

Natural Monopoly and Its Regulation

Richard A. Posner*

A firm that is the only seller of a product or service having no close substitutes is said to enjoy a monopoly.¹ Monopoly is an important concept to this Article but even more important is the related but somewhat less familiar concept of "natural monopoly." The term does not refer to the actual number of sellers in a market but to the relationship between demand and the technology of supply. If the entire demand within a relevant market can be satisfied at lowest cost by one firm rather than by two or more, the market is a natural monopoly, whatever the actual number of firms in it. If such a market contains more than one firm, either the firms will quickly shake down to one through mergers or failures, or production will continue to consume more resources than necessary. In the first case competition is short-lived and in the second it produces inefficient results. Competition is thus not a viable regulatory mechanism under conditions of natural monopoly. Hence, it is said, direct controls are necessary to ensure satisfactory performance: controls over profits, specific rates, quality of service, extensions and abandonments of service and plant, even permission whether to enter the business at all. This set of controls has been applied mainly to gas, water, and electric power companies, where it is known as "public utility regulation," and to providers of public transportation and telecommunications, where it is known as "common carrier regulation." (I shall use "regulation" or "public utility regulation" to refer to both.) The question that this Article addresses is whether natural monopoly provides an adequate justification for the imposition of these regulatory controls.²

* A.B. 1959, Yale University; LL.B. 1962, Harvard University. Associate Professor of Law, Stanford University. I wish to thank all who read and criticized earlier drafts of this Article; and to acknowledge a special debt to four with whom discussion of the issues examined herein has greatly clarified and enlarged my own thinking—William F. Baxter, Aaron Director, Leland L. Johnson, and Leonard M. Ross.

1. Throughout this Article, the terms "monopoly" and "monopolistic" will be used to refer to single-firm monopoly, rather than in the more familiar current sense in which any market that is not perfectly competitive may be said to have monopolistic elements. One should note that the market need not be nationwide. A product or service can be effectively monopolized although provided by different firms in different areas of the country, if buyers in one area are prevented by transportation or other barriers from shopping among the firms.

2. The reader may question whether natural monopoly has much to do with regulation of the transportation industries. Even in the case of the railroads, the initial regulatory thrust, at least at the federal level, was to reduce competition among the regulated firms; proponents of regulation charged that there was too much competition rather than too little. See Hilton, *The Consistency of the Interstate Commerce Act*, 9 J. LAW & ECON. 87 (1966). See generally G. KOLKO, *RAILROADS AND REGULATION 1877-1916* (1965). This theme is even clearer in the regulation of inland-water carriers, airlines, and motor carriers. See C. FULDA, *COMPETITION IN THE REGULATED INDUSTRIES: TRANSPORTATION* 12,

A critical examination of this question seems timely. The terms "public utility" and "common carrier" may have rather an antique ring, but they also have important contemporary applications. The regulated industries provide the essential infrastructure of modern industrial society. They are also on the frontiers of technological progress. The principal civilian use of nuclear energy has been electrical generation, the principal commercial application of space technology satellite communications; both are regulated services. We are also witnessing the emergence of immensely promising industries, such as cable television, that may have sufficient natural monopoly characteristics to invite extension of the regulatory principle to them. And it is even intimated that the extension of price controls to the economy at large must be seriously considered.³

As a perusal of the citations in this Article will disclose, the 1960's have seen an upsurge of scholarly interest in the regulatory field after many years of comparative neglect. The Brookings Institution is supporting an ambitious program of study in the field. Several high-level federal policy groups, including the President's Task Force on Communications Policy⁴ and the Cabinet Committee on Price Stability, have recently addressed particular aspects of regulation. But what has been lacking thus far is an attempt to evaluate its basic soundness. Much criticized in the details of its application, regulation is assumed by nearly all who work or write in the field, as by the public in general, to be fundamentally inevitable, wise, and necessary. However, personal experience as a government lawyer involved in regulatory matters made me skeptical about the validity of the assumption and this study has convinced me that in fact public utility regulation is probably not a useful exertion of governmental powers; that its benefits cannot be shown to outweigh its costs; and that even in markets where efficiency dictates monopoly we might do better to allow natural economic forces to determine business conduct and performance subject only to the constraints of antitrust policy. I would stress, however, that no general chal-

16, 20-21 (1961); L. KEYES, *FEDERAL CONTROL OF ENTRY INTO AIR TRANSPORTATION* 83, 85 (1951); 71 YALE L.J. 307, 308-09 (1961). But in all of these instances, prominent among the conditions alleged to justify regulation were those conventionally associated with tendencies to natural monopoly: excess capacity, price discrimination, and "ruinous" price wars. See, e.g., *Coordination of Motor Transportation*, 182 I.C.C. 263, 362 (1932); L. KEYES, *supra* at 90-92, 103-04. In some instances, to be sure—trucking is a good example—the allegation of natural monopoly is preposterous. One reason for regulating trucking, however, was to protect a discriminatory pattern of railroad pricing that had arisen in the era when the railroad industry had pronounced natural monopoly features. See note 121 *infra* and accompanying text. Natural monopoly is thus a basic, albeit not the only, theme of transportation regulation. To the extent that public utility regulation can be justified on grounds unrelated to natural monopoly (I cannot myself think of any such ground), the critique of this Article is inapplicable.

3. See Kaysen, *Model-Makers and Decision-Makers: Economists and the Policy Process*, THE PUBLIC INTEREST, Summer 1968, at 80, 89-90.

4. I served with the Task Force as its general counsel in 1967-68, which will explain the frequency with which my examples are drawn from the communications industry. At this writing, the report of the Task Force to the President has not been published, but it is summarized in 34 TELECOMMUNICATIONS REPORTS, Dec. 9, 1968, at 1. Needless to say, the opinions in this Article are my own.

lenge to government regulation of business is intended. One regulatory framework whose continued existence is explicitly presupposed by my analysis is, as just mentioned, the antitrust laws. Regulations enforcing standards of health or safety are instances of the many other government constraints on business activity that lie outside the scope of my critique.

The Article, in four parts, attempts to (1) identify areas of behavior (such as prices and profits) where an unregulated natural monopolist might pursue policies contrary to the welfare of society; (2) describe the regulatory process as it operates today and, in a rough way, evaluate its social benefits and costs; (3) assess the possibilities of constructive reform; (4) consider some alternatives to regulation and offer some practical suggestions.

I. THE GROUNDS FOR REGULATING PRICES, ENTRY, OR OTHER BUSINESS CONDUCT IN A NATURAL MONOPOLY MARKET

In this opening branch of the analysis, I shall have nothing directly to say about the concepts or practice of regulation. Rather, I shall ask in what respects one might expect business performance under conditions of natural monopoly to be unsatisfactory from a social standpoint. When these elements of predictably deficient performance have been isolated, it will be possible to consider the extent to which the regulatory process is responsive to actual and serious problems.

A. Monopoly Prices and Profits

Under competition, the price of a good to the consumer tends to be bid down by the sellers to its cost (including in cost such profit as is required to attract capital into the industry). Consumers, as a result, obtain many goods at prices that are appreciably lower than the actual value of the goods to them. Monopoly enables the seller to capture much of the extra value that would otherwise accrue to consumers. To illustrate, let us suppose that if aspirin is sold at 1 cent per half grain (its cost) there will be 200 purchasers and that if it is sold at 10 cents there will still be 100 purchasers. The monopolist who desires to maximize his profit will sell at 10 cents—the monopoly price—where his total cost will be \$1 and his revenue \$10, producing a supracompetitive profit of \$9. Monopoly prices are widely considered to be socially undesirable because of their alleged effects on income distribution, overall economic stability, the allocation of economic resources, and proper business incentives. The arguments in support of these grounds are briefly as follows:

The effect of charging a monopoly price is to transfer wealth from

the consumers of a product to the owners of the firm selling it.⁵ The consumers are deprived of much of the extra value that they would enjoy in a competitive market, where they would be able to purchase at cost; the stockholders are enriched by capturing a good part of that value in increased profits. Transfers or redistributions of wealth are unavoidable in a society that is not perfectly egalitarian. At the same time, one could argue that it is sound social policy to reduce disparities of income and wealth so far as compatible with maintaining proper incentives. The redistribution of wealth that monopoly profits effect seems inconsistent with that goal. Consumers as a class are probably less affluent than stockholders; and a monopoly profit performs no obvious incentive function (our definition of cost included a profit sufficient to keep the firm in business).

It is further argued that insufficient demand in the private sector, a cause of recession, could be aggravated by a transfer of income from consumers to investors. The latter, being a more affluent group, are apt to save a larger proportion of their income. In periods of declining demand, moreover, a monopolist may be slower to reduce price than a competitive firm. In addition, by creating higher prices than would prevail under competition monopolization might be thought to aggravate any inflationary tendencies. And since a monopolist (as we shall soon see) uses less of the factors of production than a competitive firm, monopoly might appear to promote unemployment.

The mere act of redistributing wealth between two classes of individuals, while possibly offensive to ideals of social justice or adverse to the proper working of the business cycle, is not inconsistent with obtaining maximum benefit from the nation's economic resources. But the means by which the monopolist seeks to maximize profits may create inefficiency. Suppose that a widget costs 4 cents to produce (regardless of quantity) and that the widget monopolist can sell 10,000 at 7 cents, 12,000 at 6 cents, 13,000 at 5 cents, and 14,000 at 4 cents. Given this demand schedule, the profit-maximizing monopolist will sell at 7 cents, where his total cost is \$400, his total revenue \$700, and his monopoly profit \$300. Whether we prefer stockholders or consumers to derive the greater benefit from the production of widgets, society as a whole is worse off when the monopoly price of 7 cents is charged rather than the competitive price of 4 cents. When 14,000 are sold at the competitive price, consumers who would have taken 10,000 widgets at 7 cents derive extra value of \$300 from being able to purchase at cost. This just offsets the monopolist's loss, but there are further gains: Consumers who would have purchased an additional 2,000 at a price of 6 cents derive a

5. Insofar as companies retain a considerable portion of their earnings, monopoly profits may also be said to transfer wealth from consumers to corporations. That effect will be considered when we discuss the political dimension of the monopoly problem. See text accompanying note 85 *infra*.

286
157

value of \$40 above what they paid at the competitive price; and those who would have paid 5 cents each for the additional 1,000 derive extra value aggregating \$10. The total consumers' surplus when the competitive price is charged is thus \$350. This sum exceeds the monopoly profit (or producer's surplus)—\$300—that the seller obtained by charging a higher price.⁶

The intuitive basis of the illustration is quite simple. Because the utility functions of individuals vary, the monopolist selling at a single price cannot capture the entire consumers' surplus that a sale at cost would produce. [The price that captures as much as possible necessarily excludes a group of potential consumers to whom the utility of the product exceeded its cost of manufacture. The monopoly price thus prevents the economic system from meeting wants that could be met perfectly well. Consumers

6. Conceivably, the \$50 difference between the monopoly profit obtained and the consumers' surplus sacrificed understates the economic cost of monopoly, for an extra dollar of income may be worth less to stockholders as a group than to consumers as a group, assuming the former to be richer. See Lerner, *The Concept of Monopoly and the Measurement of Monopoly Power*, 1 *REV. ECON. STUDIES* 157, 158-59 (1934), reprinted in *READINGS IN MICROECONOMICS* 239, 240-41 (W. Breit & H. Hochman eds. 1968). We shall disregard this possibility, however, in view of its highly conjectural and uncertain character. See T. SCITOVSKY, *WELFARE AND COMPETITION: THE ECONOMICS OF A FULLY EMPLOYED ECONOMY* 60 (1951); text following note 39 *infra*.

The concept that monopoly pricing causes welfare losses, illustrated in the text by a rather stylized arithmetical example, can also be represented, and perhaps more clearly, graphically. Let dd be the range of prices at which various quantities of widgets will sell—in other words, the demand schedule for widgets. Under competition it is evident that the equilibrium price is p_c and output O_c ; for at any higher price additional output could be sold at a remunerative price—a price that exceeded the cost of the additional output (marginal cost or MC)—while at any lower price cost would exceed revenue. When p_c is the price charged, consumers' surplus equal to the area Ap_cC is generated, representing the additional amount that consumers could be made to pay for widgets under a system of perfect discrimination. A monopolist, on the other hand, would be free to restrict his output to O_m and charge the higher price p_m , the point from which any further reduction in price would generate less additional revenue (marginal revenue or MR) than additional cost. At that price consumers' surplus is reduced to the area Ap_mD and the monopolist appropriates the area Dp_mB as monopoly profit or producers' surplus, resulting in a net diminution in welfare of $p_m p_c B$. That area represents the "deadweight loss" of monopoly—the part of consumers' surplus that the monopolist cannot appropriate but that the consumers lose. One should note, however, that this model of monopoly performance is highly simplified; for a number of refinements besides those I shall discuss in the text see J. ROBINSON, *THE ECONOMICS OF IMPERFECT COMPETITION* 143-58 (1933).

