be used to derive, from the estimates used by Arnold Harberger and others of the deadweight

12. The sign of ϵ is positive in equations (10) and (11).

13. For the special case where the firm is able to charge the optimum monopoly price, equation (10) becomes simply

$$C = \frac{-R_m}{2\epsilon}, \quad (10')$$

and equation (11) becomes

$$\frac{D}{L} = 2. \quad (11')$$

ϵ has its usual negative sign in equation (10').

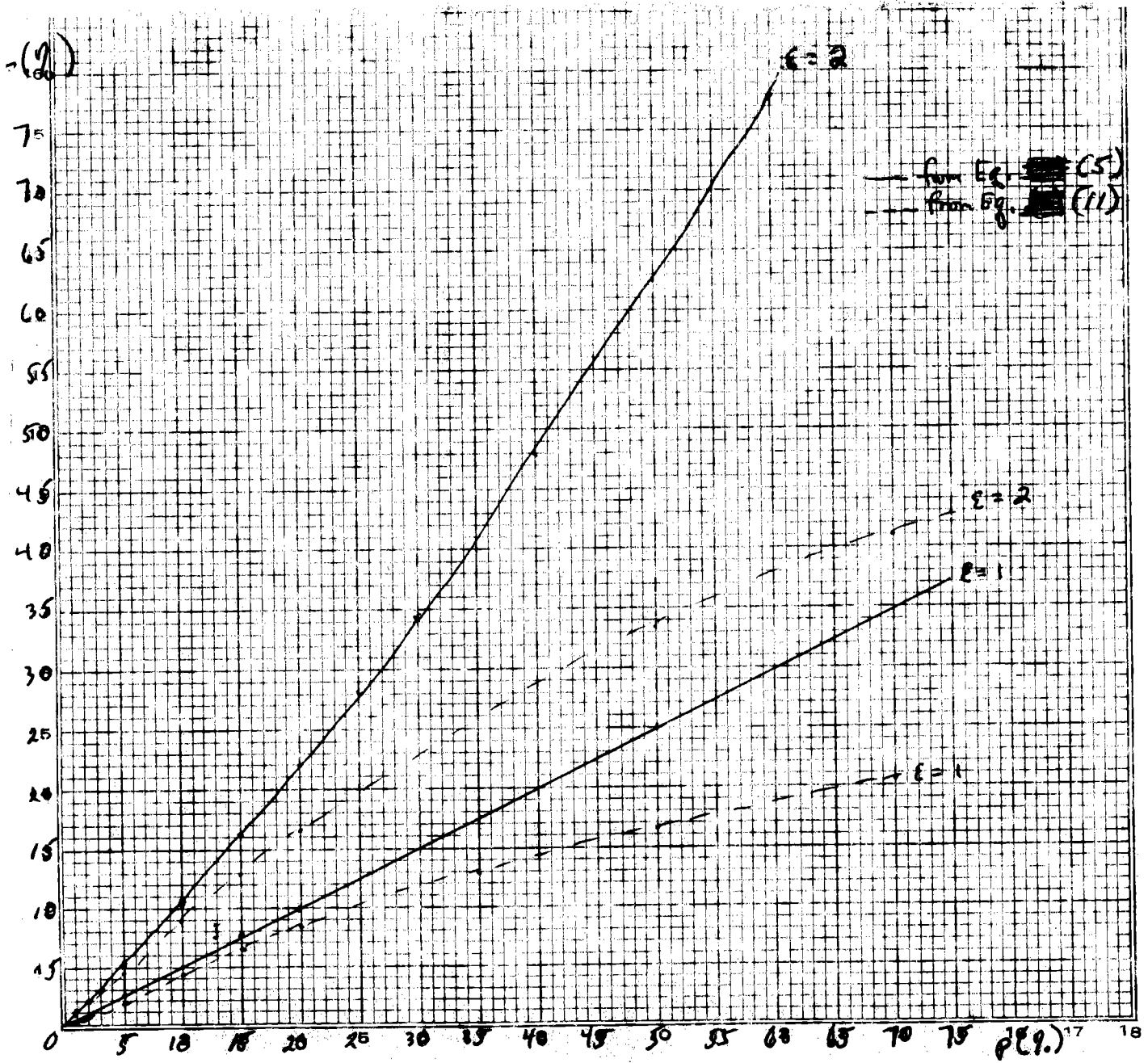


Figure 2

loss of monopoly, an estimate of the total social cost of monopoly. Harberger, estimating an average monopoly price increase of about six per cent and assuming that the elasticity of demand was constant and equal to -1 , found the deadweight loss from monopoly in the manufacturing sector to be equal to (at most) .1 per cent of GNP. ^{14/} Harberger's (implicit) k is thus $1/1.06$ or .9434, and from this and his estimate of ϵ , we calculate from equation (5) that the ratio of D to L in Harberger's analysis is .03. Hence if D is .1 per cent of GNP, L is about 3.3 per cent and C about 3.4 per cent--about \$40 billion today. David Schwartzman used similar methods and found D equal to about .1 per cent of GNP too. ^{15/} But he assumed a price increase of 8.3 per cent and an $\epsilon = -1.5$. Plugging these values into equation (5) yields $D/L = .06$. Hence if $D = .1$ per cent of GNP, $L = 1.7$ per cent and C about 1.8 per cent.

