The Social Costs of Monopoly and Regulation

Richard A. Posner

University of Chicago Law School and National Bureau of Economic Research

This paper presents a model and some highly tentative empirical estimates of the social costs of monopoly and monopoly-inducing regulation in the United States. Unlike the previous studies, it assumes that competition to obtain a monopoly results in the transformation of expected monopoly profits into social costs. A major conclusion is that public regulation is probably a larger source of social costs than private monopoly. The implications of the analysis for several public policy issues, such as appropriate policy toward mergers and price discrimination, are also discussed.

When market price rises above the competitive level, consumers who continue to purchase the sellers' product at the new, higher price suffer a loss (L in fig. 1) exactly offset by the additional revenue that the sellers obtain at the higher price. Those who stop buying the product suffer a loss (D) not offset by any gain to the sellers. This is the "deadweight loss" from supracompetitive pricing and in traditional analysis its only social cost, L being regarded merely as a transfer from consumers to producers. Loss D, however, 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 (Tullock 1967). Theft provides an instructive analogy. The transfer of wealth from victim to

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. I am grateful to William F. Baxter, Gary S. Becker, Harold Demsetz, Victor R. Fuchs, William M. Landes, Sam Peltzman, and George J. Stigler for helpful comments on previous drafts of the paper.

[*Journal of Political Economy*, 1975, vol. 83, no. 4]

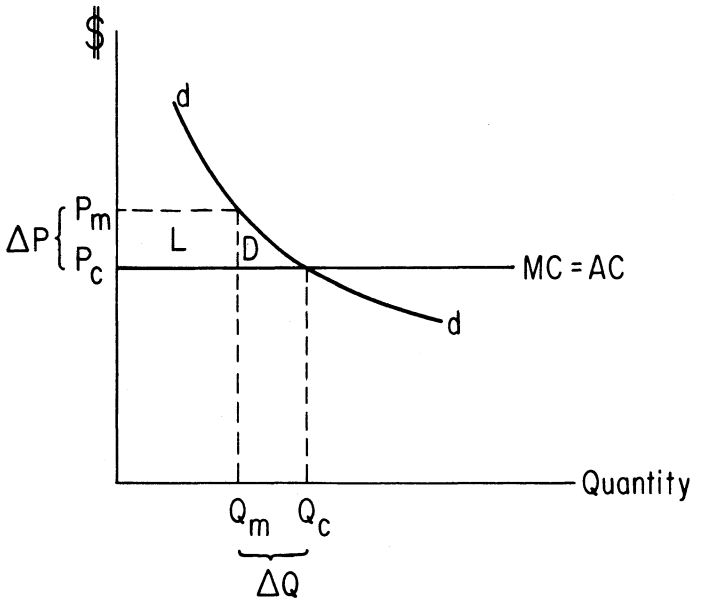

FIG. 1.—Social costs of supracompetitive pricing

thief involves no artificial limitation of output,[1] but it does not follow that the social cost of theft is zero. The opportunity for 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 (Tullock 1967; Becker 1968, p. 171, n. 3).

This sort of analysis has long been familiar in a few special contexts. Plant's criticism of the patent system, made more than 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 (Plant 1934). Telser's theory of resale price maintenance is in the same vein (Telser 1960), as is the literature on nonprice competition among members of a cartel (Stigler 1968, pp. 23–28; Douglas and Miller 1974). But, while the tendency of monopoly rents to be transformed into costs is 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 for the most part to be ignored. The present paper is an effort to rectify this neglect.[2]

[1] If a thief took three radios from a home and on the way out dropped one, which broke, the resulting loss would correspond to the deadweight loss of monopoly.

[2] See Krueger (1974) for a parallel approach to the measurement of the social costs of import licenses in India and Turkey.

Part I 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 II uses the model to estimate the social costs of monopoly in the United States and the social benefits of antitrust enforcement. The estimates are crude; their primary value may simply be to induce skepticism about the existing empirical literature on the social costs of monopoly. Part III considers the implications of the analysis for several qualitative issues relating to monopoly and public regulation.

I. A Model of the Social Costs of Monopoly

A. Assumptions

The critical assumptions underlying the model are the following:

1. Obtaining a monopoly is itself a competitive activity, so that, at the margin, the cost of obtaining a monopoly is exactly equal to the expected profit of being a monopolist. An important corollary of this assumption is that there are no intramarginal monopolies—no cases, that is, where the expected profits of monopoly exceed the total supply price of the inputs used to obtain the monopoly. If there were such an excess, competition in the activity of obtaining the monopoly would induce the competing firms (or new entrants) to hire additional inputs in an effort to engross the additional monopoly profits.

2. The long-run supply of all inputs used in obtaining monopolies is perfectly elastic. Hence, the total supply price of these inputs includes no rents.

3. The costs incurred in obtaining a monopoly have no socially valuable by-products.

The first two assumptions assure that all expected monopoly rents are transformed into social costs, and the third that these costs do not generate any social benefits.[3] But how reasonable are such assumptions?

1. The first is a standard assumption of economics and, pending better evidence than we have, seems a reasonable one in the present context. Anyone can try to obtain a patent, a certificate of public convenience and necessity, a television license, a tariff, an import quota, or a minimum-wage law; and anyone can try to form a cartel with his competitors or, if he is a member of a cartelized industry, try to engross a greater share of the monopoly profits of the industry.[4] Nonprice competition in the airline

[3] Another assumption, but one that 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 fig. 1.

[4] Other than by reducing price, a method of obtaining a larger share of the cartel's profits that would not involve a socially wasteful use of resources.