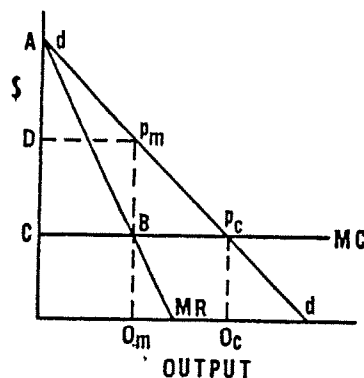


FIGURE 1

may be led to substitute more costly or less useful products merely because the cost of widgets to them is too high, although society's economic resources would be better used producing widgets rather than substitute products. It can also be shown that in limiting output the monopolist is underutilizing productive resources.

Finally, the ability to obtain very substantial profits without particular exertion, merely as a consequence of enjoying a monopoly, may be thought to dull incentives to efficient and progressive operation. A firm that is continuously and effortlessly very profitable may not feel much sense of urgency about reducing costs in order to obtain still greater profits.

The case for condemning monopoly prices and profits just outlined is less compelling than it perhaps first appears. It is not clear that an unregulated monopolist will normally charge a price that greatly exceeds what a nonmonopolist would charge for the same service; nor is it clear that society should be deeply concerned if a natural monopolist does charge an excessive price.

One possible ground for doubting that grossly excessive prices and profits are likely to flow from the possession of a monopoly can be derived from the theory that the large modern corporation does not seek to maximize profit.⁷ The revisionist theory, as one might apply it to a monopolist, may be summarized briefly as follows: Management in the large modern corporation is largely autonomous and self-perpetuating. The nominal owners, the stockholders, will assert control only if the corporation fails to produce a respectable profit, comparable to that of similar firms but not necessarily the maximum that management could extract. To be sure, if competition is sufficiently vigorous, the managers will be constrained, not by stockholders but by the market, to sell as dearly as they can while minimizing cost. Under competition, there is in theory only one profit—the return necessary to attract and hold capital—not a range of possible profits that includes a comfortable but moderate return near the bottom of that range. But it is possible that in many industries price competition is not very effective due to fewness of sellers, barriers to entry by new competitors, and other factors. Management in such industries may enjoy a broad area of discretion as to how much profit to make. Since the managers, it is argued, derive no direct

7. For a forceful recent exposition of the theory see J. GALBRAITH, *THE NEW INDUSTRIAL STATE* *passim* (1967); D. LAMBERTON, *THE THEORY OF PROFIT* (1965); J. MCGUIRE, *THEORIES OF BUSINESS BEHAVIOR* (1964); O. WILLIAMSON, *THE ECONOMICS OF DISCRETIONARY BEHAVIOR: MANAGERIAL OBJECTIVES IN A THEORY OF THE FIRM* 12-25 (1964); Alchian, *The Basis of Some Recent Advances in the Theory of Management of the Firm*, 14 *J. IND. ECON.* 30 (1965); and Machlup, *Theories of the Firm: Marginalist, Behavioral, Managerial*, 57 *AM. ECON. REV.* 1 (1967), contain excellent summaries of the earlier literature. For a lively debate on the question see Peterson, *Corporate Control and Capitalism*, 79 *Q.J. ECON.* 1 (1965); Kaysen, *Another View of Corporate Capitalism*, 79 *Q.J. ECON.* 41 (1965); Peterson, *Corporate Control and Capitalism: Reply*, 79 *Q.J. ECON.* 492 (1965). My statement of the "non-profit-maximization" or "managerial discretion" theory is perhaps overstrong. For a more cautious and hypothetical statement and analysis see O. WILLIAMSON, *supra*, *passim*.

pecuniary benefit from higher profits, they can be expected to subordinate profit maximization to objectives of more immediate personal concern, such as security, corporate image, pleasant surroundings, good labor relations, high salaries, empire building, and so forth. Such tendencies should be especially pronounced among monopolists, since they enjoy the greatest freedom from competitive pressures. From this it might seem proper to infer that an unregulated monopolist would not charge monopoly prices or collect monopoly profits.

I consider this dubious reasoning. To begin with, the view that managers of a publicly held firm are likely to maximize stockholder earnings is at least as plausible as the view that they are not. Investors do care a great deal about the earnings of the firms in which they invest, since earnings significantly affect both dividends and the market value of a stock. Large investors, at least, do have ways of impressing their concerns on management. And the take-over bid is not unknown. It constitutes an ever-present threat to the incumbent management, and like any deterrent its effectiveness cannot be measured by the frequency with which it is actually employed. Moreover, most firms require access to outside capital as at least a marginal source of funds, and diminished earnings will mean diminished funds from the sale of additional securities. Even if not coerced by stockholders or market forces to maximize earnings, business managers might adopt that course because they viewed earnings as the most appropriate criterion of business success and the surest path to prestige, security, and other elements of personal fulfillment. Not least, managers typically do own stock in their company, not enough for control but quite enough to give them a substantial personal stake in the stock's performance and therefore in the firm's earnings.

The empirical evidence on profit maximization by large and relatively secure firms is as yet inconclusive. We know, for example, that patent and copyright holders and other monopolists commonly practice price discrimination.⁸ As we shall soon see, discrimination is the profit-maximizing strategy of a monopolist. At the same time it is highly unpopular with purchasers, government agencies, and society at large. Its prevalence in these circumstances is some indication of the persistence of the profit drive among those insulated from direct competitive pressures. But it is an inconclusive indication. We shall soon see that price discrimination is consistent with other corporate goals besides maximizing the shareholders' earnings.

The evidence in support of the new theories of the firm is also impressionistic and inconclusive.⁹ Perhaps the best evidence is the fact that many

8. For examples of price discrimination by two unregulated monopolists, Alcoa (before World War II) and United Shoe Machinery, see C. KAYSER, UNITED STATES V. UNITED SHOE MACHINERY CORPORATION: AN ECONOMIC ANALYSIS OF AN ANTI-TRUST CASE 146 (1956); Machlup, *Characteristics and Types of Price Discrimination*, in BUSINESS CONCENTRATION AND PRICE POLICY 397, 417-18 (Nat'l Bureau Econ. Research 1955).

9. William Baumol characterizes the empirical basis for his theory that firms seek to maximize

corporations make charitable contributions. However, the amounts that corporations give to charity are trivial in relation to their profits,¹⁰ and one of the reasons why this is so, surely, is that stockholders would be justifiably outraged to see management divert substantial profits, properly theirs, to charitable ends of the managers' devising. At most, such evidence indicates that firms do not always seek to maximize short-run profit when to do so might undermine the firm's prosperity in the long run. A charitable contribution is fully consistent with long-run profit maximization; a modest expenditure buys an asset of some value to any firm appraising its long-term prospects—public goodwill. The corporate-gift example suggests a reconciliation of the opposing viewpoints in the debate over profit maximization: [the large corporation seeks to maximize profits, but over the long rather than the short run.]¹¹

A more critical point for our purposes is that even if the management of a monopolistic firm chooses not to maximize shareholder earnings—profits in the accounting sense—it might charge the same price that a conventional profit maximizer would charge, that is, the monopoly price. "Profit" and "profit maximization" are ambiguous concepts. To say that a firm is not maximizing profit may mean any one of a number of different things, and it is necessary to distinguish them. First, it may mean that the managers are, in effect, diverting monopoly profits to themselves in the form of salaries, bonuses, perquisites, and staff far in excess of what is required to attract and retain a competent management.¹² Such a course of

sales revenues rather than profits as "impressions gathered through casual observation." W. BAUMOL, BUSINESS BEHAVIOR, VALUE AND GROWTH 27 (rev. ed. 1967). Contrary evidence is not difficult to adduce at this level. The following is a quotation from the president of a large corporation: "We are not interested in volume unless it is highly profitable. The name of the game used to be 'how high do you stand on Fortune's 500.' We've dropped from 393 to 481 in the past three years. But we have gone from a 64¢ a share loss before special items in 1965 to a 51¢ profit in 1967 by chopping off \$30,000,000 worth of sales." INVESTOR'S READER, Sept. 4, 1968, at 17. Moreover, careful empirical study has failed to substantiate Baumol's hypothesis. See, e.g., Mabry & Siders, *An Empirical Test of the Sales Maximization Hypothesis*, 33 S. ECON. J. 367 (1967). The case studies of O. WILLIAMSON, *supra* note 7, and of R. CYERT & J. MARCH, A BEHAVIORAL THEORY OF THE FIRM (1963), are suggestive but inconclusive. (Additional studies are summarized in Williamson, *A Dynamic Stochastic Theory of Managerial Behavior*, in PRICES: ISSUES IN THEORY, PRACTICE, AND PUBLIC POLICY 11, 22-23 (A. Phillips & O. Williamson eds. 1967).) They show that under conditions of adversity firms find it possible to reduce costs appreciably. From this it is inferred that a firm not faced by adversity will allow a considerable organizational slack to build up despite the sacrifice of profits entailed thereby. But this is not a necessary inference. What is slack under adversity may be appropriate use of resources in other periods. When a firm's sales decline, for example, clearly it must adjust its expenses even though they were appropriate for the former level of output. For some recent statistical evidence that management-controlled firms may be less profitable than owner-controlled see Monsen, Chiu & Cooley, *The Effect of Separation of Ownership and Control in the Performance of the Large Firm*, 82 Q.J. Econ. 435 (1968).

10. See Hetherington, *Fact and Legal Theory: Shareholders, Managers, and Corporate Social Responsibility*, 21 STAN. L. REV. 248, 279 n.103 (1969); Schwartz, *Corporate Philanthropic Contributions*, 23 J. FIN. 479 (1968).

11. Cf. D. LAMBERTON, *supra* note 7, at 101-02; Mabry & Siders, *supra* note 9, at 377. As used in this context, the term "long run" does not have its usual connotation in economic discussion of a period within which all costs are variable; that is, long-lived assets wear out (or become obsolete) and must be replaced. The contrast I wish to suggest, rather, is between maximizing for all periods and maximizing only for the present period, or, less formally, between a time horizon of, say, 5-10 years and one of, say, 1-2 years.

12. See O. WILLIAMSON, *supra* note 7, at 129-34.

action, if pursued by the management of a monopoly firm, would require the fixing of a monopoly price in order to support the abnormal return to the managers.

Second, insistent upon only moderate profit, the management of a monopolistic firm might be slack and allow costs to drift upward. This hypothesis also assumes that prices well above the minimum attainable cost level are being charged. Third, management might try to maximize profit but fail because of uncertainty about demand, costs, and other relevant conditions. Or, baffled by the complexities of determining the precise combination of outputs and prices that maximizes profit, management might fall back on more or less crude proxies or rules of thumb to guide its decision.¹³ Presumably, however, its decision rules would be designed to approximate monopoly price.

It has also been suggested that management typically seeks to maximize sales revenues rather than profits, as an end in itself.¹⁴ It is not clear, however, that a sales-maximizing monopoly would charge a price or prices that did not return substantial monopoly profits. This is best shown graphically (a procedure that has the added advantage of introducing some concepts that will recur in later discussion). Under the cost and demand conditions pictured in Figure 2, a profit-maximizing monopolist selling at a single

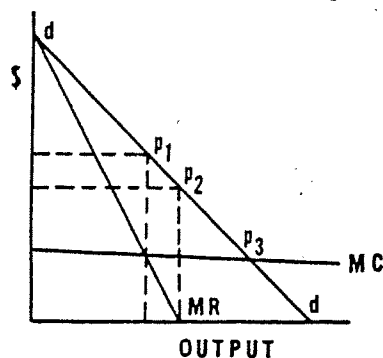


FIGURE 2

price would sell at p_1 . From any higher price (and therefore smaller output) the firm would have an incentive to move toward p_1 ; for an additional sale would generate more extra revenue (marginal revenue or MR) than extra cost (marginal cost or MC). A sales-maximizing monopolist, on the

13. See W. BAUMOL, *supra* note 9, at 29-30.

14. W. BAUMOL, *supra* note 9, at 46-48. Professor Baumol also presents a modified version of the sales-maximization hypothesis: Companies seek to maximize the rate at which their sales grow. *Id.* at 96-101. This formulation is not entirely easy to distinguish from old-fashioned profit maximization, since, as Professor Baumol stresses (*id.* at 96-97), large profits are necessary to facilitate rapid corporate expansion and to attract such outside capital as is necessary to that end. See also R. MARRIS, *THE ECONOMIC THEORY OF 'MANAGERIAL' CAPITALISM* (1964).

other hand, would sell at p_1 . Any lower price would produce negative marginal revenue—that is, his total sales revenue would decline. Under competition, finally, price would be bid down to p_1 , where price equals marginal cost. Thus, in our illustration the sales-maximizing price is well above the competitive price and includes substantial monopoly profits. But one should note that this is not a necessary characteristic of monopoly; if MC intersects dd at or above p_1 , the sales-maximizing price will be equal to or lower than the competitive price.