However, neither estimate can be taken seriously because both Harberger and Schwartzman, in determining the monopoly price increase, relied on rate-of-return statistics. Above-average rates of return were used not only to identify the monopolized industries but also to calculate the monopoly price increase. If the analysis in this paper is correct, such a procedure is improper. The monopolist is not expected to enjoy a supernormal rate of return, but only the normal

14. Arnold C. Harberger, Monopoly and Resource Allocation, 44 Am. Econ. Rev. Papers and Proceedings 77 (1954), reprinted in his Taxation and Welfare 91 (1974).

15. The Burden of Monopoly, 68 J. Pol. Econ. 627 (1960).

return ^{16/} (assuming, realistically, that the investment which he makes in obtaining the monopoly is written off over the life of the monopoly). ^{17/} This point is distinct from the objections to Harberger's procedure raised by Goerge Stigler: that monopoly profits are often capitalized into the valuation of a firm's assets and that some of the profits may be received as rents by suppliers of the firm's inputs. ^{18/}

The proper method of calculating the social costs of monopoly (deadweight plus additional loss) is to obtain estimates of the monopoly price increase, and of the elasticity of demand at the relevant points along the demand curve, from industry studies. An independent estimate of the elasticity of demand would of course be unnecessary if we could assume that, after the price increase, the price charged was the optimum monopoly price; and where an independent estimate of ϵ is available, it can serve as a check on that assumption. To illustrate, there have been a number of estimates of the percentage by which CAB regulation has increased the price of airline travel. The simple average of these estimates ^{is} ^{.66.} ^{19/} If a 66 per cent price increase

16. Subject to an important qualification, discussed later, when there is uncertainty in the market for becoming a monopolist.

17. Thus, if it costs \$10 million to obtain a monopoly that will yield a profit (net of the cost of production and sale) of \$1 million a year for 10 years, we can expect the annual expenses of the company to be adjusted upward, by depreciation or other accounting techniques, by \$1 million a year throughout the 10-year period.

18. George J. Stigler, *The Statistics of Monopoly and Mergers*, 64 *J. Pol. Econ.* 33 (1956).

19. Computed from Richard E. Caves, *Air Transport and Its Regulators* 372 (1962); William A. Jordan, *Airline Regulation in America* 110-11, 124-25 (1970); Note, *Is Regulation Necessary?*, *California Air Transportation and National Regulatory Policy*, 74 *Yale L. J.* 1416, 1435-36 (1965).

over competitive levels is assumed to raise the price of air travel to the optimum monopoly level, then the elasticity of demand at the monopoly price can be calculated from the formula which equates marginal cost to marginal revenue, ^{20/} $MC = P_m(1 + \frac{1}{\epsilon})$. Since $MC = P_c$ and $P_c = .6P_m$, $\epsilon = -2.5$ at the monopoly price. An independent estimate of the long-run elasticity of demand for air travel made by Houthakker and Taylor is ^{21/} -2.36 , which is virtually identical to our calculation.

We have two choices now. We can assume constant elasticity and solve for D/L using equation (5'), or we can assume a linear demand curve and solve for D/L using equation (10'). With the assumption of constant elasticity, $D \approx 2.04 L$, and (from equation (6')) it is readily calculable that the total social cost of the airline monopoly is approximately equal to 2.48 times the total revenue at the monopoly price.

However, the assumption of a linear demand curve seems more plausible than the assumption of constant elasticity, especially for large relative price increases, which one expects to find associated with a rising elasticity of demand as substitutes become increasingly attractive.

20. This was essentially the procedure used by David Kamerschen, in *An Estimation of the Welfare Losses from Monopoly in the American Economy*, 4 W. Econ. J. 221 (1966), to estimate the deadweight loss from monopoly in manufacturing. He has been criticized, and rightly so, for assuming that firms in concentrated industries subject to the Sherman Act's prohibition of collusive pricing are typically able to charge the profit-maximizing monopoly price.

21. H. S. Houthakker & Lester D. Taylor, *Consumer Demand in the United States, 1929-1970*, at 124 (1966). This must be the elasticity of demand at the regulated price, since only a small part of the airline industry is not subject to CAB regulation.

And if equations (10') and (11') are used, D is one half as great as L , and C is equal to 20 per cent of the total revenues of the industry at the monopoly price. Still, this indicates a very large social loss from the regulation-induced airline monopoly.

The foregoing exercise brings out an important general point that has been completely ignored in previous studies of the cost of monopoly to the economy: much of this cost probably originates in the regulated sector of the economy. The ability of firms to maintain supracompetitive prices must be greater in industries in which a regulatory agency limits entry and price competition than in the manufacturing sector where express collusion is forbidden by the Sherman Act. Yet all of the previous studies of the cost of monopoly to the economy have been based on supposed monopoly pricing in manufacturing alone.

Table 2 collects estimates of the regulation-induced price increase, and the elasticity of demand at the current price, for several industries for which these data are available. Two estimates of elasticity are given; one (ϵ_1) is derived from the price-increase data, on the assumption that the industry is charging the optimum monopoly price, and the other (ϵ_2) is an independent estimate of elasticity. The estimates of the total social costs of the regulation in question, which appear in the last two columns of the table, ^{22/} are based on the assumption that the industry's demand curve is linear in the relevant region, and are expressed as a percentage of the total revenues of the industry. Our airline example appears in the last row of the table.