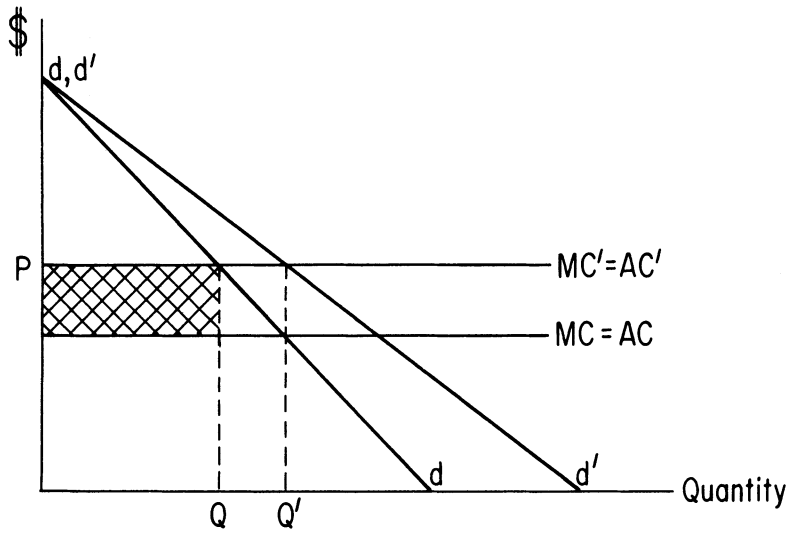

FIG. 2.—Nonprice competition when market price exceeds competitive level

industry illustrates the last point. If the Civil Aeronautics Board places a floor under airline prices that exceeds the marginal cost of providing air transportation under competitive conditions, the situation initially is as depicted in figure 2 and is unstable. Since 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 fig. 2) where the industry is earning only a normal return (see Douglas and Miller 1974). The result will be the transformation of the monopoly profits initially generated by the regulatory price floor—the shaded rectangle—into higher costs for the industry. The demand curve shifts to the right because the increased expenditures on service improve the product from the standpoint of the consumer. But the additional consumer surplus is not great enough to offset the higher costs—otherwise the higher level of service would have been provided without the spur of monopoly pricing.

If nonprice competition were forbidden (say, at zero cost) or were somehow not feasible, it would not follow that our assumption that monopolizing is a competitive activity would be overthrown. It would mean simply that the expected profits of the airline business would be greater than if the airlines could expect those profits to be dissipated in nonprice competition. Hence, more resources would be devoted to obtaining a license from the CAB in the first place. The expected profits from monopoly pricing of air transportation would still be zero.

2. Although the assumption that obtaining monopolies involves constant costs seems plausible as a first approximation—there seems little

reason to think that it involves using resources whose long-run supply is inelastic—a more important point is that the assumption may not be a crucial one. Assume that suppliers of inputs into monopolizing do obtain rents. In the long run, the availability of such rents will attract additional resources into the production of those inputs, and these resources will be wasted from a social standpoint. Some possible exceptions are considered in part III(7). Clearly, however, the production function of monopolies requires greater attention than I give it in this paper. The assumption of a perfectly elastic long-run supply may fail for an input as foreign to conventional economic analysis as political power.

3. In the airline example, the expenditures on monopolizing had a socially valuable by-product (improved service), although the value was less than its cost. However, the possibility that expenditures on monopolizing will yield such by-products will be ignored in the development of the model, and its principal relevance, therefore, is to methods of monopolizing that have little or no social value. The formation of a cartel, the procuring of a tariff or other protective legislation, and the merging of competing firms in a market to produce a monopoly (where the merger does not enable economies of scale or other efficiencies to be realized) are examples of such methods. (Even in these cases, there will be some socially valuable by-products [e.g., information] if, for example, the cartel agreement fails to limit nonprice competition.) At the opposite extreme, obtaining a monopoly by cutting costs or prices or by innovation will normally yield social benefits greater than the expenditures on monopolizing.

Several more preliminary points should be noted briefly.

1. Legal and illegal monopolies must be distinguished. The threat of punishment can be used to increase the expected costs of monopolizing and thereby reduce the amount of resources invested in the activity. To the extent that enforcers' resources are merely substituted for monopolizers', there will be no social savings (see Becker 1971, p. 101); but the literature on punishment (e.g., Becker 1968) suggests that activities such as monopolizing can be deterred at low social cost by combining heavy monetary penalties (i.e., transfer payments) with modest resources devoted to apprehending and convicting offenders.[5] Hence, under an optimum system of penalties, the social costs of *illegal* monopolies might be quite low.

2. As an extension of the last point, note that the observed monopoly profits in an industry may actually underestimate the social costs of monopoly in that industry. Considerable resources may have been

[5] This could, to be sure, merely shift the problem to a new level: the opportunity to obtain substantial rents from apprehending and convicting monopolists will induce enforcers to pour resources into enforcement activities. This problem is analyzed in Landes and Posner (1975).

expended by consumers or enforcers to reduce those profits. Monopoly profits in an industry could be zero, yet the social costs of monopoly in that industry very high, if enforcement of antimonopoly measures were both expensive and effective.

3. Given uncertainty, the expected monopoly profits of any firm seeking a monopoly may be much smaller than the actual monopoly profits, and so will its expenditures. If 10 firms are vying for a monopoly having a present value of $1 million, and each of them has an equal chance of obtaining it and is risk neutral, each will spend $100,000 (assuming constant costs) on trying to obtain the monopoly. Only one 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. If the market for monopoly is in fact characterized by a high degree of uncertainty, this would explain why the costs of obtaining monopoly have largely eluded detection. 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.

4. 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, since a bribe is a pure transfer. In fact, however, 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 (Krueger 1974, pp. 292–93).