Thus far, it has been assumed that a monopolist would sell at a single price. Figure 3 shows, however, that a profit-maximizing monopolist able

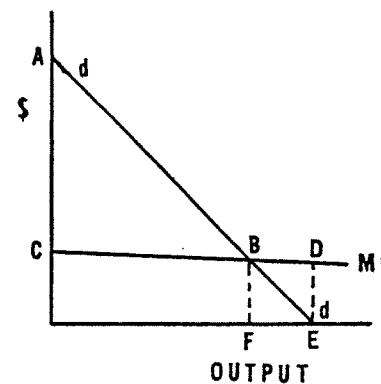


FIGURE 3

to discriminate perfectly (we shall see later that natural monopolists are normally in a good position to discriminate finely, although not perfectly) will sell at the range of different prices lying along dd between A and B (the latter being the point at which MC intersects dd) and will obtain a monopoly profit of ABC . The sales-maximizing monopolist will proceed likewise but will then continue down the demand curve, selling additional output at prices ranging from B down to a point just above E , and will thereby obtain additional revenue BEF (at a loss equal to BDE).

From a monopolist's decision to maximize sales, therefore, it does not necessarily follow that he will not obtain monopoly returns. A final possibility is that management might, out of pure benignity, forgo any monopoly profit and sell at the same price that competition would dictate. But there is no evidence that such a tendency is common, and it would be surprising if it were. In sum, unless this last and least plausible version of the revisionist theory is adopted, even a "non-profit-maximizing" monopolist is quite likely to charge the monopoly rather than the competitive price.¹⁵

15. Whether monopoly profit is taken directly or transformed into unnecessary expenses (e.g., slack) will not affect monopoly pricing but may affect other dimensions of the monopoly problem.

The relationship between the profit motive and pricing policy is loose in another respect. A price that does not maximize profit may be either lower or higher than the profit-maximizing price, and if it is higher the impact on efficient resource allocation will be even more adverse than if the firm charged the profit-maximizing price.

The controversy over whether firms insulated from strong competitive pressure maximize profit thus sheds little light on monopoly pricing. On the other hand, the distinction mentioned earlier between short-run and long-run profit maximization is highly pertinent. Business policies designed to maximize the present period's earnings may be short-sighted in their neglect of events that could drastically impair future profits, such as restrictive legislation or entry by new competitors into the firm's markets. A monopolist maximizing long-run profit may or may not charge the monopoly price determined without regard for more or less distant contingencies. On the one hand, he may decide to sell at a somewhat lower price in order to discourage entry by potential competitors or for other strategic reasons.¹⁶ On the other hand, he may charge the monopoly price but divert a portion of the abnormal return to expenditures designed to improve the firm's long-run position. An example would be an advertising campaign designed to generate public goodwill (rather than to expand sales) in order to ward off possible political interference with the continued enjoyment by the firm of its monopoly position, or, as previously mentioned, a charitable contribution designed with a similar effect in mind.

To summarize our discussion of profit maximization, the traditional assumption that a monopolist will strive to charge the monopoly price remains quite plausible, subject to one important qualification: The monopoly price depends on management's time horizon. In the long run, a persistently very large spread between price and cost may spur entrepreneurs to devise ingenious methods of challenging or supplanting the monopolist and legislatures of curbing him. The long-run monopoly price—a price fixed with these dangers in mind—may thus be significantly lower than the short-run monopoly price, although still well above cost. There is

Thus, diversion to workers of monopoly profit in the form of an excessively generous wage settlement does not detract from the inefficiency of the firm's charging a monopoly price, but it does mitigate distributive inequality, since presumably workers, unlike stockholders, are not as a group richer than consumers. On the general subject of the pricing implications of various theories of the firm see Williamson, *supra* note 9.

16. Alcoa apparently followed a policy of limit pricing when it had a monopoly of aluminum production. See *United States v. Aluminum Co. of America*, 148 F.2d 416, 426 (2d Cir. 1945). See also M. BOWMAN & G. BACH, *ECONOMIC ANALYSIS AND PUBLIC POLICY* 388-89 (2d ed. 1949); Coase, *Some Notes on Monopoly Price*, 5 REV. ECON. STUDIES 17, 26 (1937). For other strategic considerations supporting self-restraint in pricing see W. BAUMOL, *supra* note 9, at 43, 46. One might ask why entry would not be as effectively deterred simply by the knowledge that the monopolist could, if threatened by new entry, reduce his price, and therefore why limit pricing is a rational strategy. A possible answer is that in contemplating entry a firm is more likely to look to the price being charged in the market than to the costs of the firms selling there, because it is ordinarily much easier for a firm to determine the price charged by, than the costs of, another firm.

at least this much validity to the revisionist view of the modern corporation: It is plausible to suppose that the management of a typical monopolist would identify its own interests with the long-run interests of the corporation and would have sufficient independence from stockholders to fix the long-run monopoly price without fear of being dislodged. Although no more than plausible, this theory does have some empirical support.¹⁷

Let us turn now to a second respect on which the conventional assumption that monopoly results in excessive prices requires not contradiction but careful qualification. It is this: In attempting to determine the degree to which monopoly prices can reasonably be deemed excessive, it is improper to compare returns under monopoly with returns under fully competitive conditions. Competitive returns may no longer be the norm in our economy, given the prevalence¹⁸ of oligopoly. Many economists believe that firms in an oligopolistic market, a market in which a few firms account for most sales, tend to avoid vigorous price competition. Each one realizes that a price cut by it will cut so deeply into the sales of the others as to evoke prompt matching responses, resulting in lower profits for all.¹⁹ No economist believes that the number of firms in a market is the only index to whether such behavior is likely; for example, unless entry is for some reason difficult, tacit collusion to maintain a supracompetitive price level by the existing firms in the market, however few they may be, is unlikely to succeed for very long. Nor is it at all clear how concentrated a market must be for oligopolistic interdependence to emerge. But it is at least plausible that in many, perhaps most, markets today the pattern of prices and profits lies somewhere between that of monopoly and that of competition—and perhaps closer to the former than to the latter in a nontrivial number of cases—due to oligopoly.

Even so, one might reply, the proper course would be to restructure oligopolistic markets where necessary to restore competitive conditions and to force the prices of natural monopolists down to the competitive level through regulation; and if restructuring is impossible there is still no reason to allow the natural monopolist to earn more than a competitive return. Such reasoning raises some serious questions. Although this is not the place to explore the matter in detail, it is not clear that a broad restructuring of oligopolistic markets would be justified. To penalize by dissolution firms

17. Studies of the pre-World War II Alcoa monopoly offer some support. See *United States v. Aluminum Co. of America*, 148 F.2d 416, 426 (2d Cir. 1945); L. WEISS, *ECONOMICS AND AMERICAN INDUSTRY* 221-22 (1961). But cf. D. WALLACE, *MARKET CONTROL IN THE ALUMINUM INDUSTRY* 225-63 (1937).

18. See J. BAIN, *INDUSTRIAL ORGANIZATION* 133-49 (2d ed. 1968); C. KAYSER & D. TURNER, *ANTITRUST POLICY* 26-41 (1959).

19. This assumes that demand for the firms' product is neither growing rapidly nor highly sensitive to price decreases. If either condition holds, all of the firms may be better off at a lower price level.

191

that succeed in obtaining large market shares could have a very bad effect on the incentive to compete aggressively. Moreover, the implementation of a policy of restructuring would probably be costly and prolonged. An administrative or judicial determination would have to be made of the minimum firm size²⁰ in each market; even in theory, this is an exceedingly difficult determination to make.²¹ These objections are weighty because the benefits of restructuring are somewhat speculative. There is considerable uncertainty about the actual effect of oligopoly on pricing. Conceivably it is not very great.²²

If we assume, for these or other reasons, that our economy is likely to remain highly oligopolistic (a stiff law prohibiting mergers that contribute to market concentration has been on the books for 18 years now without perceptibly reducing the degree of concentration),²³ there are compelling reasons for not attempting to reduce a natural monopolist's profit below whatever is the prevailing level in such an economy. It would be a curious policy that devoted substantial resources to compressing profits to the competitive level in a relatively small sector of the economy while countenancing in a much larger sector profits that may be substantially supracompetitive. Such special treatment could not be justified by any difference in the importance of the services provided by natural monopolists. What could be more vital than drugs and medicines, currently produced by a highly oligopolistic industry that enjoys an exceptionally high profit level?²⁴

In addition, to eliminate prices that exceed competitive levels in one industry while tolerating their continuance in many others is inefficient; it will cause excessive migration of resources to the former. Consumers will buy more of the product because the cost to them is now reduced, even though a substitute product made by another industry meets the same need better and at lower cost to society. From the standpoint of efficient allocation, a more sensible objective for an economy permeated by excessive returns may be to proportionalize the excess-profit factor in the prices of goods and services so as not to distort consumer choice (though even this is subject to a number of qualifications). But whatever may be the "second best" solution when the best (all prices at competitive levels) is unattain-

20. Not plant size, because there may be substantial multiplant economies to single-firm operation—for example, in marketing.

21. See generally Stigler, *The Economics of Scale*, 1 J. LAW & ECON. 54 (1958), reprinted in G. STIGLER, *THE ORGANIZATION OF INDUSTRY* 71 (1968).

22. See Stigler, *A Theory of Oligopoly*, 72 J. POL. ECON. 44 (1964), reprinted in G. STIGLER, *supra* note 21, at 39. But see, e.g., N. COLLINS & L. PRESTON, *CONCENTRATION AND PRICE-COST MARGINS IN MANUFACTURING INDUSTRIES* 115-16 (1968).

23. *Hearings on the Status and Future of Small Business in the American Economy Before the Senate Comm. on Small Business*, 90th Cong., 1st Sess. 475, 484 (1967) (statement of Dr. Willard F. Mueller, Director, Bureau of Economics, Federal Trade Commission). The law, of course, is the Celler-Kefauver Antimerger Act, amending section 7 of the Clayton Act, 15 U.S.C. § 18 (1964).

24. Profits after taxes of more than 20 percent of stockholders' equity are not uncommon. Merck, for example, had a 25.4 percent profit in 1967. *FORTUNE*, June 15, 1968, at 192-93.

able, it is not to eliminate profits lopsidedly and thereby create improper price signals.²⁵

Nevertheless, it is probably the case that unregulated natural monopolists could extract profits somewhat higher than those prevailing in oligopolistic industries; and to eliminate *that* excess might well be a step in the right direction. As noted earlier, the precise impact of oligopoly on price competition is not known and may not be very great after all, especially in the many industries that are only loosely oligopolistic in structure. And oligopolists lack the advantage over potential competitors that the natural monopolist has by virtue of the economies of scale in a natural monopoly market. The monopoly return might, consequently, be higher than one would judge desirable. But the magnitude is uncertain. Moreover, so long as oligopolistic interdependence, governmentally sanctioned restrictive practices (as in agriculture), and other serious market imperfections result in substantial disparities between price and marginal cost in *some* important markets not wholly unrelated to the natural monopoly markets, we cannot be *sure* that the elimination of monopoly pricing in the latter markets would improve the efficient allocation of resources or even that it would not worsen it. Analysis of the "second best" problem has demonstrated the pitfalls of the piecemeal approach.