22. C_1 is computed from ϵ_1 , C_2 from ϵ_2 .

Table 2
The Social Costs of Regulation

Industry	Regulatory Price Increases	ϵ_1	ϵ_2	C_1	C_2
Physicians' services	.67 ^a	-1.43	-.5753 ^b	.29	.35
Eyeglasses	.34 ^c	-.394	-.45 ^d	.13	.24
Milk	.11 ^e	-10.00	-.339 ^f	.05	.10
Motor carriers	.62 ^g	-2.63	-1.14 ^h	.19	.30
Oil	.65 ⁱ	-2.5	-.9 ⁱ	.20	.32
Airlines	.66	-2.5	-2.36	.20	.19

a Unpublished study by Reuben A. Kessel.

b. H. S. Houthakker & Lester D. Taylor, Consumer Demand in the United States, 1929-1970, at 99 (1966)(short-run).

c Lee Benham, Price Structure and Professional Control of Information, p. 19 (March 1973).

d Id. at 30 (simple average).

e Reuben Kessel, Economic Effects of Federal Regulation of Milk Markets, 10 J. Law & Econ. 51, 73 (1967).

f H. S. Houthakker, New Evidence on Demand Elasticities, 33 Econometrica 277, 286 (1965). This estimate is for all food; an estimate limited to dairy products in the Netherlands was not significantly different. Robert Ayaynian, A Comparison of Barten's Estimated Demand Elasticities With Those Obtained Using Forsch's Method, 37 Econometrica 79 (1969).

g Average of estimates in Department of Agriculture studies cited in Thomas Gale Moore, Freight Transportation Regulation 72 (American Enterprise Institute 1972); Richard N. Farmer, The Case for Unregulated Truck Transportation, 46 J. Farm Econ. 398 (1964).

h Simple averages of various estimates for transportation in Scandinavia. See Regnar Forsch, A Complete Scheme for Computing All Direct Costs and Cross Demand Elasticities in a Model With Many Sectors, 27 Econometrica 177 (1959); Richard W. Parks, Systems of Demand Equations: An Empirical Comparison of Alternative Functional Forms, 37 Econometrica 629, 649 (1969).

i Cabinet Task Force on Oil Import Control, The Oil Import Question (1970).

These estimates are crude, although they have the virtue of being improvable. They do suggest, however, that the social costs of regulation are probably extremely high, given that about 17 per cent of GNP comes from industries such as agriculture, transportation, communications, power, banking, insurance, and medical services that contain the sorts of controls over competition that might be expected to lead to supracompetitive prices. ^{23/} True, a much higher percentage of GNP--30 per cent--originates in manufacturing and mining, a highly concentrated sector of the economy, and the conventional wisdom associates high concentration with supracompetitive pricing. But only about a fifth of the output of this sector comes from industries in which four firms account for 60 per cent or more of sales, and there is little theoretical basis for believing that the sellers in less concentrated industries could collude effectively without engaging in behavior prohibited by the Sherman Act. ^{24/} Not all violations of the Sherman Act are detected and punished, of course, but the secret conspiracies that escape detection are probably not very effective--even the great electrical conspiracy, a highly elaborate and long-continued conspiracy

23. This figure may seem an overstatement, since not all of the markets in the regulated industries are in fact subject to the relevant regulatory controls (almost half of the trucking industry, for example, is exempt from regulation by the ICC). On the other hand, tariffs and similar restrictions (e.g., the oil import quotas) are excluded from the estimate of the percentage of GNP affected by regulation.

24. Thus, Kessel's study of underwriting costs shows that an increase beyond eight in the number of bids does not reduce those costs; and an industry where the four largest firms have less than 60 per cent of the market is apt to contain at least eight significant competitors. Reuben Kessel, A Study of the Effects of Competition in the Tax-Exempt Bond Market, 79 J. Pol. Econ. 706, 723 (1971).

among a very small group of firms, apparently succeeded in raising the price level by only 8.8 per cent on average. ^{25/} It seems highly unlikely that the price level of the manufacturing and mining sector as a whole is more than about two per cent above the competitive level.^{26/} Assume that it is two, and that the average elasticity of demand for the products of this sector, at current prices, is -1.1607. ^{27/} Then, from equation (10), the total social costs of monopoly in this sector are 1.9 per cent of the total revenues generated in the sector, and while this is substantial, it amounts to a total dollar loss substantially smaller than that generated in the regulated sector. ^{28/}

25. U.S. Cong., Jt. Comm. on Internal Revenue Taxation, Staff Study of Income Tax Treatment of Treble Damage Payments Under the Antitrust Laws 39 (Nov. 1, 1965).

26. If it is assumed that only in industries where the four-firm concentration ratio exceeds .6 is effective undetectable collusion likely, and that it allows these industries to maintain, on average, prices five per cent above the competitive level, while in the rest of the manufacturing and mining sector the average price level is only one per cent above the competitive level, then average prices for the entire sector would only be 1.83 per cent above the competitive price level. (Statistics on the distribution of output among industries in different four-firm concentration ratio groups are from the 1963 Census of Manufacturers.)

27. This figure is a simple average of the long-run price elasticities for nine product groups within the manufacturing and mining sector estimated by Houthakker & Taylor, *supra* note 21, at 72, 74, 83, 112-14, 116, 128-31.

28. The simple average of the social-cost estimates presented in Table 2 is 20.4 per cent of the total revenues of the regulated industry. Assuming that 50 per cent of the output of that sector is produced in markets that are regulated in a manner similar to the industries in Table 2 and that the average social cost of regulation in each such market is 20.4 per cent of total revenue, then it would follow that the social costs of regulation were equal to 1.7 per cent of GNP, while the social costs of monopoly in manufacturing and mining would be equal to only .6 per cent of GNP. These estimates exclude, of course, the administrative costs of regulation and antitrust enforcement in the two sectors respectively, and the benefits of monopoly. The benefits are likely to be greater in the manufacturing and mining sector, where much concentration and the resulting monopoly pricing may be due to efficiencies of various

Another use to which the analysis developed here can be put is estimating the social benefits of the antitrust laws. Table 3, which is constructed on the same basis as Table 2, presents estimates of the social costs of several cartels not subject to the prohibitions of the American antitrust laws.