B. The Model

Given the assumptions explained above, the total social costs of monopoly prices in figure 1 are simply $D + L$, and since $D \simeq \frac{1}{2}\Delta P \Delta Q$ and $L = \Delta P(Q_c - \Delta Q)$, the relative sizes of D and L are given by

$$\frac{D}{L} \simeq \frac{\Delta Q}{2(Q_c - \Delta Q)}. \tag{1}$$

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

$$\frac{D}{L} \simeq \frac{p}{2(1/\varepsilon - p)}. \tag{2}$$

The partial derivatives are

$$\begin{aligned} \frac{\partial(D/L)}{\partial \varepsilon} &\simeq \frac{2p}{(2 - 2p\varepsilon)^2} > 0; \\ \frac{\partial(D/L)}{\partial p} &\simeq \frac{2\varepsilon}{(2 - 2p\varepsilon)^2} > 0. \end{aligned} \tag{3}$$

In words, the ratio of D to L is smaller, the less elastic the demand for the industry's product at the competitive price and the smaller the percentage price increase over the competitive level. At moderate elasticities and percentage price increases, D is only a small fraction of L (and hence of the total costs of monopoly). For example, at an elasticity of one[6] and a price increase over the competitive level of 10 percent, D is only 5.6 percent of L.

Observe that the model does *not* assume that the actual supracompetitive price being charged (P_m in fig. 1) is the optimum monopoly price for the industry (otherwise the supracompetitive price increase would not be determined independently of the elasticity of demand, as in [2]). The rationale of this procedure is that perfect monopoly is presumably rare; it will, however, be considered as a special case later.

Using R_c to denote total sales revenues at the competitive price, C, the total social costs of monopoly, is approximated by

$$D + L = pR_c - \tfrac{1}{2}\Delta P \Delta Q \tag{4a}$$

$$= R_c(p - \tfrac{1}{2}\varepsilon p^2). \tag{4b}$$

The partial derivatives of C are (approximately)

$$\frac{\partial C}{\partial R_c} = p - \tfrac{1}{2}\varepsilon p^2 > 0 \text{ iff } \varepsilon p < 2;$$

$$\frac{\partial C}{\partial p} = R_c(1 - \varepsilon p) > 0 \text{ iff } \varepsilon p < 1; \tag{5}$$

$$\frac{\partial C}{\partial \varepsilon} = -\tfrac{1}{2}p^2 R_c < 0.$$

In words, the social costs of monopoly will usually—not always—be higher, the larger the industry's sales revenues at the competitive price and output and the greater the percentage price increase over the competitive level. And they will always be higher, the less elastic the demand for the product at the competitive price—the costs of monopoly being greatest when demand is totally inelastic at the competitive price.

Formulas (2) and (4b) are accurate only for small changes in the price level. Yet monopolization might result in large price increases. Hence (1) and (4a) remain useful. For purposes of empirical estimation, it is helpful to derive two additional formulas: one for the case where data on the deadweight loss, the elasticity of demand, and the monopoly price increase are available and the elasticity of demand is assumed to be constant, and the other for the case where data on the monopoly price increase, the monopoly output, and the elasticity of demand at the

[6] Throughout this paper, ΔQ is treated as a positive number. Therefore, ε [$= (\Delta Q/\Delta P)/(Q/P)$] is also positive.

monopoly price are available and the demand curve is assumed to be linear.

1. For the case of constant elasticity, let $k \equiv P_c/P_m$ and $R_m \equiv$ total sales revenue at the monopoly price and output. Then, since $Q_c = \alpha P_c^{-\varepsilon}$ and $Q_m = \alpha P_m^{-\varepsilon}$, and therefore $\Delta Q = \alpha(P_c^{-\varepsilon} - P_m^{-\varepsilon})$, D/L and C are approximately

$$\frac{D}{L} = \frac{(kP_m)^{-\varepsilon} - P_m^{-\varepsilon}}{2P_m^{-\varepsilon}} = \frac{k^{-\varepsilon} - 1}{2}; \tag{6}$$

$$C = D + L = D\left(1 + \frac{2}{k^{-\varepsilon} - 1}\right) = R_m(1 - k)\left(\frac{k^{-\varepsilon} + 1}{2}\right).^{7} \tag{7}$$

The partial derivatives of D/L are (approximately)

$$\begin{aligned} \frac{\partial(D/L)}{\partial k} &= \frac{-\varepsilon}{2k^{\varepsilon+1}} < 0; \\ \frac{\partial(D/L)}{\partial \varepsilon} &= \frac{-k^{-\varepsilon} \ln k}{2} > 0. \end{aligned} \tag{8}$$

[7] For the special case where the firm is able to charge the optimum monopoly price for the industry, so that $P_c = MC = P_m(1 - 1/\varepsilon)$, equation (6) becomes

$$\frac{D}{L} = \frac{(1 - 1/\varepsilon)^{-\varepsilon} - 1}{2} \tag{6'}$$

and equation (7) becomes

$$C = \frac{R_m[(1 - 1/\varepsilon)^{-\varepsilon} + 1]}{2\varepsilon}. \tag{7'}$$

Since a demand curve of constant elasticity is nonlinear, the question arises whether the linear approximation of the deadweight loss used in equations (6) and (7) (and [6'] and [7']) introduces a source of serious inaccuracy. It appears not to, at least in the simple case where $\varepsilon = 1$ and therefore

$$\frac{D}{L} = \frac{\int_{Q_m}^{Q_c} P\,dQ - P_c\,\Delta Q}{(P_m - P_c)Q_m} = \frac{\ln(1/k) - 1 + k}{1 - k}. \tag{6''}$$

Table 1, which compares D/L as calculated from equation (6) (with $\varepsilon = 1$) and from equation (6''), shows that the linear approximation overestimates the deadweight loss, but not seriously.