Finally, one should note that the possession of a monopoly does not always enable a firm to charge the monopoly price. Although only a single natural gas pipeline company can efficiently supply a new increment of demand in the Los Angeles retail market (let us say), more than one pipeline company in the western United States is in a position to construct and operate that pipeline. The retail gas company serving the Los Angeles market can therefore (I am assuming no regulatory constraints on the bargaining process) invite bids from several competitors for the opportunity to serve that market. Unless the bidding process is collusive, the pipeline that wins the long-term contract to supply the city's new demand should be the one whose price for the term is closest to cost and contains the least monopoly profit.²⁶ But the bargaining process cannot be relied upon as a complete solution to the problem of monopoly price. Unless the parties

25. For discussions of the vexing problem of "second best" solutions to allocative inefficiency see 2 J. MEADE, *THEORY OF INTERNATIONAL ECONOMIC POLICY* 102-18 (1955); Baumol, *Monopolistic Competition and Welfare Economics*, 54 AM. ECON. REV. PAPERS & PROCEEDINGS 44, 46 (1964); Lipsey & Lancaster, *The General Theory of Second Best*, 24 REV. ECON. STUDIES 11, 16-17, 25 (1956); Mishan, *Second Thoughts on Second Best*, 14 OXFORD ECON. PAPERS (N.S.) 205, 214-17 (1962).

One would not be concerned with the allocative consequences of forcing down a monopolist's price to marginal cost when other products were being sold at prices in excess of their marginal costs if there were no substitutes for the monopolist's product. There would in that case be no allocative consequences. But if monopoly markets are that insulated from other markets, neither is there any reason to expect that monopoly pricing has any misallocative effects, i.e., that it shifts demand to substitutes that cost society more to produce.

26. See Demsetz, *Why Regulate Utilities?*, 11 J. LAW & ECON. 55 (1968); cf. *United States v. El Paso Natural Gas Co.*, 376 U.S. 651 (1964).

build in private regulatory devices such as renegotiation with arbitration in the event of disagreement, the process may create an inefficient bias toward contract terms longer than risk conditions justify. More important, the buying side may be too fragmented to bargain effectively (as is true of telephone subscribers, for example). Still, bargaining may be an antidote to monopoly pricing in some cases, and it is an especially significant factor to be borne in mind when contemplating the extension of regulation to a new industry such as cable television, where the opportunity of local government, representing the subscribers, to drive a hard bargain with the would-be monopolist may be a viable alternative to conventional methods of regulation.²⁷

Admitting the force of the foregoing points, it is nonetheless plausible to assume that an unregulated monopolist will typically set prices and obtain profits that are in a meaningful sense excessive, albeit less so than popularly supposed. But it is a fair question whether the eradication of such excess profits is necessary or important to the social goals of promoting equitable income distribution, overall economic stability, efficient allocation of resources, and incentives to innovation and cost reduction.

Society condemns certain forms of income redistribution because of the means employed and without inquiry into the impact of the redistribution on the pattern of incomes; an extreme example is larceny. The first question in evaluating the distributive effect of unregulated natural monopoly, therefore, is whether the natural monopolist's extraction of consumers' surplus is the kind of conduct that should be prevented regardless of the actual magnitude or direction of its effect on incomes. If the answer is no, we proceed to the next question, whether unregulated natural monopoly causes or contributes to a socially undesirable income distribution.

In considering the first question we may begin by noting that a monopoly profit is simply a form of "economic rent," a term economists apply to any return obtained by virtue of controlling a scarce or unique factor of production. The profit that an individual realizes when he sells his home in an area where the pressure of increasing population has made real estate more scarce and therefore more valuable than when he bought is a classic example of economic rent. While the receipt of economic rents was once a burning social issue in America,²⁸ it is no longer. The conventional reply to a comparison of monopoly to other rents is that monopoly rents are the result of an artificial, contrived scarcity rather than a natural scarcity, and that prices inflated by a rent factor serve a valuable purpose in rationing naturally scarce resources such as land or petroleum while the monopolist's rents serve no comparable social purpose.

27. See text accompanying notes 199-201 *infra*.

28. For the story of Henry George's crusade against land rents see H. GEORGE, *PROGRESS AND POVERTY* (1954); 3 J. DORFMAN, *THE ECONOMIC MIND IN AMERICAN CIVILIZATION* 142-49 (1949).

olistic rents serve no comparable social purpose. The reply is compelling if one's image of the monopolist is of the classical "engrosser" who buys up all of the available corn on the way to the market and then forces up the price by withholding an adequate supply, or of the holder of a government franchise that limits entry of competitors. Our concern is with the unregulated *natural* monopolist. His market power flows from the cost and demand characteristics of the market in which he is selling, rather than from unfair or restrictive tactics or from legal privileges. Moreover, we shall see that the natural monopolist is well situated to adopt a method of pricing—discrimination—that maximizes profit without necessarily restricting output.²⁹ Thus, although a natural monopolist should be able to extract large profits, it is difficult from an ethical standpoint to distinguish an individual who obtains a high return by virtue of an interest in a natural monopoly firm from one who owns a strategically located plot of land and watches its value rise year after year without any skill or effort on his part.

Neither is it true that monopoly profits, unlike other forms of economic rent, serve no useful function in the regulation of the economy, although the function they serve is not rationing. Under competition, we need worry little about a firm's incentives to price efficiently, to minimize its costs, and to innovate. If it is inefficient the firm may be badly hurt or even destroyed by its rivals; the possibility should provide enough motivation for good performance. The "stick" of competitive displacement is absent under monopoly, or at least smaller. But supracompetitive profits provide a substitute incentive that may be nearly as effective, although in the form of a "carrot." To anticipate subsequent discussion, the unregulated monopolist has a strong incentive to price efficiently, to minimize costs, and to innovate, because these tactics will enable him to increase his profits. Deny the monopolist the opportunity to obtain profits in excess of his costs and you may destroy his incentive to better his performance. In principle, one would like to distinguish between those supracompetitive profits that reward a monopolist for superior performance and those that constitute a simple mulcting of the consumer. In practice such a separation seems impossible.³⁰

Our comparison of land and natural monopoly rents suggests a general formulation of the difference between redistributions that are condemned regardless of their actual effect on the pattern of incomes in society and

29. See text accompanying notes 43-47 *infra*. One should note, however, that to effectuate price discrimination a monopolist may have to impose on his purchaser restrictions against resale, in which event he would be interfering with the workings of a free market. But we shall see that such restrictions are typically not necessary in natural monopoly markets because the product is a service, which by nature is not readily transferable.

30. See text accompanying notes 175-84 *infra*. The term "monopoly profits" will be used throughout to embrace both kinds of supracompetitive return.

those that are condoned unless that effect is harmful. Some activities are, at best, worthless to society. Examples are the manufacture of burglary tools and the formation and enforcement of cartels. These activities are to be discouraged quite apart from any effect on income distribution. It is thus sufficient condemnation of the redistributions to which burglary or the monopolization of competitive markets gives rise that they constitute inducements to socially undesirable conduct.³¹ But the effort of a businessman to monopolize a market by producing at a cost so low as to drive out his competitors and deter new entry or, the monopoly achieved, to improve his return by lowering his costs still further is not at all reprehensible.³² It is conduct we want to encourage, and supracompetitive profits provide the inducement to engage in it. While, to repeat, it would be nice to be able to distinguish between those supracompetitive profits that provide the necessary inducement to efficiency and those that are pure windfalls, in the absence of a reliable method for making the distinction it is unreasonable to equate the profits of natural monopoly with those of antisocial conduct.

The remaining question is whether the profits of natural monopolists cause or aggravate an undesirable pattern of incomes in society. To answer this question, we must first determine what the income effect of monopoly profits is. The conventional assumption that they redistribute income from a poorer class—consumers—to a richer—stockholders—cannot be maintained without careful qualification. On the consumer side of the equation, one should note that many purchasers of natural monopoly services are business firms, which sometimes will, but sometimes will not, be able to pass on the bulk of a cost increase to *their* customers. At some point, moreover, many of the "essential" services provided under conditions of natural monopoly become luxuries. Examples are colored telephones, water for swimming pools, electricity for air conditioning, and long-distance telephony for casual chit-chat. Stated another way, natural monopoly services seem to some extent income elastic: Wealthier people tend to buy more and poorer less of these services. A profit-maximizing monopolist in these circumstances will try to design a rate schedule that enables the poorer consumer to purchase at a price closer to the marginal cost of serving him than the wealthier consumer is charged, lest the former be deterred from taking service by a high price. Since, as explained later, a natural monopolist's marginal cost is lower than his average total cost, this kind of price discrimination—discrimination in favor of the less affluent and against the

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

32. One could, of course, argue that the opportunity to reap natural monopoly profits may bias private inventive activity in the direction of process or product innovations that lend themselves to monopolistic exploitation because they involve large economies of scale.

more affluent consumer—may in some instances enable the less affluent to obtain natural monopoly services at lower rates than those corresponding to average total cost. Competition, if it were viable under these conditions, would prevent the seller from loading a disproportionate amount of his total costs on a group of wealthier customers able and, in the absence of good substitutes, willing to shoulder them;³³ it would force him to charge everybody average total cost.³⁴

On the shareholder side of the redistribution equation, one should note that a significant proportion of the equity capital in our society is owned by employee pension funds, by universities and other charitable foundations, and by individuals of moderate means.³⁵ Some monopoly profits, moreover, are probably distributed to individuals other than shareholders—for example to workers.³⁶ Most important, the degree to which wealthy individuals can actually increase their wealth by virtue of monopoly profits depends critically on the structure of the tax laws. To illustrate, let us suppose that company X, in a competitive market, has net income of \$200,000 per year before federal corporate income tax. To simplify computation the tax will be assumed to be a flat 50 percent. Company X's net income after tax will therefore be \$100,000. Assume that the market value of its common stock is \$2 million, all owned by individuals in the 70 percent bracket of the federal personal income tax, and that X distributes 50 percent of its after-tax income as dividends. Suppose that X obtains a monopoly of its market, and is able to increase its net income before tax by, say, 50 percent, or \$100,000. Since one-half of the additional income is taxed away by the corporation tax, X's net income after tax will increase to \$150,000.

The stockholders will not be greatly enriched by the added dividends—\$25,000—that accrue to them annually as a result of the acquisition of a monopoly, because \$17,500 will be taxed away. They will be enriched for

33. For a fuller discussion of price discrimination by natural monopolists see text accompanying notes 43-49 *infra*.

34. A further point is that the creation of a natural monopoly will—paradoxically—usually make the consumer better off than he was before, even if he must pay a very high monopoly price. Most natural monopolies have arisen not from changes in the methods of producing existing products or services but from the creation of new services—such as telegraphy, telephony, and electric power. The cost of a new service to the consumer, including whatever monopoly profit the seller is able to include, must be lower than that of the service it displaced; otherwise it would not have displaced the old service. But perhaps the proper comparison is not between today's consumers and yesterday's but among present-day consumers. Also, one should note that the displacement of an existing by a new service may harm some consumers—those who preferred the former service but could not sustain it by themselves when most of their fellow consumers switched to the new.

35. At the end of 1967, pension funds and nonprofit institutions held roughly 12 percent (by market value) of the stock listed on the New York Stock Exchange, and mutual insurance and savings institutions held additional substantial amounts. Computed from NEW YORK STOCK EXCHANGE, 1968 FACT BOOK 42 (1968). More than 50 percent of the shareholders of public corporations had a reported household income of less than \$10,000 per year. Computed from NEW YORK STOCK EXCHANGE, 1965 CENSUS OF SHAREOWNERS 15 (1965). On the other hand, all but 3 percent of the corporate stock (by market value) owned by consumer units in this country in 1962 was owned by the wealthiest 20 percent of those units. INEQUALITY AND POVERTY xxii (E. Budd ed. 1967).

36. See note 15 *supra* and text accompanying note 52 *infra*.

another reason. Assuming that the price-earnings ratio before the monopoly was acquired remains unchanged (for reasons that we need not dwell on here, it might well be higher or lower), the market value of *X* common stock will rise from \$2 to \$3 million. If the stockholders (all of whom, we shall assume, have owned the stock for more than six months) sell their shares at the new price, they will realize a gain of \$1 million, of which only \$250,000 will be taxed away. Although they will not, of course, receive the added dividends that they would have received had they kept the stock, \$7,500 per year after taxes is obviously a poor swap for a lump sum after taxes of \$750,000.

What is involved here is a gaping loophole in the federal personal income tax that enables individuals largely to escape the progressive feature of the tax by capitalizing future earnings. Were there no difference in treatment between long-term capital gains and other income, the principal beneficiary of monopoly profits would be the United States Government. Presumably revenues from this source would be expended by the government in accordance with public needs, including that of distributive justice.