Table 3
Social Costs of Cartelization

Industry	Cartel Price Increase	ϵ_1	ϵ_2	C_1	C_2
Nitrogen	.75 ^a	-2.3256	-1.4493 ^b	.21	.30
Sugar	.30 ^c	-4.3276	-.339 ^d	.12	.22
Aluminum	1.00 ^e	-2	n.a.	.25	--
Aluminum	.38 ^f	-3.6311	n.a.	.14	--
Rubber	1.00 ^g	-2	n.a.	.25	--
Electric bulbs	.37 ^h	-3.7023	n.a.	.14	--
Copper	.31 ⁱ	-4.2499	n.a.	.12	--

a George W. Stocking & Myron W. Watkins, *Cartels in Action* 163 (1946).

b Id. at 166.

c Id. at 46.

d H.S. Houthakker, *New Evidence on Demand Elasticities*, 33 *Econometrica* 277, 286 (1965) (food --obviously a much too low estimate for one food product sold at a cartel price!).

e Stocking & Watkins, supra note a, at 228.

f Id. at 251.

g Id. at 64-65.

h Id. at 343.

i. George W. Stocking & Myron W. Watkins, *Cartels or Competition?* 127 (1948).

28. continued

sorts. See Harold Demsetz, *Industry Structure, Market Rivalry, and Public Policy*, 16 *J. Law & Econ.* 1 (1973). There is no accepted theory which attributes social benefits to regulation limiting entry and price competition; but see note 29 infra.

Presumably, collusive price increases of this magnitude, and the attendant very substantial social costs, are deterred by the American antitrust laws. However, a complete cost-benefit analysis of antitrust law would also require estimation of (1) the costs of administering the antitrust laws and (2) the large social costs imposed by the many perverse applications of those laws, which may be an inevitable by-product of having antitrust laws.

The conclusions that emerge from our estimation exercises are necessarily tentative; they are that:

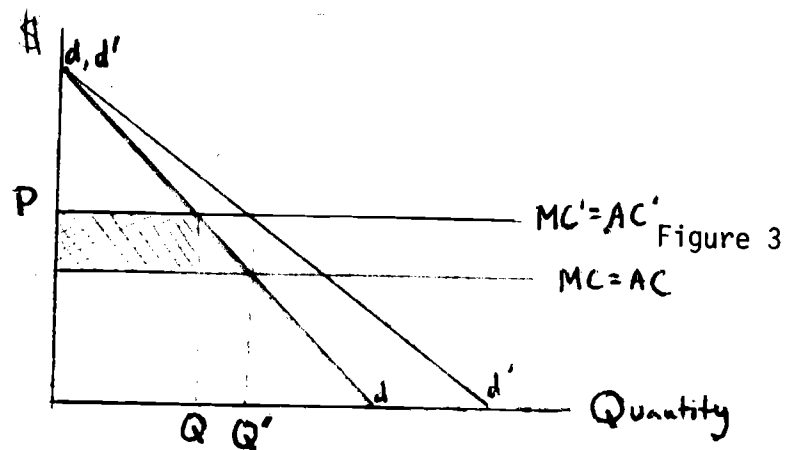
1. the social costs of public regulation, in increasing prices in the regulated industries above competitive levels, are probably very great--perhaps two per cent of our entire GNP;
2. the social costs of monopoly in manufacturing and mining are, in contrast, probably much smaller even though a larger sector of the economy is affected;
3. the enforcement of the antitrust laws against cartels has probably generated large social-cost savings, but at what cost we do not know.

B

The validity of the approach of this paper, and of the estimates presented in the preceding subpart, depend, of course, on the soundness of our basic assumptions that the market for becoming a monopolist is a

competitive one and that marginal costs in that market are constant; and it is time we examined those assumptions critically. The first is a standard assumption of economics, and, pending better evidence than we have, seems reasonable in the present context. Anyone can try to obtain a patent, a certificate of public convenience and necessity, a television license, a minimum-wage law, and anyone can try to enter a cartelized industry, or if he is a member of such an industry try to engross a greater share of the monopoly profits of the industry. To be sure, where there are no barriers to entry the expenditure of resources on becoming a monopolist will lead to an increase in the output of the market and a consequent reduction in price. But these situations are properly analyzed as ones where the expected monopoly profits are small and hence where the social costs of monopoly are small too.

As one example of the mechanisms by which the potential transfer of wealth from consumers to the owners of a monopolistic firm is transformed into a social cost, consider nonprice competition in the airline industry. Assume that the CAB places a floor under airline prices that exceeds the marginal cost of providing air transportation under competitive conditions. If the regulation is not anticipated, the situation initially is as depicted in Figure 3, and is unstable. Since there are several airlines, and nonprice competition is not constrained, the airlines will expend resources on such competition (better service, etc.) until the marginal costs of air transportation rise to the level (p in Figure 3) where the industry is earning only a normal return.