TABLE 1

P^* (%)	D/L†	
	Eq. (6)	Eq. (6'')
5	.025	.025
10	.050	.049
15	.075	.072
20	.100	.094
50	.250	.216

* Monopoly price increase.
† Ratio of deadweight to additional loss.

In words, 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) and greater, the more elastic the demand.

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 the demand curve can be approximated by a straight line, we begin by determining the slope of the demand curve at the monopoly price:

$$\frac{\Delta Q}{\Delta P} = \frac{\varepsilon Q_m}{P_m}. \tag{9}$$

Since the slope of a linear demand curve is constant, this equation can be used to find ΔQ and hence C and D/L:

$$C = R_m(1 - k)[1 + \tfrac{1}{2}\varepsilon(1 - k)]; \tag{10}$$

$$\frac{D}{L} = \frac{\varepsilon(1 - k)}{2}.^{8} \tag{11}$$

The estimates produced by our two formulas for the ratio of the deadweight to the additional loss from monopoly—equations (6) and (11)—turn out not to be very different for price increases of less than 25 percent, and even for much larger price increases if the elasticity of demand is no greater than one (see fig. 3).

II. Empirical Estimates

The formulas developed in the preceding part can be used to derive, from the estimates of the deadweight loss of monopoly made by Arnold Harberger and others, an estimate of the total social cost of monopoly. Harberger (1954), estimating an average monopoly price increase of about 6 percent and assuming that the elasticity of demand was constant and equal to unity, found the deadweight loss from monopoly in the manufacturing sector to be equal to (at most) 0.1 percent of GNP. Harberger's (implicit) k is 0.9434, and from equation (6) the ratio of D to L in Harberger's analysis is, therefore, 0.03. Hence, if D is 0.1

[8] In the special case where the firm is able to charge the optimum monopoly price,

$$C = \frac{R_m}{2\varepsilon}; \tag{10'}$$

$$\frac{D}{L} = \frac{1}{2}. \tag{11'}$$

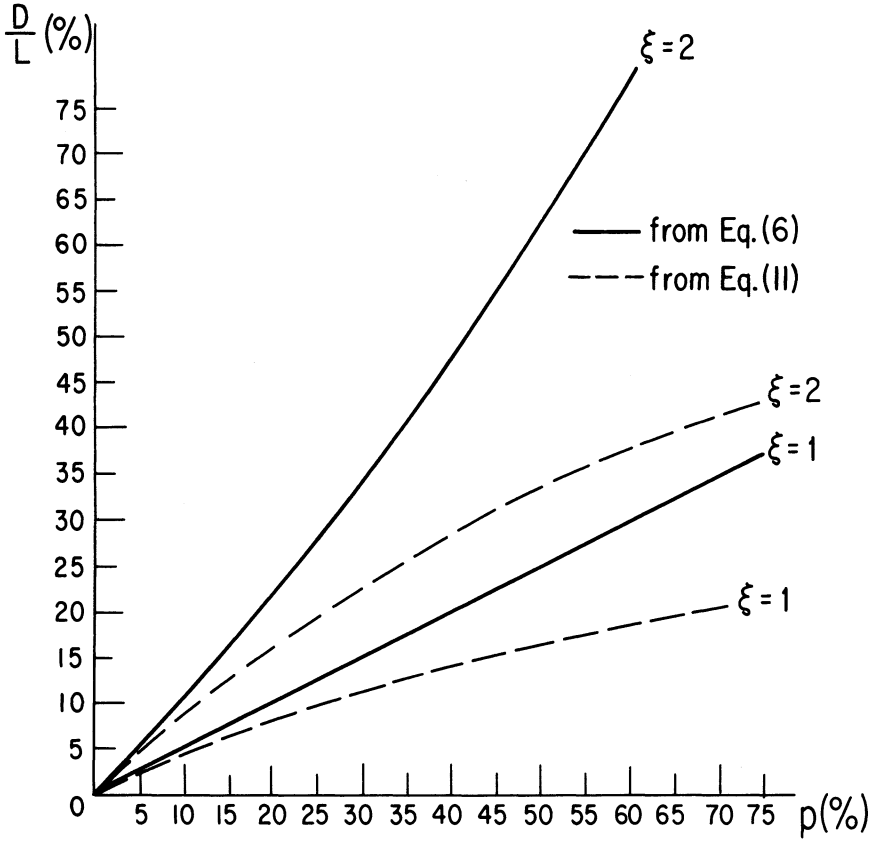

FIG. 3.—Ratio of deadweight to additional loss of monopoly, for different price increases and demand elasticities.

percent of GNP, L is about 3.3 percent and C about 3.4 percent of GNP. Schwartzman (1960) used similar methods and found D equal to about 0.1 percent of GNP too. But he assumed a price increase of 8.3 percent and an elasticity of demand of 1.5. Plugging these values into equation (6) yields $D/L = 0.06$. Hence, if $D = 0.1$ percent of GNP, $L = 1.7$ percent and $C = 1.8$ percent.

Neither estimate can be given much credence, however, because of the method that both Harberger and Schwartzman employed to determine the monopoly price increase. Persistently above average rates of return were used both (1) to identify the monopolized industries and (2) to calculate the monopoly price increase. If the approach of this paper is correct, such a procedure is improper, especially the second step. Because of uncertainty, many monopolists may enjoy supernormal rates of return ex post, but those rates will understate the percentage of the monopolist's

revenues that is attributable to monopoly pricing, unless no cost whatever was incurred in obtaining (or maintaining) the monopoly.[9]

A better method of calculating the social costs of monopoly (deadweight plus additional loss) is to obtain from industry studies estimates of the monopoly price increase and of the elasticity of demand at the relevant points along the demand curve. An independent estimate of the elasticity of demand would 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 (computed from Caves 1962, p. 372; Jordan 1970, pp. 110–11, 124–25; and Yale Law Journal 1965, pp. 1435–36). If a 66 percent price increase 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,[10] to be 2.5 at the monopoly price. An independent estimate of the long-run elasticity of demand for air travel made by Houthakker and Taylor (1966, p. 124) is 2.36,[11] which is virtually identical to my calculation.