In sum, if the tax system were really designed to further distributive justice, the distributive effect of monopoly as of other profits would be adequately corrected; as discussed earlier, there is no reason to draw invidious distinctions between natural monopoly and other forms of rent or income. If, on the other hand, the tax system is unprogressive, special treatment of the profits of natural monopolists will do little to achieve social justice. The natural monopoly sector is a small part of the economy.³⁷ The opportunities for altering the distribution of wealth by profit maximization in that sector pale by comparison with those afforded by the long-term capital-gains and other tax loopholes. And if special treatment of natural monopoly profits is nevertheless desired, we shall see that it can quite possibly be achieved by minor modification of the tax laws at less social cost than by a system of direct regulatory controls.³⁸

One can question, finally, whether income equalization is sound social policy, at least in the sense that would justify efforts to eliminate natural monopoly profits.³⁹ Most contemporary economists, for example, would

37. In 1967 the percentage of Gross National Product contributed by electrical, gas, telephone, and water companies, the principal natural monopolists, was well under 5. Railroads provided another 1.2 percent. Computed from United States Department of Commerce, Office of Business Economics, 48 SURVEY OF CURRENT BUSINESS, July 1968, at 27.

38. See text accompanying note 196 *infra*.

39. For a variety of perspectives relevant to the general question of distributive justice see *EQUALITY AND POVERTY* (E. Budd 1967); R. DAHL & C. LINDBLOM, *POLITICS, ECONOMICS, AND WELFARE* 134-61 (1953); G. KOLKO, *WEALTH AND POWER IN AMERICA—AN ANALYSIS OF SOCIAL CLASS AND INCOME DISTRIBUTION* (1962); A. LERNER, *THE ECONOMICS OF CONTROL: PRINCIPLES OF WELFARE ECONOMICS* 23-40 (1944); R. MUSGRAVE, *THE THEORY OF PUBLIC FINANCE: A STUDY IN PUBLIC ECONOMY* 19-22, 98-110 (1959); P. SAMUELSON, *FOUNDATIONS OF ECONOMIC ANALYSIS* 243-49 (1947); T. SCITOVSKY, *supra* note 6; Rahl, *Distributive Justice*, in *PHILOSOPHY, POLITICS AND SOCIETY—THIRD SERIES* 58 (P. Laslett & W. Runciman eds. 1967).

be unwilling to assert that a more equal distribution of wealth would increase the sum of human welfare or happiness. One can, to be sure, imagine cases where a redistribution from a wealthier to a poorer individual probably would increase the well-being of the latter more than it diminished the well-being of the former: A dime is doubtless worth more to a beggar than to most millionaires. But if we ask whether a redistribution of \$1,000 in annual income from a family whose income is \$20,000 to a family whose income is \$10,000 would have a similar effect, we shall indicate the difficulty of making interpersonal comparisons of utility, except in extreme cases. Not only do individuals with larger incomes tend to have larger expenses and a different conception of what is a necessity and what a luxury, but a higher income may compensate for the absence of nonmonetary satisfactions (such as greater leisure or less responsibility) that a lower-paying occupation might yield.

A theory having greater intuitive appeal is that the individual from whom wealth is redistributed (whether he is more or less affluent) will usually feel a sense of loss that is greater than the recipient's sense of gain—that people value the wealth they have more than new increments. This theory, however, lends little support to a policy of eliminating natural monopoly profits. We noted earlier that conditions of natural monopoly are historically associated with the creation of new services.⁴⁰ A consumer will not patronize a new service unless it makes him better off to do so, and if it does, he will not feel that his wealth has diminished even if the price of the new service includes a substantial monopoly profit.

The utilitarian ethic, then, the ethic that underlies the economist's conception of social welfare, does not imply a goal of equalizing incomes, save perhaps to the extent necessary to eliminate the extreme inequality we call poverty. On the contrary, the economist would be concerned, and rightly so, with the possible social cost in reduced incentives of equalizing the rewards of economic activity, not to mention the possible diminution in human satisfaction that might result from forced income uniformity among individuals of widely different tastes and ambitions. Since welfare economics is not the only source of social values, we cannot end our inquiry here. But other normative systems appear to yield a similar answer. Contemporary conceptions of fairness and social justice may be thought to require that the community assure all individuals the monetary resources necessary to maintain a decent minimum level of existence. And one can bolster this ethical notion with the political scientists' perception that poverty breeds social unrest and with the economists' that poverty has harmful spillover effects on the rest of society, for example, in the form of higher crime rates. It is also possible that extreme concentrations of wealth are a

40. See note 34 *supra*.

72X
591

threat to political stability. What is elusive is any broader goal of income equality than that implied by the preceding discussion. Certainly it is not to be found in the traditional American ideal of equality of economic opportunity, which is not at all the same thing as equality of economic rewards. Equality of rewards seems, if anything, inconsistent with equality of opportunity, and with other basic values such as personal freedom and individualism.

Less abstractly, in a society that is generally affluent even wide disparities of income may be quite tolerable, and necessary to foster individual incentive. The distributive objective in such a society shifts from greater equality of incomes as such to closing the gap between the majority of people, who are reasonably well off, and the minority who do not share in the general affluence. The income disparities that trouble our society today are not between individuals who have large unearned incomes and the rest of us, but between the average middle-class American and individuals who live in poverty. This points up the irrelevance of control of monopoly profits to any currently significant goal of income equalization. Redistribution of the profits of natural monopoly to consumers would alleviate the burdens of poverty to only a trivial extent. Unregulated land rents almost certainly are a much greater factor in the plight of the poor.

The reader might object that the foregoing critique of distributive justice undermines progressive taxation as well as control of monopoly profits. But that would be an erroneous inference. The Government must raise money somehow, and it is difficult to conceive of a method of doing so that would not have some effect on the distribution of income. In a context where distributive effects are probably unavoidable, it may be appropriate to indulge a preference for equality—though how far, and indeed whether income taxation is the best method of doing so, are matters of legitimate debate. What emerges from our discussion is the absence of any tenable principle upon which to base special measures to alter the distribution of wealth as between a natural monopolist and its customers.

Concern with the impact of natural monopoly on economic stability also seems misplaced. Apart from the fact that the federal government has powerful weapons in its monetary and fiscal policies for preventing depressions or recessions, assuring full employment, and curbing excessive inflation, the natural monopoly markets are probably much too small a sector of the economy to affect overall stability materially. An increase in price and constriction of output in one market by reason of a changeover from competition to monopoly should cause an expansion of output and decrease in price in others. Workers would flow from the monopoly to competitive markets. Since the market price of the monopolist's stock would be bid up, so that subsequent purchasers obtained only a normal

return, fears of excessive savings by wealthy shareholders seem exaggerated. If the monopolist responded to a general decline in consumer demand by raising his price still further, prices in competitive markets would simply decline more rapidly than would otherwise be the case. In short, so long as the natural monopoly markets remain a small sector of the economy, neither the formation nor subsequent behavior of natural monopolies is likely to aggravate business cycles significantly even in the absence of effective countercyclical policies.⁴¹

The argument that monopoly prices lead to a misallocation of resources requires qualification in three respects. First, as mentioned earlier it is difficult to assert confidently that the correction of excessive prices in one area of the economy will actually improve the efficiency of resource allocation; it may have the opposite effect. Second, some studies (contradicted, however, by others) indicate that the impact of allocative inefficiency on the nation's productivity may be slight, even if large monopoly profits are assumed.⁴² Third, the theory is based on an assumption that is peculiarly vulnerable as applied to a natural monopoly. It is that the seller will charge a single price. As pointed out earlier, the reason why a monopolist, in order to maximize profit, must fix a price that excludes consumers perfectly willing to pay him a normal profit and more is that there is no single price that captures the entire consumers' surplus, individual consumers having different utility functions.⁴³ Suppose that the monopolist is not required to charge a single price but is free to charge different prices for the same product regardless of cost, and that his purchasers are unable to resell. In that event the profit-maximizing monopolist will want to charge each consumer who will pay, at a minimum, a price that returns the monopolist a normal or competitive profit (below which the monopolist could employ his resources more profitably elsewhere) as much as the particular consumer, considering his individual needs and alternatives, is willing to pay.⁴⁴ The widget example shows that this is indeed the profit-maximizing strategy. By charging 7 cents for the first 10,000 widgets, 6 cents for the next 2,000, 5 cents for the next 1,000, and 4 cents for the last 1,000, the monopolist obtains a monopoly profit of \$350. Were he required to charge a single price to all purchasers, his maximum monopoly profit, as we saw earlier, would be only \$300.

41. See J. BAIN, *PRICE THEORY* 238-40 (1952); Stigler, *Administered Prices and Oligopolistic Inflation*, 35 J. BUS. U. CHI. 1, 8-9 (1962), reprinted in G. STIGLER, *supra* note 21, at 235.

42. See W. BAUMOL, *WELFARE ECONOMICS AND THE THEORY OF THE STATE* 161 (2d ed. 1965); Leibenstein, *Allocative Efficiency vs. "X-Efficiency"*, 56 AM. ECON. REV. 392-97 (1966), and studies cited therein. But see Kamerschen, *An Estimation of the "Welfare Losses" from Monopoly in the American Economy*, 4 W. ECON. J. 221 (1966).

43. See text accompanying note 6 *supra*.

44. Discrimination may also take the form of charging a customer different prices for different quantities of the good or service in question, since the strength of the purchaser's demand may vary with the quantity taken.

[Not only is price discrimination the profit-maximizing strategy of a monopolist, but under conditions of natural monopoly it may be the only feasible method of pricing consistent with an efficient allocation of resources. Natural monopoly refers to a market whose entire demand can be met at lowest cost by a single firm. This implies that before a firm can begin to do business it must sink large sums in a plant that is large enough or can readily be expanded to serve the entire market. Once the heavy initial fixed or overhead expenses are incurred, the cost of serving a particular customer is relatively slight. If the firm charged every customer the cost of employing idle capacity to produce an additional unit of output, it would not recover its overhead costs. Faced with such a situation, the firm could charge a single price that included a proportional share of the overhead costs as well as the additional cost of producing the unit. But such pricing would violate efficient allocation. It would exclude customers perfectly willing and able to pay the actual cost of expanding production to meet their demands, but no more. One solution, perhaps efficient but surely unrealistic, is for the government to pay the firm a subsidy enabling it to charge *all* purchasers the cost of producing an additional unit. Another efficient—and more realistic—solution is discrimination. Those who will pay only the additional cost are charged that amount. Other purchasers are charged as much as they will pay. In this manner the monopolist can recover its total costs without turning away anyone willing to pay the minimum cost of producing the units that he takes.⁴⁵

To illustrate, let us suppose that coal is discovered at Coaltown, 200 miles from the nearest market for coal (Markettown). Railroad *R* builds a line to Coaltown and fixes a rate (we shall assume no regulation) that covers both the total costs of the rail line—that is, the fixed costs (interest on bonds, real estate taxes, etc.) that are incurred regardless of whether any coal is actually hauled plus the expenses involved in the hauling—and the additional value that the coal operators, considering alternative transportation means and the price at which they can sell coal, are willing to pay the railroad to carry their product. Fixed costs are \$10,000 a year, operating expenses \$10 per ton, and the monopoly profit the railroad is able to exact \$5 per ton. The railroad hauls 1,000 tons of coal from Coaltown each year. Since it averages the fixed costs over this quantity, the rate is \$25 per ton.

The next year a lumber mill is built midway between Coaltown and Markettown. Because the lumber mill can truck its lumber products to

45. There is a good discussion of these points in Henderson, *The Pricing of Public Utility Undertakings*, 15 MANCHESTER SCHOOL OF ECON. & SOCIAL STUDIES 223 (1947). For a more recent treatment see Vickrey, *Some Implications of Marginal Cost Pricing for Public Utilities*, 55 AM. ECON. REV. PAPERS & PROCEEDINGS 605 (1965).

Markettown for \$17.51 per ton, it will not pay *R* \$25. In these circumstances *R*, if sensible, will offer to carry the lumber mill's products for \$17.50. At that price *R* covers the additional cost of serving this new customer (\$10, since the coal operators are defraying the entire fixed costs of the line) and obtains a monopoly profit of \$7.50.

Some years later a competing railroad, *R'*, builds a line to Coaltown and offers to carry coal to Markettown for \$12.51. Should *R* meet that rate or abandon the line? It should meet the rate. If the lumber mill is providing it with 1,000 tons of business a year, then at a rate of \$12.50 to the coal operators and \$17.50 to the lumber mill *R* will cover its total costs (\$10,000 plus \$10 per ton). Even if the lumber mill yields a smaller volume, so that *R* cannot cover its total costs at any price it can exact, it should not abandon the line, since both rates cover variable costs⁴⁶ and make some contribution to fixed or overhead costs. Fixed costs—those incurred independently of actual operations—by definition cannot be avoided by a cessation of operations. The railroad would be worse off by abandonment, since it would continue to owe the full \$10,000 a year.