(Judging from the profit statements and stock prices of the regulated airlines, their return is indeed no higher than normal.) By this process, the monopoly profit initially generated by the regulatory price floor--the shaded rectangle--is transformed into higher costs for the industry, and these costs are a social cost of regulation. The demand curve shifts to the right because the increased expenditures on service presumably improve the product from the standpoint of the consumer. The shift produces some additional consumer surplus but not enough to offset the higher costs--otherwise the higher level of service would have been provided without the spur of monopoly pricing. ^{29/}

This discussion suggests an important (and familiar) point about expenditures on monopolizing. If for some reason the free market is not expected to produce the optimum output of a particular good or service, the creation of a monopoly in that market may, by attracting

29. This does suggest, however, that not all of the social costs of airline regulation estimated in Part III A, supra, represent pure waste.

additional resources into the market, result in the correct output of the good. This is the economic justification of the patent system and the (nonmonopolistic) explanation for resale price maintenance; it is not a justification for the regulation of the airline industry.

Observe that if the regulatory price floor had been anticipated, the initial cost curve in Figure 3 would be incorrectly drawn. The airlines would have bid away the anticipated monopoly profits in jockeying for admission to the regulatory cartel.

Figure 3 also illustrates the second critical assumption underlying our analysis, the constancy of the marginal costs of becoming a monopolist.^{30/} The assumption seems appropriate in most contexts--there seems little reason to believe that obtaining monopolies involves the use of resources the long-run supply of which is inelastic. Some exceptions are discussed in the next part of the paper.

How is the analysis affected if there is uncertainty in the activity of becoming a monopolist or if bribery is the method used to obtain the monopoly? Given uncertainty, the expected monopoly profits of any firm seeking a monopoly will be smaller than the actual monopoly profits, and therefore the expenditures of each firm will also be smaller than the actual monopoly profits. Suppose ten firms are vying for a monopoly having a present value of \$1 million and each of them has an equal chance of obtaining it. If risk neutral, each will spend \$100,000 (assuming constant costs) on trying to obtain the monopoly. Only one

30. With rising marginal costs, the total social costs of becoming a monopolist would be less than the private social costs = expected monopoly profits; the difference would be rents received by suppliers of resources specialized to the activity of obtaining monopolies.

will succeed, and his costs will be much smaller than the monopoly profits. But the total costs of obtaining the monopoly--counting losers' expenditures as well as winners'--will be the same as under certainty--\$1 million.

If we assume that the market for monopoly is characterized by a high degree of uncertainty--a plausible assumption--we can explain why the costs of obtaining a monopoly have largely eluded observation. Most of the costs are incurred in unsuccessful efforts to obtain a monopoly--the lobbying campaign that fails, the unsuccessful attempt to obtain a bank charter or form a cartel. This point also suggests that the use of above-average accounting rates of return to identify (though not measure) monopoly power may not be entirely unsound after all. The firm (or industry) that obtains monopoly under conditions of uncertainty will, like the winner of a lottery, enjoy windfall profits ex post.

It might seem that where monopoly is obtained by bribery of government officials, the additional loss of monopoly with which this paper is concerned would be eliminated--a bribe is a pure transfer. But this conclusion would be incorrect. Bribery merely shifts the monopoly profits from the monopolist to the officials receiving the bribe, and draws real resources into the activity of becoming an official who is in a position to receive these bribes.

IV

The analysis of the costs of monopoly presented in this paper has implications for a number of issues of public policy.

1. In a recent paper Comanor and Smiley attempt to show that a large part of the inequality in the distribution of wealth in contemporary America is attributable to monopoly. ^{31/} They use the Harberger and other studies discussed in Part III of this paper to determine the aggregate wealth transfer from consumers to the owners of monopoly firms, and by a series of additional assumptions concerning the incomes of consumers and shareholders, family size, the savings rate, etc., derive an estimate of the distributive impact of monopoly. Many of the assumptions are questionable but even if their correctness were conceded the conclusion would be wrong. There is no reason for thinking that monopoly has a significant distributive effect. Consumers' wealth is not transferred to the shareholders of monopoly firms; it is dissipated in the purchase of inputs into the activity of becoming a monopolist.

2. Oliver Williamson has argued that the refusal of the courts to recognize a defense of economies of scale in merger cases under the Clayton Act is questionable because, under plausible assumptions concerning the elasticity of demand, only a small reduction in the merging firms' costs is necessary to offset any deadweight loss created by the price increase that the merger enables the firms to make. ^{32/} The nature of the comparison is shown in Figure 4.

31. William S. Comanor & Robert H. Smiley, Monopoly and the Distribution of Wealth, forthcoming in the Quarterly Journal of Economics.