If we assume a constant elasticity of 2.5 and solve for D/L using equation (6′), $D = 1.29L$, and (from equation [7′]) it is readily calculable that the total social cost of the airline monopoly is equal to 92 percent of the total revenue of the industry 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. If, therefore, equations (10′) and (11′) are used instead of (6′) and (7′), $D = 0.5L$ and $C = 0.2R_m$—still a very large social loss from the regulation-induced airline monopoly. (These estimates ignore, however, the partially offsetting benefits of excessive nonprice competition in the airline industry.)

[9] This point is distinct from the (also valid) objections to Harberger's procedure raised by Stigler (1956)—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.

[10] This was essentially the procedure used by Kamerschen (1966) to estimate the deadweight loss from monopoly in manufacturing. He has been criticized, rightly, 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. The assumption is more plausible with regard to a regulated industry in which entry and price competition are limited by the regulatory agency and the Sherman Act is inapplicable.

[11] This is presumably the elasticity of demand at the regulated price, since only a small part of the airline industry is exempt from CAB regulation.

TABLE 2

SOCIAL COSTS OF REGULATION

INDUSTRY	REGULATORY PRICE INCREASE (%)	ELASTICITY		COSTS (AS % OF INDUSTRY'S SALES)	
		ε_1	ε_2	C_1*	C_2*
Physicians' services	.40†	3.500	0.575‡	.14	.31
Eyeglasses	.34§	0.394	0.450‖	.13	.24
Milk	.11#	10.000	0.339**	.05	.10
Motor carriers	.62††	2.630	1.140‡‡	.19	.30
Oil	.65§§	2.500	0.900§§	.20	.32
Airlines	.66	2.500	2.360	.20	.19

* C_1 based on ε_1; C_2 based on ε_2.
† Kessel 1972, p. 119.
‡ Houthakker and Taylor 1966, p. 99 (short run).
§ Benham 1973, p. 19.
‖ Benham 1973, p. 30 (simple average).
Kessel 1967, p. 73.
** Houthakker 1965, p. 286. This estimate is for all food; an estimate limited to dairy products in the Netherlands was not significantly different (Ayaynian 1969).
†† Average estimates in Department of Agriculture studies cited in Moore (1972) and Farmer (1964).
‡‡ Simple averages of various estimates for transportation in Scandinavia (see Frisch 1959 and Parks 1969, p. 649).
§§ Cabinet Task Force on Oil Import Control 1970.

All of the previous studies of the cost of monopoly to the economy have been based on supposed monopoly pricing in manufacturing alone. Yet 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. 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; the other (ε_2) is an independent estimate of elasticity. The estimates of the total social costs of the regulation in question (C_1, where ε_1 is the estimate of elasticity used, and C_2, where ε_2 is used) 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.

These estimates are, of course, very crude, but they do suggest that the total costs of regulation may be extremely high, given that about 17 percent of GNP originates in 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.[12] Indeed, the costs of

[12] Of course, 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 Interstate Commerce Commission). On the other hand, tariffs and similar restrictions (e.g., the oil import quota) are excluded from the estimate of the percentage of GNP affected by regulation.

regulation probably exceed the costs of private monopoly. To be sure, a higher percentage of GNP—30 percent—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 one-fifth of the output of this sector comes from industries in which four firms account for 60 percent 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.[13] Not all violations of the Sherman Act are detected and punished, but the secret conspiracies that escape detection are probably not very effective—even the great electrical conspiracy, an elaborate and relatively durable conspiracy among a very small group of firms, apparently succeeded in raising prices by less than 10 percent on average (see U.S. Congress 1965, p. 39). It would be surprising if the price level of the manufacturing and mining sector as a whole were more than about 2 percent above the competitive level.[14] Assume that it is 2 percent, and that the average elasticity of demand for the products of this sector, at current prices, is 1.1607.[15] Then the total social costs of monopoly in this sector are 1.9 percent of the total revenues generated in the sector (from equation [10]). This amounts to a total dollar loss substantially smaller than that generated in the regulated sector.[16] And this is true even if we assume that prices in the manufacturing and mining sector are, on average, 4 percent above the competitive level, rather than 2 percent.[17]

This comparison excludes, of course, both the relative costs of regulation

[13] Thus, Kessel's study of underwriting costs (1971, p. 723) shows that an increase beyond eight in the number of bids does not reduce those costs substantially—and an industry where the four largest firms have less than 60 percent of the market is apt to contain at least eight significant competitors.

[14] If we assume that only in industries where the four-firm concentration ratio exceeds 60 percent is effective, undetected collusion likely, and that collusion allows these industries to maintain prices, on average, 5 percent above the competitive level while in the rest of the manufacturing and mining sector the average price level is only 1 percent above the competitive level, then average prices for the entire sector would be only 1.83 percent 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 Manufactures.)

[15] This figure is a simple average of the long-run price elasticities for nine product groups within the manufacturing and mining sector estimated in Houthakker and Taylor (1966, pp. 72, 74, 83, 112–14, 116, 128–31).

[16] The simple average of the social-cost estimates presented in table 2 is 19.8 percent of the total revenues of the regulated industry. Assuming that 50 percent 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 19.8 percent of total revenue, the social costs of regulation would be equal to 1.7 percent of GNP, while the social costs of monopoly in manufacturing and mining would be equal to 0.6 percent of GNP.