The salient point is that the prices are efficient even though the price differential favoring the coal operators is not nicely proportionate to the cost of service (the lumber yard is closer and was established after the railroad incurred its heavy fixed costs in establishing the line to Coaltown, yet pays more) but only to the differing values that the respective customers place on the service. No customer willing to pay the minimum cost of serving him is denied service. And that was true when the price differential went the other way and both prices included substantial (but different) monopoly profits. A monopolist able to discriminate perfectly will not include in his price a monopoly profit so large that he will lose the customer.

[Unhappily for efficiency (if not for other social values), perfect price discrimination is rarely possible. A monopolist ordinarily cannot bargain with or otherwise ascertain the demand elasticity of each potential customer for each individual unit of output; and if he could society might find the procedure intolerable because of its extortionate flavor. As a practical matter, the monopolist must establish classifications, and unless these are very fine, in the process of attempting to maximize his profits from each class he may end up restricting output by as much as—or even more than—he would have done by selling at a single price.⁴⁷ To be sure, since natural

46. Much "value of service" pricing in the regulated industries fails to do this. See, e.g., J. MEYER, M. PECK, J. STENASON & C. ZWICK, *THE ECONOMICS OF COMPETITION IN THE TRANSPORTATION INDUSTRIES* 182 (1959). For discriminatory pricing to be efficient, the price to the favored purchaser must not be less than the true cost of serving him.

47. See J. ROBINSON, *THE ECONOMICS OF IMPERFECT COMPETITION* 190-95 (1933). But Mrs. Robinson's conclusion is that even imperfect price discrimination is more likely to increase than to reduce output, *Id.* at 200-02.

SE
t 91

monopolists are typically sellers of services that can be metered and are not readily transferable, they should be able to practice a highly refined, although not perfect, form of discrimination. And intuitively one would suppose that a highly refined discriminatory pricing system would result in greater output than if the single monopoly price were charged. This may be true but it cannot be proved rigorously. We can say only that the output of a discriminating natural monopolist will not necessarily be suboptimal, and that the degree to which it is suboptimal will vary from market to market.

To complete our discussion of discrimination, let us briefly consider the objections to it. It is commonly said to have undesirable secondary effects on the allocation of resources. Thus, if a change from a single price to price discrimination raises transportation charges to aluminum producers, the increased cost of aluminum will tend to shift demand to substitute products although the actual cost to society of transporting aluminum has not risen. On the other hand, the monopolist has good reasons of self-interest for not carrying discrimination to the point where major substitution effects occur. If a railroad raises its rate to aluminum producers by so much that aluminum users reduce their purchases, there will be less business for the railroad. Our earlier point, then, governs: If the monopolist can discriminate perfectly he will not charge prices that result in turning away any remunerative business. If he cannot discriminate perfectly, discrimination may have undesirable secondary effects.

A monopolist may have difficulty enforcing a finely discriminatory rate structure. Those who purchase at lower rates will have an incentive to resell to those in higher rate brackets. To prevent this kind of arbitrage, the monopolist may be forced into policing activity that is costly and that may run afoul of the long-standing public policy, held by the Supreme Court to be implicit in the Sherman Act,⁴⁸ against restraints on alienation. However, these problems would probably not be general in natural monopoly industries, since, as mentioned, the output of such industries typically consists of services that are not readily transferable. Even in the absence of formal measures to prevent arbitrage, then, it would be unlikely to erupt on a very large scale.

Discrimination is also challenged as an unfair method of competition, but we shall see later in this Article that the charge is questionable.⁴⁹ In sum, discrimination may be consistent with and even necessary to allocative efficiency. It is also the policy one would expect an unregulated monopolist to adopt voluntarily since it would maximize his profit. By that very token, discrimination aggravates the distributive effects of monopoly; it enables

48. Most recently in *United States v. Arnold, Schwinn & Co.*, 388 U.S. 365, 377-78 (1967).

49. See text accompanying notes 127-29 *infra*.

the monopolist to appropriate even more of the consumers' surplus than if he charged a single price. But that does not alter our point that, given discrimination, the extent to which unregulated natural monopoly leads to allocative inefficiency is uncertain.

The argument, finally, that possession of monopoly profits dulls the incentive to make additional profits—so that a comfortably prosperous firm will seek less assiduously for ways of reducing its costs in order to increase its profits than a lean firm—is plausible but, when one reflects on the actual financial structure of a publicly held corporation, unconvincing. The only individuals in a position to reap monopoly profits are those who own stock at the time that the monopoly is first obtained or first becomes valuable. As soon as it becomes known that a firm has a valuable monopoly, the price of its stock will rise as a means of discounting the anticipated future profits. Subsequent purchasers of the stock will not earn a monopoly return on *their* investment, nor will original owners derive any additional benefit from the firm's monopoly position beyond that reflected in the present value of their stock, until and unless the firm increases its profits. Current owners of a monopolist thus have the same incentive to improve the firm's earnings as the owners of a competitive firm.

B. Internal Inefficiency

In discussing the implications of monopoly for efficiency I have heretofore been concerned with how the price system allocates the nation's stock of economic resources among different industries to meet consumer wants at the lowest social cost. Another important, and to the layman a more familiar, kind of efficiency is cost minimization by the firm, which I shall call "internal efficiency."⁵⁰ I limit the term to mean the best possible use of a firm's resources within the existing state of technology. Efforts to reduce costs through advancing the state of the art are discussed in the next subpart under "innovation."

In a competitive market, the drive to minimize costs has aspects both of the carrot and of the stick. By reducing costs, the firm can obtain greater profits, either by continuing to sell at the market price or by shading that price and thus increasing its volume of sales. But the benefits are likely to be short-lived as competitors match the cost reductions and adjust price as their own costs fall in order to take maximum advantage of the new cost level. It is concern for survival that provides the strongest incentive to cost reduction by the competitive firm. If it fails to match a rival's cost reduction, possibly if it fails to anticipate a rival's cost reduction, it may find itself fatally disadvantaged.

50. For the distinction between allocative and internal (or, to the economist, "technical" or "X") efficiency see Leibenstein, *supra* note 42; Williamson, *Economics as an Antitrust Defense: The Welfare Tradeoffs*, 58 AM. ECON. REV. 18 (1968).

158
168

In the case of a monopolist the carrot is larger but the stick smaller. Cost reduction will enable the monopolist to increase its profits, and with less concern that the effect will be short-lived, since it has no rivals. But concerns for survival ordinarily play no part. The assumptions one makes about a monopolist's corporate objectives are thus quite important here. Certainly a monopolist who is a strict profit maximizer will be powerfully motivated to minimize his costs. It is difficult to argue that his motivation will be significantly less, on balance, than that of the competitive firm. Regrettably, a strong motivation to be efficient does not guarantee efficiency. Firms differ in their ability to minimize cost. Under competition, a firm either learns from its most efficient rival or goes under; either way production ends up at the least-cost level. The situation is more complex under monopoly. On the one hand, few entire industries (defining an industry as all the firms in the country that sell a particular good or service) are natural monopolies. Even the Bell System, which comes close to monopolizing the telephone industry, is a federation of semiautonomous regional operating companies rather than a monolith. Generally it is the regional or local market that can accommodate only a single firm. Within an industry, then, there will be a number of firms operating in separate markets and each firm will have a strong incentive to reduce costs. There should be sufficient diversity to produce many useful examples for emulation by the others, much as under competition.

On the other hand, conditions of cost and demand may vary significantly from market to market, and that will complicate efforts to borrow from efficient counterparts. A firm studying the methods employed in another market may have difficulty in determining whether lower costs in the other market stem from external factors or superior methods, and if the latter whether they are applicable to the problems that it faces in its own market. In short, even assuming that monopolists are assiduous profit maximizers in the conventional sense and hence strongly motivated to minimize their costs, one would still be concerned that those monopolists who, despite motivation, lacked great talent for cost minimization might have trouble imitating their more efficient cousins.

If managers of a monopoly firm exploit their opportunities to pursue ends other than immediate profit maximization, additional problems of internal efficiency may arise. Let us suppose that Firm *A*, a monopolist, is managed by Mr. *X*, who owns no stock in the corporation and who, because the stock of the corporation is widely dispersed and because all of the directors of *A* are members of management, controls the firm with minimum regard for the stockholders' interests. In a good year, with profits running to 30 percent of equity capital after taxes, *X* raises his salary by an amount equal to one-half of the firm's profits. Formally this is a sub-

stitution of a cost item (salary) for profits, and increases the firm's costs. Actually it is no such thing; it is a diversion of monopoly profits to *X*. The firm has not consumed any economic resources unnecessarily, but has simply distributed part of its profits to someone other than the stockholders.

X's action causes a murmur of disapproval among the stockholders, so the next year, rather than skim off some of the firm's monopoly profits in the form of salary, he splendidly refurbishes his office at a cost to the company of \$100,000. It is possible that this expenditure, too, represents nothing more than a diversion of monopoly profits to *X*, but that would be true only if *X*, had he felt free to take a slice of the firm's profits in money, would have devoted \$100,000 of his own money to refurbishing his office. He may have better things to do with \$100,000. He may derive less utility from refurbishing his office than he would from refurbishing his home. If so, the translation of monopoly profits into a business expense wasted resources: The economic welfare of society (of which *X*, of course, is a member) would have been greater if *X* had been given the \$100,000 directly.

Managerial self-indulgence of the kind illustrated in these examples may not be terribly serious from the standpoint of internal efficiency. It is largely (although as the last example shows not entirely) a matter of how monopoly profits are allocated between stockholders and managers, rather than how efficiently the firm is run.⁵¹ Moreover, it is not clear that managerial self-indulgence is either particularly widespread among major firms or involves large amounts of money. The growing professionalism and bureaucratization of corporate management should prevent gross excesses in this area. Most important, in a corporation with annual revenues of hundreds of millions of dollars, the amounts diverted by management to its own use (whether directly or in perquisites) above reasonable compensation are not likely to be substantial in relation to the corporation's sales or even profits.

What could be more serious is the subordination of immediate profit maximization to long-term firm and managerial interests in security, prestige, entrenchment, and political power and acceptability. A management not forced to reduce costs to the bone in order to survive is free to take a more strategic attitude toward corporate and personal destiny than one constrained by the market to pursue cost minimization and immediate profit maximization. Such a management may see value in acceding to the demands of labor unions after only nominal resistance in order to enlist union support in Congress or state legislatures behind legislation favorable to the firm's interests. It may decide to spend large sums on public relations

51. It could, however, distort the allocation of managerial talent as between monopoly and non-monopoly firms. Good managers might gravitate to the former because of the greater rewards available. This might lead to a general bidding up of managerial costs in relation to those of other factors of production.

297
691

in order to generate a favorable climate of opinion that might some day be useful in warding off legislation that the firm opposes or obtaining legislation it favors. It may give favorable pricing treatment to politically powerful purchasers such as the federal government. It may use its own purchases as a way of dispensing patronage to potentially useful allies in the business community. It may locate plants with a view toward maximizing the political support that it can generate in furtherance of its objectives. It may over-innovate in order to impress the public with its progressiveness.⁵³

This danger may be termed industrial politicization. Instead of pursuing a single-minded policy of profit maximizing in the short term, the firm recognizes the long-run value of building political support through corporate image and influence building and invests substantial sums in that pursuit. Such policies may require the firm to operate at an inefficient level of expenditure, although that is not a necessary implication. An excessively generous wage settlement with a union may represent simply a distribution of a portion of the firm's monopoly profits to its workers, analogous to the distribution to managers discussed earlier. (A settlement involving an agreement not to lay off unneeded workers might, in contrast, represent a real cost.) Nor would it be sound to regard all corporate efforts to influence the political process as wasteful or improper. If other groups use the political process to advance their economic welfare—as of course they do—business firms cannot reasonably be asked to abstain. One way of building political goodwill, finally, is to forgo monopoly profits, although that might leave some important potential sources of support, such as labor, unappeased.

In the absence of any systematic empirical study one can only guess at the gravity of the problem under discussion. I suspect that monopoly power is not the crucial variable. Competitive firms, after all, do many of the same things through trade associations. A more important variable may be the industry's involvement with government. Although the textile industry is competitive, one would expect a textile manufacturer having government contracts to consider the probable reaction of powerful Congressmen very carefully before relocating a plant or making some other major business move that could have political repercussions, and one dependent on continued government curtailment of imports to weigh carefully the probable reaction of the White House to any attempt to reduce labor costs.