32. Economics as an Antitrust Defense: The Welfare Tradeoffs, 58 Am. Econ. Rev. 18 (1968).

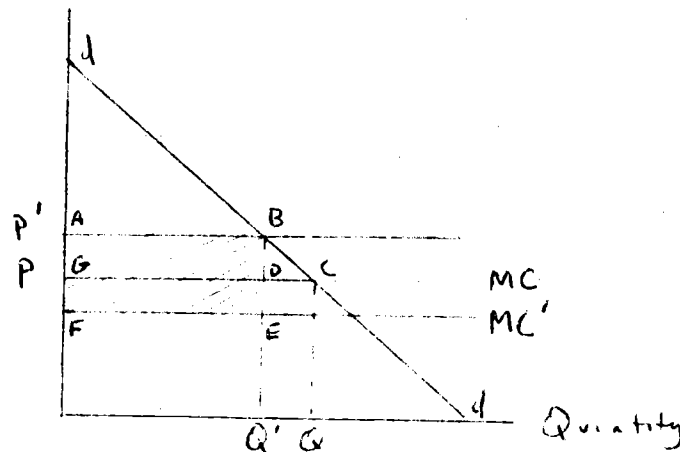


Figure 4

This analysis is incomplete, however. The expected profits of the merger--ABEF--will generate an equivalent amount of costs as firms vie, through service competition or whatever, to engross these profits (alternatively, the merging firms might expend resources equal to the rectangle in attempting to negotiate the merger, convince the government to permit them to make it, or contest an antitrust case brought to prevent or undo the merger). The total social cost of the merger, at least as a first approximation, is $ABEF + BCD$ and exceeds the cost savings ($GDEF$) made possible by the merger. Of course, the curves could be drawn in such a way that the merger would generate net cost savings. My point is only that there is no presumption that anti-competitive mergers generate net savings and this, combined with the high cost of litigating issues of cost savings, provides some justification for the refusal of the courts to admit a defense of efficiencies in a Clayton Act merger case.

3. It is frequently argued that the antitrust laws should not concern themselves with practices that are merely methods of price discrimination, since there is no basis for thinking that discrimination

increases the deadweight loss of monopoly, and it may reduce it (to zero, if discrimination is perfect). ^{33/} The conclusion may be justifiable by reference to the costs of administering antidiscrimination rules, but the basis on which it has been defended by its proponents is incorrect. Even when price discrimination is perfect, so that the deadweight loss of monopoly is zero, the total social costs of a discriminating monopoly are greater than those of a single-price monopoly. ^{34/} This is because, under perfect price discrimination, C is the entire area between the demand curve and the marginal (= average) cost curve, and must exceed the sum of L and D at any single price (see Figure 1).

4. It is widely believed that the case for antitrust enforcement has been gravely weakened by the theory of the second best. The theory teaches that the elimination of one monopoly in an economy containing other monopolies (or other sources of divergence between price and marginal cost, such as taxation) may reduce the efficiency of resource allocation; hence antitrust enforcement may increase rather than reduce D. The true economic basis for antitrust enforcement, however, is not D but $D + L$, and we have seen that under plausible assumptions as to the elasticity of demand D is only a small fraction of $D + L$, at least for moderate increases in price above the competitive level. The social costs measured by L, like the social costs of theft, are largely unaffected by the existence of second-best problems.

33. This is a major theme in Ward S. Bowman, Jr., *Patent and Antitrust Law* (1973).

34. I abstract from the costs of administering the price-discrimination scheme; these increase the costs of discriminating monopoly relative to those of a nondiscriminating monopoly.

This discussion implies that monopolizing would be a socially undesirable activity even if D were zero, and this is correct. Criminal or socially undesirable activity may in fact be defined as activity that does not add to the social wealth, but merely transfers it involuntarily. ^{35/} Monopoly, like theft, is unproductive of anything other than involuntary wealth transfers, and hence any resources devoted to obtaining monopoly are socially wasted. ^{36/}

5. The analysis in this paper suggests a possible explanation for the positive correlation that has been found between concentration and advertising (I assume, without much conviction, that the studies finding such a correlation will withstand the assaults of their critics). ^{37/} It may be easier to collude on price than on the amount of advertising. Although there is no great trick to establishing an agreed-upon level of advertising and detecting departures from it, the incentives to violate any such agreement are strong because the gains from a successful advertising campaign may be difficult to offset immediately and hence offer promise of a more durable advantage than a price cut would. If so, the situation is similar to nonprice competition in the airline industry. Price is fixed by the cartel but the level of advertising is not, or at least not effectively. As a result, the monopoly profits generated by the cartel price are transformed into additional expenditures on

35. Some pure wealth transfers may, of course, be socially desirable. But presumably public redistribution is both more efficient and more equitable than the redistribution brought about by criminal activity.

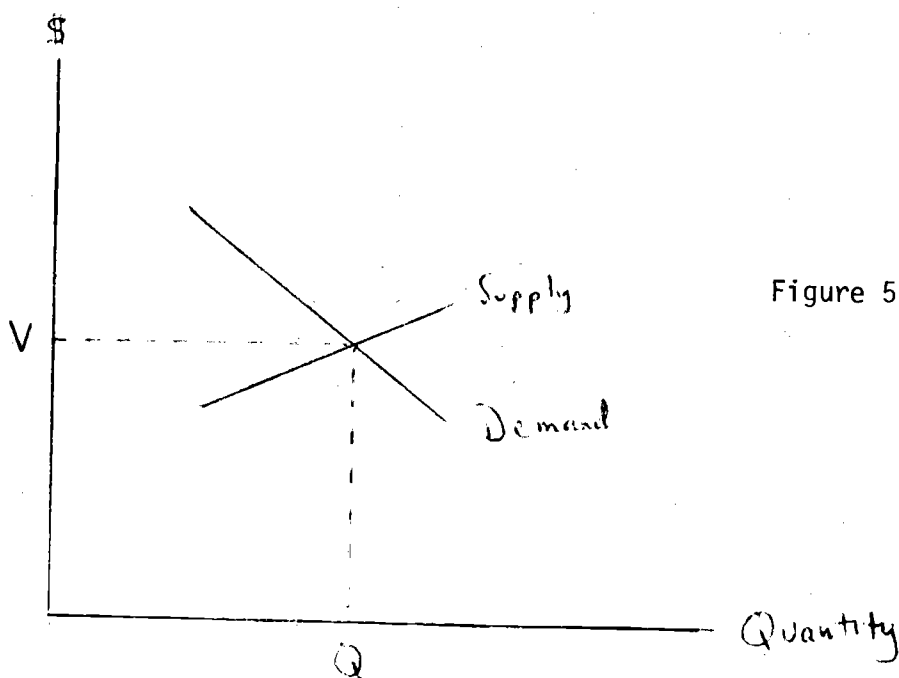
36. Subject to the qualification stated earlier in the discussion of patent monopolies.