[17] In which event the social costs of monopoly in that sector would be about 1.2 percent of GNP.

TABLE 3

SOCIAL COSTS OF CARTELIZATION

INDUSTRY	CARTEL PRICE INCREASE (%)	ELASTICITY		COSTS (AS % OF INDUSTRY'S SALES)	
		ε_1	ε_2	C_1	C_2
Nitrogen	0.75*	2.3256	1.4493†	.21	.30
Sugar	0.30‡	4.3276	0.3390§	.12	.22
Aluminum	1.00‖	2.0000	...	.25	...
Aluminum	0.38#	3.6311	...	.14	...
Rubber	1.00**	2.0000	...	.25	...
Electric bulbs	0.37††	3.7023	...	.14	...
Copper	0.31‡‡	4.2499	...	.12	...
Cast-iron pipe	0.39§§	3.5641	...	.14	...

* Stocking and Watkins 1946, p. 163.
† Stocking and Watkins 1946, p. 166.
‡ Stocking and Watkins 1946, p. 46.
§ Houthakker 1965, p. 286; obviously a much too low estimate for one food product sold at a cartel price!
‖ Stocking and Watkins 1946, p. 228.
Stocking and Watkins 1946, p. 251.
** Stocking and Watkins 1946, p. 64–65.
†† Stocking and Watkins 1946, p. 343.
‡‡ Stocking and Watkins 1948, p. 127.
§§ United States v. Addyston Pipe & Steel Co., 85 F. 271 (6th Cir. 1898).

and of antitrust enforcement and the relative benefits of monopoly in the two sectors.[18] Were these additional factors included, however, it is doubtful that the comparison would become more favorable to the regulated sector. In particular, while there are theoretical reasons for believing that concentration in unregulated markets is associated with economies of scale and other efficiencies (Demsetz 1973), there is no accepted theory or body of evidence that ascribes social benefits to regulation limiting entry and price competition.

The analysis developed here can also be used to estimate 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 well-organized (mainly international) private cartels.[19]

Presumably, collusive price increases of this magnitude and the attendant very substantial social costs are deterred by current enforcement of the American antitrust laws. A complete cost-benefit analysis of the antitrust laws would, however, also require estimation of (1) the costs of administering those laws[20] and (2) the large social costs imposed by the

[18] To recur to an earlier point, the assumed monopoly price increase in the manufacturing and mining sector may underestimate the social costs of monopoly in that sector. Those costs may be reflected in expenditures by consumers and enforcers in preventing monopoly pricing.

[19] As distinct from the sorts of covert conspiracies that might escape detection under present enforcement of the Sherman Act (see Stigler 1968, pp. 268–70).

[20] A point to be kept in mind is that, while these costs are incurred annually, private—unlike governmentally protected—cartels eventually collapse (although they often re-form later). Hence, table 3 gives an exaggerated picture of the *average* annual costs of cartelization as it would exist in the absence of the Sherman Act.

many perverse applications of antitrust laws that are, perhaps, an inevitable by-product of having such laws.

A very large disclaimer concerning the accuracy of the estimates presented in this part of the paper needs to be entered at this point. Quite apart from any reservations about the realism of the assumptions on which the model used to generate these estimates is based, the crudeness of the data on price increases and elasticities of demand precludes treating the estimates of the costs of the monopoly and regulation as anything more than suggestive. The suggestions are, however, interesting ones: (1) previous studies of the costs of monopoly may have grossly underestimated those costs; and (2) the costs of monopoly are quite probably much greater in the regulated than in the unregulated sector of the economy, despite the greater size of the latter sector.

III. Other Applications

1. In a recent paper Comanor and Smiley (in press) attempt to show that a large part of the inequality in the distribution of wealth in contemporary America is attributable to monopoly. They use studies such as Harberger's (1954) 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 highly doubtful. There is no reason to think 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 (1968) 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 (see fig. 4).

This analysis is incomplete, however. The expected profits of the merger (*ABEF*) will generate an equivalent amount of costs as the firms vie to make such mergers or, after they are made, to engross the profits generated by the higher postmerger price through service competition or whatever. As a first approximation, the total social cost of the merger is *ABEF* + *BCD* and exceeds the cost savings (*GDEF*) made possible by it. The curves could, of course, be drawn in such a way that the merger would generate net cost savings; the point is only that there is no presumption that anticompetitive mergers generate net savings. This consideration, together with the high cost of litigating issues of cost

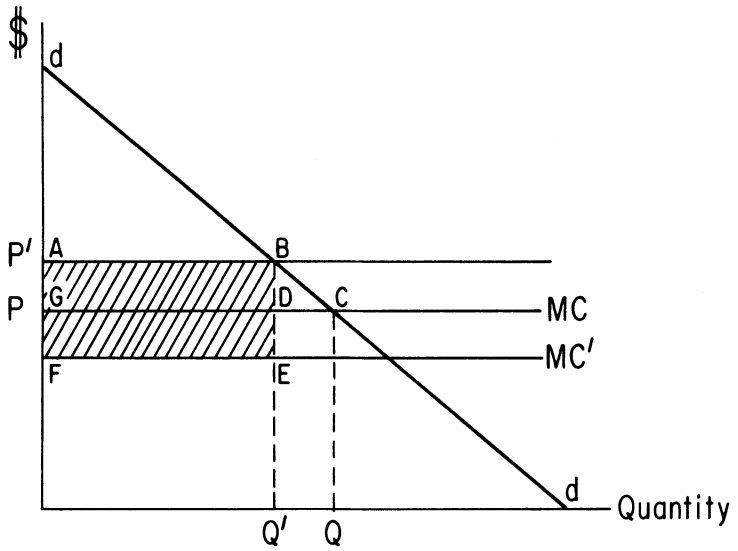

FIG. 4.—The costs of mergers

savings, may provide a justification for refusing to recognize a defense of efficiencies in merger cases where the merger is likely to produce a substantial increase in monopoly power.