One might be concerned, finally, that a monopolist who lacked the discipline of profit maximization might simply allow costs to drift upward toward his monopoly price, tolerating inefficiency until his profits were deeply eroded. But this assumes that the firm that does not maximize profit

52. For some evidence relating to regulated monopolists see note 102 *infra*.

has no other maximands. Whether management is seeking to line its own pockets, to build pyramids, or to accumulate political support, failure to exercise close cost control will only impair its objectives. A more plausible hypothesis is that the organizational characteristics of the modern large firm preclude effective cost control except in response to conditions of adversity; but convincing evidence is thus far lacking.⁵⁴

Although we do not know the extent to which internal inefficiency is a serious problem of monopoly, it could be substantially more serious than the more familiar problem of monopoly profits. Quite apart from our earlier point that monopoly profits may not deserve a great deal of worry, one should note that to incur an unnecessary expense wastes more of society's resources than jacking up price by the same amount in order to return investors' monopoly profits. By increasing price, the higher costs produce the same restriction of output as if the price had been inflated by an equivalent monopoly profit. Other than possibly restricting output, however, a monopoly profit merely transfers wealth from the buyer to the seller; society's stock of resources is not directly diminished. But money expended to hire more of the factors of production than actually needed to conduct a business diverts resources from more productive activities, and this effect is not only additive to, but could be many times greater than, the social cost in allocative inefficiency.⁵⁵

C. Failure To Optimize the Rate and Direction of Technological Change

Although technological progress has been enormously important in increasing the standard of living in advanced countries, we know relatively little about the market environment most conducive to such progress. Formidable difficulties in measuring technological progress and in disentangling multiple causes have made empirical study thus far inconclusive. We are remitted largely to theory.

Innovation exhibits several rather special characteristics. First, it is expensive; the costs of inventive activity, which are frequently substantial, must be incurred before—often long before—any revenues can be realized. Second, innovation is a risky activity for a private firm to undertake; both cost and success are difficult to predict. From these facts it follows that firms

53. See discussion in note 9 *supra*. In speaking of "cost" control in this context, I am, of course, distinguishing between those costs that constitute the managers' expense preferences (e.g., fancy offices), and all other costs. Management will by definition not seek to minimize the former category of costs, but it has every incentive to minimize all other costs; and we earlier saw that managerial expense preferences are very often not real costs at all but simply an indirect form of monopoly profit.

54. Assuming a relatively inelastic demand, the welfare loss occasioned by a price increase, due to market power, of 20 percent may be completely offset by a reduction in cost of 1 percent. See Williamson, *supra* note 50, at 22-23. Conversely, modest cost increases create more serious welfare losses than relatively large price increases arising solely from market power. See also Comanor & Leibenstein, *Allocative Efficiency, X-Efficiency and the Measurement of Welfare Losses* (to be published in *AM. ECON. REV. PAPERS & PROCEEDINGS* (May 1969)).

170

are unlikely to innovate unless the payoff from successful innovation is quite large. In addition, the high degree of uncertainty that characterizes inventive activity implies the importance of pursuing a number of diverse approaches toward the desired breakthrough, since any one is quite likely to fail. A third distinctive characteristic of invention is that its essence is knowledge. Once used, knowledge can readily be appropriated by others. The successful innovator may have difficulty in reaping private benefits equal to the social benefits of his work. If he cannot do so the rate of innovation may be suboptimal.

The foregoing factors define the essential elements of sustained and effective inventive activity by private firms: the resources to enable heavy expenses to be incurred well in advance of any possible payoff; the incentive to incur the costs and the risks of innovation, which in turn depends both on a large payoff if the benefits of the innovation can be appropriated by the inventor and a reasonable prospect that he will in fact be able to appropriate them; and a sufficient diversity of paths to breakthrough. To what extent are these conditions fulfilled in a monopolistic as compared to a competitive environment?

By virtue of enjoying monopoly profits, a monopolist at any given moment may have relatively more resources to devote to inventive activity than a firm whose profits are limited by competition. But possession of resources does not dictate their use for a particular purpose. Moreover, if a competitive firm has reason to anticipate that innovation will yield a substantial profit it should be able to raise the required funds in the capital market. Thus, if the prospects of innovation seem bright, both the monopolist and the competitive firm should be able to finance the necessary R & D, the former because it has, and the latter because it has access to, the necessary resources. This comparison seems a standoff.

At first blush, one might imagine that the competitive firm would have more to gain from successful innovation than a monopolist, and hence a greater incentive to innovate. An innovation that reduces the cost of a product sold under competitive conditions enables the innovator to reduce his price, and if by doing so he can drive out his competitors and obtain a monopoly of the market, he will be able to appropriate as monopoly profit a great deal of the extra value, above cost, that consumers attach to the product. The monopolist, in contrast, is presumably already capturing much of the consumers' surplus available in his market. A reduction in his costs would enable him only to capture some more.⁵⁵

55. I had thought, on a first reading, that this was the argument made by Arrow, *Economic Welfare and the Allocation of Resources for Invention*, in *THE RATE AND DIRECTION OF INVENTIVE ACTIVITY: ECONOMIC AND SOCIAL FACTORS* 609, 619-22 (Nat'l Bureau Econ. Research 1962). While the argument still seems to me an interesting one, I am now convinced that it is not the argument presented by Arrow. If I understand it correctly, his argument is that an inventor of a process that reduced the cost

This point requires qualification in two important respects. First, it is primarily applicable to innovations whose only consequence is to reduce the cost of the monopolist's existing product. A monopolist has an incentive equal to a competitive firm's so far as inventions applicable to markets not presently monopolized by him are concerned. And even within his monopolized markets he has a very strong incentive to product innovation; for if he develops a better product his demand curve may shift sharply to the right—that is, consumers may attach much greater value to what he produces than previously. If so he will be able to appropriate a good deal more consumers' surplus than before, much as a competitive firm could. Thus, before A. T. & T. laid the first undersea telephone cable in 1956, its international telephone service was not greatly in demand because the quality of radiotelephone service was poor and the capacity was limited. Innovation created a service that was much more valuable to the consumer and that in consequence began immediately to make substantial inroads into substitute services such as telegraphy.⁵⁶

Second, while the potential payoff from cost-reducing if not from product-improving innovations may be greater for the competitive firm than for the monopolist, the likelihood that the competitive firm can appropriate all or most of the potential gain is often less. Whatever gains accrue from a cost reduction in a monopoly market are securely the monopolist's. He has no rivals to cancel the gains by promptly imitating the innovation and adjusting price accordingly. The extent to which a competitive firm can appropriate the fruits of its inventive activity depends on whether and how long it can keep the innovation secret, how complete the protection obtainable under the patent laws is, and, failing either of these protections, how valuable a headstart over rivals proves to be. Secrecy is an uncertain protection and in many instances out of the question. Patent

of a product produced under competitive conditions could demand a royalty equal to a fraction less than the difference between the industry's former costs of production and its new, lower costs and that this royalty would exceed the additional profit that a monopolist of the same market would obtain from the same innovation. The reason why the monopolist's gain from innovation is smaller, however, is that the monopolist's output is deemed, by the principle that monopolists restrict output, to be smaller than that of the competitive industry. The same reduction in unit cost, applied to a smaller output, yields a smaller gain from innovation. Therefore, a monopolist will devote fewer resources to innovation than a competitive firm. But this is no more than a special case of the general proposition that a monopolized industry tends to use fewer resources—whether labor, capital, managerial, scientific, or whatever—than an equivalent industry that is competitive; in restricting output, the monopolist reduces his inputs. The implications of this familiar characteristic of monopoly for technological progress are unclear. If output is reduced in one industry because it is monopolized, it will be expanded in others as consumers shift their demand; if fewer resources are invested in innovation in one industry because it is monopolized, if output is reduced, and if therefore the gains from innovation are also reduced, one would expect more resources to be devoted to innovation in other industries, where output is now greater and the gains from innovation correspondingly increased. The overall level of inventive activity should not be greatly affected. This point is developed in a forthcoming article by Harold Demsetz in the *Journal of Law and Economics*.

56. See *Hearings on Merger of International Telegraph Carriers Before Senate Comm. on Interstate and Foreign Commerce*, 86th Cong., 1st Sess. 29-31 (1959).

29
151

"monopolies" cannot be equated with economic monopolies; not only is it frequently feasible to "invent around" a patent but patent rights may be costly to defend. And a headstart may or may not be a significant protection against prompt imitation by rivals. In sum, the competitive firm may have somewhat more dramatic prospects than the monopolist from a successful cost-reducing innovation, but it also has less assurance of realizing them.

The tradeoff becomes even more complex when we recall that a monopolist can subordinate short-term to long-term profit goals. Because of the enormous prestige of science and technology in this society—our almost religious veneration for material progress—a management not constrained by competition to minimize costs in the short run may attach high (perhaps excessive) importance to rapid and productive innovation as a matter of self-esteem and corporate image. The example of the Bell Telephone Laboratories, perhaps the foremost privately owned industrial laboratory in the world, indicates the dividends in public goodwill that a monopolist can obtain by supporting a substantial R & D effort. It should also be noted that a cost reduction permits a monopolist to reduce price without sacrificing profits. Indeed, profit maximization *requires* that the monopolist reduce his price when his costs decline, albeit not by the full amount of the cost reduction.⁵⁷ Consumers who see the price of a service falling (or the quality of the service improving) are unlikely to complain vociferously about monopoly prices and profits. That fact should make innovating highly attractive to a firm concerned with the long-run political viability of its monopoly.

57. This is best illustrated graphically: Under cost condition MC , the profit-maximizing monopolist will sell at price p , the price at which marginal revenue equals marginal cost. Suppose costs decline to MC' . If the monopolist remains at price p , his monopoly profit is the area $ApCD$; if he reduces price to p' , where his marginal revenue is equal to his new marginal cost, MC' , then his monopoly profit is the area $Ep'BD$. It can be demonstrated mathematically that this area is always larger than $ApCD$. However, if the demand curve is shifting to the right at the same time costs are decreasing, the new profit-maximizing price may be higher than the old.

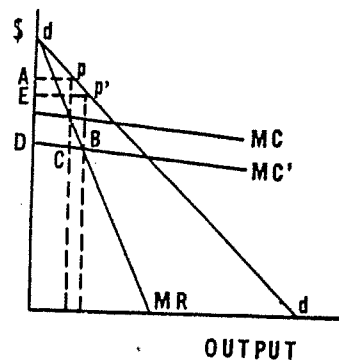


FIGURE 4

This conclusion may seem to contradict J. R. Hicks' well-known dictum, "The best of all monopoly profits is a quiet life."⁵⁸ A pure *ipse dixit* when offered, the remark has never been substantiated; I earlier expressed my doubt whether the image of the sated monopolist corresponds to reality.⁵⁹ In any event, one would suppose that a monopolist who wanted to enjoy a "quiet life" *would* innovate, sacrificing some immediate profits for long-run security against technological displacement. No natural monopoly can safely be assumed by owners or managers to be ordained to last forever, impervious to changes in technology and consumer taste. The monopolist must always reckon with the possibility of being supplanted as a result of technological change. It behooves him to anticipate such change through an active R & D program. And conducted on a substantial scale by a substantial firm, R & D is hardly so adventurous or unpredictable as to require a gambler's temperament. There is a good correlation between increased R & D expenditures and enhanced profitability,⁶⁰ and no reason, therefore, why a moderately cautious firm should be deterred from an adequate innovative effort.

On the other hand, concern for survival provides a greater incentive to rapid innovation for the competitive than for the monopoly firm. A firm that fails to anticipate a competitor's innovation may be destroyed, and although some monopolists have suffered grievously from the innovations of potential competitors (such as Western Union vis-à-vis the telephone companies) one would expect a monopolist to feel less concern about being preempted. In this respect, the motivation to innovate is stronger under competition than under monopoly.⁶¹

How do the diverse incentives of competitors and of monopolists net out? What market structure provides the greatest overall incentive to innovate? One can only guess at the answer. Before leaving the subject of incentives, I should mention the prevalent notion that a monopolist will not introduce an innovation as early as would a competitive firm for fear of being unable to recover its existing investment. It is true that a monopolist will not introduce a new process unless the total cost of the new is less than the marginal cost of the old. These are the respective current costs of the processes and that is the only comparison relevant in determining a firm's conduct; sunk costs are bygones. But the same principle governs the introduction of innovations in a competitive setting. A firm with a new

58. *Annual Survey of Economic Theory: The Theory of Monopoly*, 3 *ECONOMETRICA* 1, 8 (1935), reprinted in *READINGS IN PRICE THEORY* 361, 369 (Am. Econ. Ass'n 1952).