37. See e.g., Robert B. Ekelund, Jr. & William P. Gramm, Advertising and Concentration: Some New Evidence, 15 Antitrust Bull. 243 (1970).

advertising. Therefore we can expect, other things being equal, to find more advertising in cartelized than in competitive industries.

6. In discussions of the "social responsibility" of large corporations, it is generally assumed that a firm (or a group of firms) having some monopoly power could decide to incur somewhat higher costs in order to discharge its social responsibilities, without courting bankruptcy. Thus, glancing back at Figure 1, even if MC rose to P_m , the firm would still be covering its costs. However, if the analysis in this paper is correct, and the expected profits of monopolizing are zero, it follows that the entire area L in Figure 1 will represent fixed costs to the firm, unless the monopoly was obtained under conditions of uncertainty. In the latter case the fixed costs will be somewhat lower, but in the former case any increase in MC will place the firm in danger of bankruptcy.

7. If one views the activity of obtaining monopolies as a competitive industry, the amount and social costs of monopoly are as shown in Figure 5.



The number of potential monopolies in the economy is plotted on the horizontal axis and the value of each monopoly on the vertical. The demand curve has a negative slope, on the theory that there is not an unlimited supply of lucrative monopolies. Our earlier assumption of constant costs is abandoned, and the supply (marginal cost) curve is shown as having a positive slope.

The intersection of the demand and supply curves determines the output of monopolies (Q). The total revenues of monopoly, VQ (equivalent to L in Figure 1), are equal to the total private (although now not the total social) costs of the activity of obtaining monopolies.^{38/} It is plain from Figure 5 that the social costs of monopoly can be reduced in any of three ways: by reducing the average value of a monopoly (as by eliminating regulatory barriers to entry); by increasing the marginal costs of obtaining a monopoly; and by reducing the elasticity of the supply of inputs into that activity.

This analysis provides a framework for analyzing a variety of public policy issues, such as the recurrent proposal to replace the present method of assigning television licenses (awarded to the applicant who convinces the FCC in a formal hearing of his superior ability to serve the "public interest") by an auction system. This proposal is frequently supported on distributive grounds--why should the licensee rather than the public receive the rents due to the limited amount of spectrum that has been allocated for television? This paper suggests

38. The social costs are the area under the supply curve to the left of Q .

a nondistributive justification for the auction proposal. The auction would substitute a transfer payment for a real cost--the expenditures on the hearing process by competing applicants. To be sure, there is a danger that the auction mechanism might be subverted by expenditures designed to rig the bidding. An applicant would be willing to expend in real resources up to the amount of the transfer that would be necessary to obtain the license, in order to obtain the license without having to make the transfer. However, such expenditures could be discouraged by appropriate legal penalties.^{39/} The objective of punishment would be to increase the expected costs of obtaining the license (other than by an honest bid), which includes any expected punishment costs, to the point where the applicants are induced to make the costless transfer rather than to expend real resources on trying to obtain the license outside of the auction process.

The patent laws embody a somewhat similar economizing technique. Were there no such laws, inventors would expend substantial resources on preserving the secrecy of their inventions; their efforts would generate indirect as well as direct social costs by retarding the spread of knowledge. The patent laws, by providing a legal remedy for

39. To be sure, the enforcement of legal penalties is not a costless activity. However, at least in principle, the private marginal cost of criminal activity can be raised to arbitrarily high levels far exceeding the expected value of that activity to the criminal, at negligible cost, by establishing a schedule of penalties that are so severe that even though only slight resources are devoted to detection and punishment of criminals the expected punishment cost faced by the criminal is so large as to deter him. See Gary S. Becker, *Crime and Punishment: An Economic Approach*, 76 *J. Pol. Econ.* 169 (1968).

infringement, reduce the level of such expenditures, in much the same way as the existence of legal penalties for theft reduces the level of resources that people devote to protecting their property from thieves.

An interesting method of reducing the social costs of monopoly is used, perhaps unwittingly, by labor unions. The existence of a monopoly wage might be expected to induce the expenditure of more and more resources by workers seeking entry to the union, until the expected benefits of union membership were reduced to zero. However, unions traditionally have rationed membership in a way that greatly reduces the marginal benefits of expenditures on obtaining membership, and hence the resources expended in that pursuit, by conditioning membership on a status that is difficult or impossible for the job seeker to buy at any price--such as being white, or the son of a union member, or an apprentice for seven years.^{40/} In the limit, this method of rationing would reduce the elasticity of the supply of inputs into obtaining union membership, and hence the social costs of labor monopolies excluding deadweight loss, to zero. Thus, the costs of labor monopolies may be lower than the costs of other types of monopoly, even for the equivalent price increases and scale of activity, because other markets in monopoly have not used such efficient (!) methods of rationing monopoly power.

40. The use of these methods by unions is being increasingly limited by government regulations designed to eliminate racial discrimination.