3. It has been argued (e.g., Bowman 1973) 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 (it will reduce it to zero if discrimination is perfect). 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.[21] Under perfect price discrimination, C is the entire area between the demand curve and the marginal (= average) cost curve, and it is greater than $D + L$ at any single price (see fig. 1).

4. It is occasionally suggested that the case for antitrust enforcement has been gravely weakened by the theory of the second best. Since the elimination of one monopoly in an economy containing other monopolies (or other sources of divergence between price and marginal cost, such as

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

taxation) may reduce the efficiency of resource allocation, 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 (i.e., the opportunity costs of thieves' and policemen's time and of the labor and capital inputs into locks, burglar tools, etc.), are unaffected by the existence of second-best problems (cf. Markovits 1972).

5. The analysis in this paper suggests a possible explanation for the positive correlation that has been found between concentration and advertising.[22] 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. In that event the situation is similar to nonprice competition in the airline industry. If price is fixed by the cartel but the level of advertising is not, or at least not effectively, the monopoly profits generated by the cartel price will be transformed into additional expenditures on advertising. Cartelization is presumably more common in concentrated industries.

This analysis suggests, incidentally, a possible difficulty in distinguishing empirically between Telser's theory of resale price maintenance (1960) and an alternative explanation which stresses cartelization by dealers. In Telser's theory, manufacturers impose resale price maintenance in order to induce dealers to provide services in connection with the resale of the manufacturer's brand. If Telser's theory is correct, we would expect to find resale price maintenance imposed where the efficient merchandising of a product involved the provision of extensive point-of-sale services. However, a dealer's cartel might also result in the dealers' competing away the cartel profits through service competition.

6. Discussions of the "social responsibility" of large corporations generally assume that a firm (or group of firms) having some monopoly power could, without courting bankruptcy, decide to incur somewhat higher costs in order to discharge its social responsibilities. Thus, in 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

[22] The finding has been questioned, however (e.g., Ekelund and Gramm 1970).

of uncertainty. In the latter case the fixed costs will be somewhat lower, but in the former *any* increase in *MC* will jeopardize the firm's solvency.

7. Assuming that the decision to create or tolerate a monopoly has been made, it may still be possible to prevent the expected monopoly profits from being completely transformed into social costs. The basic technique is to reduce the elasticity of supply of the inputs into monopolizing. (Thus, the present discussion modifies my original assumption of perfect supply elasticity.) Consider, for example, a market that is a natural monopoly. If the monopolist is permitted to charge a monopoly price—and suppose that he is—he may set a price that exceeds the average costs of new entrants, albeit those costs are higher than his; and new entry will presumably occur. The resulting increase in the average costs of serving the market is an example of the social costs of monopoly (independent of the welfare triangle). These costs can be reduced, however, by a rule limiting entry. Such a rule will reduce the responsiveness of a key input into monopolizing—capacity to produce the monopolized product—to increases in the expected value of the monopoly. But the rule is not very satisfactory. Prospective entrants will have an incentive to expend resources on persuading the agency to change or waive the rule—and the monopolist to expend them on dissuasion. Moreover, the more efficient the rule is at keeping out new entrants at low cost to the monopolist, the greater will be the expected value of having a natural monopoly—and, hence, the greater will be the resources that firms expend on trying to become the first to occupy a natural-monopoly market.[23]

As another example, consider the recurrent proposal to replace the present method of assigning television licenses (now awarded to the applicant who convinces the Federal Communications Commission 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 generated by the limited allocation of electromagnetic spectrum for broadcasting? But there is also an efficiency justification for the proposal. The auction would substitute a transfer payment for a real cost, the expenditures on the hearing process by competing applicants. To be sure, these expenditures might simply be redirected into rigging the bidding. But this could be discouraged, possibly at low cost, by appropriate legal penalties. The objective would be to increase the expected costs of obtaining the license (other than by an honest bid), which include 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 the auction

[23] This is the obverse of the situation discussed in Demsetz (1968), where competition to become a monopolist results in a competitive price level.

process. As mentioned earlier, in an optimum system of penalties the resources expended on enforcement would be slight.

The patent laws embody a somewhat similar economizing technique. In their absence inventors would expend substantial resources on preserving the secrecy of their inventions. Their efforts in this direction would generate indirect as well as direct social costs, by retarding the spread of knowledge. By providing a legal remedy against "stealing" inventions, the patent laws 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 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 into 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.[24] 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 the welfare triangle), to zero, disregarding the costs resulting from the exclusion of possibly better qualified workers who do not meet the membership criterion. Yet even this method may not be ultimately effective in preventing the transformation of monopoly rents into social costs. The more profitable union membership is, the greater are the resources that workers will be willing to invest (e.g., in forgone earnings due to being on strike) in union-organizing activities.

8. One reason why most students of tax policy prefer income to excise taxes is that the misallocative effect of an income tax is believed 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 correct, this does not mean 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 represents potential gain to the taxpayer, and he will expend real resources on trying to avoid the tax until, at the margin, cost and gain are equated. A critical question in comparing the costs of income and excise taxation is therefore the shape and location of the supply curves for avoiding income tax liability and excise tax liability, respectively. In the case of a highly progressive income tax system in which

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

expenses for the production of income are deductible, the comparison is likely to be unfavorable to income taxation. Were the marginal income tax rate in the highest bracket 90 percent (as it once was in this country), the taxpayer would continue expending resources on tax avoidance until the expected value of a dollar so expended fell below 10 cents. Thus, he might spend as much as 10 times his marginal 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 and on his resources and attitude toward risk.) 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 the deduction of expenses on avoidance. The main problem would be to distinguish legitimate from illegitimate avoidance efforts.[25] Still, 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 loss of monopoly.

References

Ayaynian, Robert. "A Comparison of Barten's Estimated Demand Elasticities with Those Obtained Using Frisch's Method." *Econometrica* 37 (January 1969): 79–94.

Becker, Gary S. "Crime and Punishment: An Economic Approach." *J.P.E.* 76, no. 2 (March/April 1968): 169–217.

———. *Economic Theory.* New York: Knopf, 1971.

Benham, Lee. "Price Structure and Professional Control of Information." Mimeographed. Univ. Chicago Graduate School Bus. (March 1973).

Bowman, Ward S., Jr. *Patent and Antitrust Law: A Legal and Economic Appraisal.* Chicago: Univ. Chicago Press, 1973.

Cabinet Task Force on Oil Import Control. *The Oil Import Question.* Washington: Government Printing Office, 1970.

Caves, Richard E. *Air Transport and Its Regulators.* Cambridge, Mass.: Harvard Univ. Press, 1962.

Comanor, William S., and Smiley, Robert H. "Monopoly and the Distribution of Wealth." *Q.J.E.* (in press).

Demsetz, Harold. "Why Regulate Utilities?" *J. Law and Econ.* 11 (April 1968): 55–65.

———. "Industry Structure, Market Rivalry, and Public Policy." *J. Law and Econ.* 16 (April 1973): 1–9.

Douglas, George W., and Miller, James C., III. "The CAB's Domestic Passenger Fare Investigation." *Bell J. Econ. and Management Sci.* 5 (Spring 1974): 204–22.

Ekelund, Robert B., Jr., and Gramm, William P. "Advertising and Concentration: Some New Evidence." *Antitrust Bull.* 15 (Summer 1970): 243–49.

Farmer, Richard N. "The Case for Unregulated Truck Transportation." *J. Farm Econ.* 46 (May 1964): 398–409.

[25] It would make no sense to punish everyone who believed that some provision of the Internal Revenue Code was not intended to apply to his activity.

Frisch, Ragnar. "A Complete Scheme for Computing All Direct Costs and Cross Demand Elasticities in a Market with Many Sectors." *Econometrica* 27 (April 1959): 177–96.

Harberger, Arnold C. "Monopoly and Resource Allocation." *A.E.R.* 44 (May 1954): 77–87.

Houthakker, H. S. "New Evidence on Demand Elasticities." *Econometrica* 33 (April 1965): 277–88.

Houthakker, H. S., and Taylor, Lester D. *Consumer Demand in the United States, 1929–1970.* Cambridge, Mass.: Harvard Univ. Press, 1966.

Jordan, William A. *Airline Regulation in America.* Baltimore: Johns Hopkins Univ. Press, 1970.

Kamerschen, David. "Estimation of the Welfare Losses from Monopoly in the American Economy." *Western Econ. J.* 4 (Summer 1966): 221–36.

Kessel, Reuben A. "Economic Effects of Federal Regulation of Milk Markets." *J. Law and Econ.* 10 (October 1967): 51–78.

———. "A Study of the Effects of Competition in the Tax-exempt Bond Market." *J.P.E.* 79, no. 4 (July/August 1971): 706–38.

———. "Higher Education and the Nation's Health: A Review of the Carnegie Commission Report on Medical Education." *J. Law and Econ.* 15 (April 1972): 115–27.

Krueger, Anne O. "The Political Economy of the Rent-seeking Society." *A.E.R.* 64 (June 1974): 291–303.

Landes, William M., and Posner, Richard A. "The Private Enforcement of Law." *J. Legal Studies* 5 (January 1975): 1–46.

Markovits, Richard S. "Fixed Input (Investment) Competition and the Variability of Fixed Inputs (Investment): Their Nature, Determinants, and Significance." *Stanford Law Rev.* 24 (February 1972): 507–30.

Moore, Thomas Gale. *Freight Transportation Regulation.* Washington: American Enterprise Inst., 1972.

Parks, Richard W. "Systems of Demand Equations: An Empirical Comparison of Alternative Functional Forms." *Econometrica* 37 (October 1969): 629–50.

Plant, Arnold. "The Economic Theory Concerning Patents." *Economica* 1 (n.s.) (February 1934): 30–51.

Schwartzman, David. "The Burden of Monopoly." *J.P.E.* 68, no. 6 (November/December 1960): 627–30.

Stigler, George J. "The Statistics of Monopoly and Mergers." *J.P.E.* 64, no. 1 (January/February 1956): 33–40.

———. *The Organization of Industry.* Homewood, Ill.: Irwin, 1968.

Stocking, George W., and Watkins, Myron W. *Cartels in Action.* New York: Twentieth Century Fund, 1946.

———. *Cartels or Competition?* New York: Twentieth Century Fund, 1948.

Telser, Lester. "Why Should Manufacturers Want Fair Trade?" *J. Law and Econ.* 3 (October 1960): 86–105.

Tullock, Gordon. "The Welfare Costs of Tariffs, Monopolies, and Theft." *Western Econ. J.* 5 (June 1967): 224–32.

U.S. Congress, Joint Committee on Internal Revenue Taxation. *Staff Study of Income Tax Treatment of Treble Damage Payments under the Antitrust Laws.* Washington: Government Printing Office, 1965.

Williamson, Oliver E. "Economics as an Antitrust Defense: The Welfare Tradeoffs." *A.E.R.* 58 (March 1968): 18–36.

Yale Law Journal. "Is Regulation Necessary? California Air Transportation and National Regulatory Policy." *Yale Law J.* 74 (July 1965): 1416–47.