8. One reason why most students of tax policy prefer income to excise taxes is that the misallocative effect of an income tax is assumed to be less than that of an excise tax: the cross-elasticity of demand between work and leisure is assumed to be lower than that between a commodity and its substitutes. Even if this is correct, it does not follow that the total social costs of collecting a given amount of revenue by means of an income tax are lower than those of an excise tax. The amount of the tax transfer-- T in Figure 6--represents potential gain to the taxpayer, and he will expend real

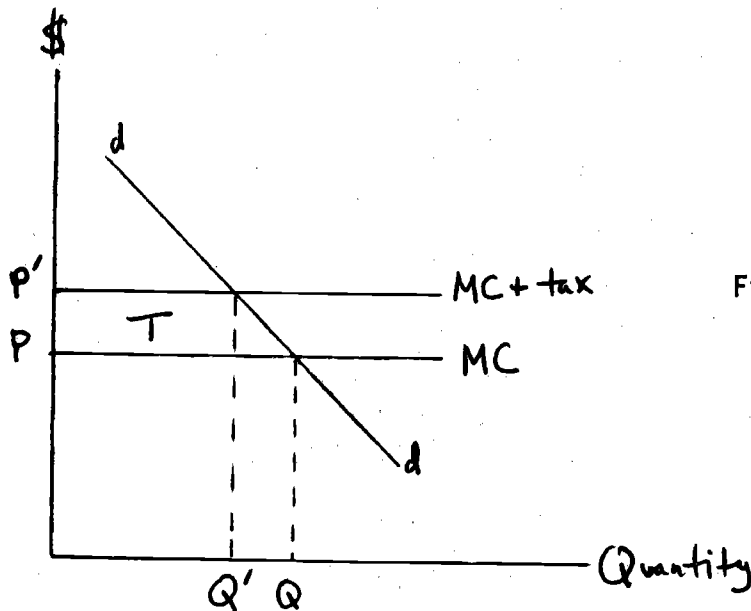
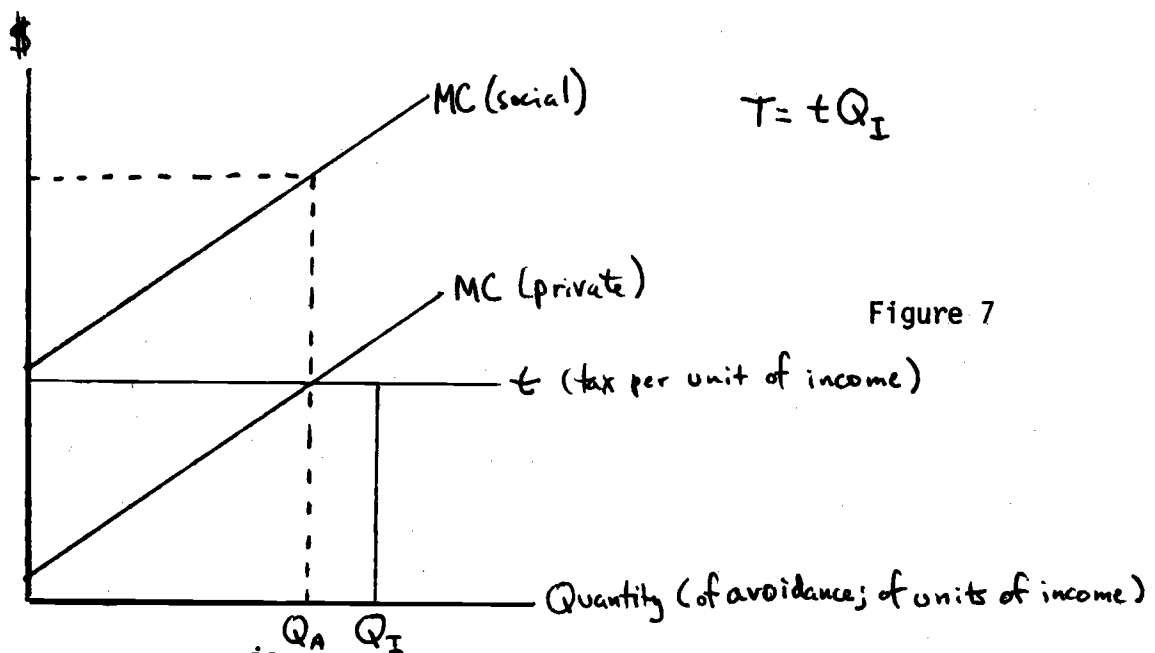


Figure 6

resources on trying to avoid the tax until, at the margin, cost and gain are equated. The critical question in comparing the costs of income and excise taxation, therefore, is the shape and location of the marginal cost curves for avoiding income tax liability and excise tax liability, respectively. In the case of a highly progressive income tax system in which expenses for the production of income are

deductible, the comparison is likely to be distinctly unfavorable to income taxation. Imagine that the marginal income tax rate in the highest bracket is 90 per cent (as it once was in this country). Then the taxpayer will continue expending resources on tax avoidance until the expected value of a dollar so expended falls below ten cents; conceivably, he might spend as much as ten times his tax liability in order to reduce that liability to zero. How much he would actually spend would depend on the location and shape of the supply curve for avoidance, but, as shown in Figure 7, his expenditures could easily exceed his tax liability (the T of Figure 6).



This analysis^{is} not conclusive against the income tax. It might be possible to increase the private marginal costs of avoidance by punishment or by disallowing expenses on tax avoidance. The main

problem would be to distinguish legitimate from illegitimate avoidance efforts. ^{41/} The point remains that no general presumption that excise taxation is less costly than income taxation can be derived from an analysis limited to the allocative costs of taxation, corresponding to the deadweight losses of monopoly.

41. One would hardly want to punish everyone who believed that some provision of the Internal Revenue Code was not intended to apply to his activity.