59. See text between notes 49 and 50 *supra*.

60. See E. MANFIELD, *THE ECONOMICS OF TECHNOLOGICAL CHANGE* 65-67, 106 (1968); E. MANFIELD, *INDUSTRIAL RESEARCH AND TECHNOLOGICAL INNOVATION—AN ECONOMETRIC ANALYSIS* 199-201, 203-04 (1968).

61. See Scherer, *Research and Development Resource Allocation Under Rivalry*, 81 *Q.J. ECON.* 359 (1967).

172
183

process will not introduce it unless its total cost is below the marginal cost of competitors using the old. Unless that condition is fulfilled the new process is not competitive, since as we saw in our railroad example a firm in pricing will ignore sunk costs if necessary to repel a new entrant.⁶² Hence one would expect the monopolist and the competitive firm to have identical incentives with respect to the timing of the introduction of new processes.

More troublesome than any supposed lack of incentive is the possible lack of diversity of approaches to technological breakthrough under monopoly. The process of research and development is to a significant extent one of trial and error. There is advantage, therefore, in the simultaneous pursuit of a variety of approaches to the desired end. Diversity may be difficult to achieve within a single firm due to the homogeneity of its personnel and the standardization of its procedures. Several firms quite different in organization and interests may achieve in the aggregate a more rapid rate of innovation than a single firm that spends the same amount on R & D as the several firms together.

While this is a forceful point, and is corroborated by what empirical studies we have,⁶³ it does not necessarily imply that innovation in an industry having a monopolistic structure will be suboptimal. As mentioned earlier, it is rare that an entire industry is a natural monopoly. A series of local or regional monopolists engaged in furnishing the same service should provide, therefore, some diversity of approach. Natural monopoly, moreover, is generally a phenomenon of distribution; manufacturers of the equipment used by the natural monopolist are an additional and very important source of diversity in inventive activity. The communications industry is atypical in the foregoing respects. Most of the regional or local monopolists at the distribution level are part of the Bell System, which also controls the major manufacturer of communications equipment, Western Electric. Even so, there is considerable diversity in the relevant R & D, since the concepts, components, and systems of modern communications are quite similar to those of the highly dynamic electronics, computer, and aerospace industries. This example illustrates the important principle that the relevant market in which to appraise the

62. See text accompanying notes 45-46 *supra*; Fellner, *The Influence of Market Structure on Technological Progress*, 65 Q.J. ECON. 556, 572-73 (1951), reprinted in *READINGS IN INDUSTRIAL ORGANIZATION AND PUBLIC POLICY* 277, 292-93 (Am. Econ. Ass'n 1958). A well-known empirical study of the electric-lamp industry is sometimes cited as supporting the proposition that a monopolistic firm will lack incentive to develop new products or processes when it is heavily committed to the old. The study indeed states: "The incentives of General Electric have not been so strong, however, for the rapid development of new light sources for general illumination, which would jeopardize its vested interest in the older incandescent lamp." A. BRIGHT, *THE ELECTRIC-LAMP INDUSTRY: TECHNOLOGICAL CHANGE AND ECONOMIC DEVELOPMENT FROM 1800 TO 1947*, at 455 (1949). This conclusion is not well supported by the study, however, or by other studies of the industry. See pages referenced in *id.* at 456; J. JEWKEE, D. SAWERS & R. STILLERMAN, *THE SOURCES OF INVENTION* 298-301 (1958).

63. See, e.g., J. JEWKEE, D. SAWERS & R. STILLERMAN, *supra* note 62, at 222, 246-47; Devons, *The Aircraft Industry*, in 2 *THE STRUCTURE OF BRITISH INDUSTRY* 45 (D. Burn ed. 1958).

diversity of innovative approaches is generally broader than the usual product or service market. Indeed, the more far-reaching the innovation, the more likely that it emerged in quite a different industry from the one in which it was first exploited commercially. Nylon was invented by the chemical, not the textile, industry. The transistor was invented by the telephone industry, not the radio or television or computer industries. The synchronous communications satellite was invented by the aerospace rather than by the communications industry.⁶⁴ The fact that an industry is a monopoly does not mean that only one firm is pursuing R & D in its technology.⁶⁵

The importance of external sources reinforces our earlier suggestion that a monopolist will feel pressure to innovate in order to forestall the emergence of competitors. An electronics firm engaged in research into the nature of electromagnetic radiation may discover a technique of communications that enables it to supplant a communications carrier in one of its markets. The carriers have every interest in anticipating such a development. The railroad industry would have benefited greatly from developing the truck and introducing trucking as an extension, rather than a competitor, of rail transportation.

A mainly theoretical analysis has yielded little support for the view that monopolists are on balance less likely to innovate than more competitive enterprises.⁶⁶ The picture is much the same when we turn to the empirical literature. The evidence to date yields no clear relationship between technological progressiveness and any particular kind of market structure.⁶⁷

64. This is a good example of the importance of diversity, even without competition. Both U.S. domestic long-haul communications and U.S. international satellite communications are monopolies, but of different companies—A.T.&T. and Comsat. A.T.&T. did not believe that synchronous satellites were feasible; Comsat did, and its judgment has been vindicated.

65. The importance of external sources is stressed by E. MANSFIELD, *THE ECONOMICS OF TECHNOLOGICAL CHANGE* 110-12 (1968).

66. Since, as noted earlier, the achievement of the least-cost level by monopolists may be more difficult than by competitive firms, see text following note 50 *supra*, it is quite possible that the diffusion of an innovation throughout an industry composed of a series of regional or local monopolists will be slower than in a competitive industry. That is not to say, however, that the state of the art is advanced less rapidly under monopoly, but only that it may take longer for all firms in a noncompetitive industry to take advantage of the inventive efforts of their most progressive counterparts or suppliers, as of other cost-saving opportunities.

67. This is the conclusion of a good recent survey of the empirical studies. E. MANSFIELD, *supra* note 65, at 215-17. To similar effect see D. HAMBERG, *R&D: ESSAYS ON THE ECONOMICS OF RESEARCH AND DEVELOPMENT* 68 (1966); R. NELSON, M. PECK & E. KALACHEK, *TECHNOLOGY, ECONOMIC GROWTH AND PUBLIC POLICY* 66-72 (1967); Brozen, *R&D Differences Among Industries*, in *ECONOMICS OF RESEARCH AND DEVELOPMENT* 83 (R. Tybout ed. 1965); Scherer, *Comment in id.* at 129. Most of the empirical studies have been comparisons of various degrees of oligopoly; there have been few studies of innovation by monopolists. Peck, *Inventions in the Postwar American Aluminum Industry*, in *THE RATE AND DIRECTION OF INVENTIVE ACTIVITY: ECONOMIC AND SOCIAL FACTORS* 279, 294 (Nat'l Bureau Econ. Research 1962), attributes the greater rate of invention in the aluminum industry after World War II to the replacement of Alcoa's monopoly by a three-firm oligopoly. He notes that Reynolds and Kaiser, although together only about the size of Alcoa, were both responsible for as many inventions; and he reasons that Alcoa would not have been thrice as inventive had it been twice as large. Perhaps not, but I have difficulty understanding the basis of the conjecture. Kendrick, in a study of productivity growth in the American economy, found that the regulated industries have done better than the national average. *Productivity Trends in the U.S. Private Economy and in the Public Utilities, 1948-1966*,

Handwritten signature and initials.

Very possibly other factors, such as absolute corporate size or the pattern of research outside of the corporate sector, are much more important to technological progress than the degree of competition. One should note, however, that if monopolistic industries are less progressive than competitive industries, the consequences could be quite serious. Technological change, which has created many valuable new products and often reduced the costs of existing products by entire orders of magnitude, is probably more important to the economic welfare of society than static efficiency, either allocative or internal.⁶⁶

D. Arbitrary Refusals To Serve, Inferior Goods and Service, and Unresponsiveness to Consumer Wants

One of the common beliefs about monopolists is that they are unresponsive to the consumer's desires because he has no choice—that they decline on capricious grounds to serve particular customers, are rude, and sell shoddy goods and provide poor service. The charge is seriously overstated, although it has a core of truth.

The argument that a monopolist is likely to be arrogant or capricious in his treatment of the consumer is supported by analogy to the treatment that minor governmental functionaries occasionally mete out to the hapless citizen. The charge makes more sense in the latter than in the former case. A minor functionary, protected in his job by political influence or civil service rules, may have nothing to gain from adopting a cooperative and polite attitude toward the members of the public with whom he deals, and he may derive psychological satisfaction from abusing them. A monopolist has a different set of incentives. The management of the Bell System or of the Pacific Gas & Electric Company can derive little psychic satisfaction from alienating customers and can fire any employees who do. If a monopolist arbitrarily refuses service to an individual, it not only gratuitously impairs public goodwill, but loses the profits that it would have obtained by serving him. A refusal to deal, therefore, is likely to reflect a yielding to powerful forces (such as intense and widespread racial prejudice in the community) that would be equally effective against competitive firms.

The notion that a monopolist will produce a less durable good than a competitive firm or render poorer service or otherwise degrade the quality of what he sells is true only in this limited sense: Since reduced quality usually means reduced cost, a firm that acquires a monopoly of a good or service formerly sold in competition may be able to make a monopoly

Apr. 24, 1968 (unpublished). This is far from conclusive, since these industries may have been the passive beneficiaries of the inventive efforts of others.

66. See R. NELSON, M. PECK & E. KALACHEK, *supra* note 67, at 16-18; E. MANSFIELD, *supra* note 65, at 4-5; cf. J. SCHUMPETER, *CAPITALISM, SOCIALISM AND DEMOCRACY* 83 (1942).

profit by holding price constant and reducing quality as well as by holding quality constant and increasing price. That is not to say, however, that the monopolist is indifferent to quality. The decision whether to degrade quality or increase price will be guided by the cost of different levels of quality and the value that the consumer attaches to them. If consumers in the aggregate will pay \$1,000 for widgets that cost \$600 to produce and \$1,100 for a better grade that costs the monopolist \$680, he will produce the better grade. If they will pay only \$1,079, indicating an unwillingness to shoulder the extra expense of the better product, the monopolist will be guided by that preference.⁶⁹

Far from being indifferent to quality, then, the monopolist has a strong incentive to determine consumers' reactions to various quality-price combinations. Nor is it cogent to argue that in the absence of competitive choice consumers' wants are difficult to gauge accurately. There is nothing to prevent a monopolist from probing them through the same devices used by competitive firms to develop new markets—market research, advertising, sales promotions, and test marketing. He has every incentive to be ingenious in anticipating and responding to consumers' wants.

E. Ruinous or Wasteful Competition

Thus far I have been discussing the equilibrium state of a natural monopoly market or industry: a single firm supplying the market's entire demand. Sometimes, however, several firms may find themselves in such a market. A market that once supported several firms, each operating at efficient scale, may, by reason of imperfectly anticipated technological change, become a natural monopoly before the firms (minus one) have made graceful exits. If these firms compete vigorously, competition will be short-lived. The most efficient firm will survive and the others fail or be acquired by it. If, however, by outright collusion or by adopting a policy of "live and let live" the firms in a natural monopoly market refrain from vigorous—and literally destructive—competition, production will persist at an inefficient scale, since by definition the most efficient way of supplying the market's entire demand is by one firm.

The possibility that more than one firm will find itself selling in a natural monopoly market is not, however, a substantial basis of concern about performance under natural monopoly. The situation is inherently unstable. Either there will be a brief flurry of fierce competition that leaves one firm in clear command of the field, forcing the others to merge with it or to fail, or—even more likely, one would think—there will be mergers without a period of fierce competition. The firms will realize that they can do much

69. See Stigler, *A Theory of Oligopoly*, 72 J. POL. ECON. 44, 61 (1964), reprinted in G. STIGLER, *THE ORGANIZATION OF INDUSTRY* 39, 62 (1968).

better by merging, operating at an efficient scale, and reaping monopoly profits than by either (1) competing to the death or (2) continuing multi-firm production at higher costs and lower profits than if they were consolidated.

If this reasoning is correct, there is not much substance to the conventional view, upon which regulation of the transportation industries is largely founded,⁷⁰ that in an industry where the economies of scale are substantial unregulated competition will cause chronic excess capacity leading to sustained and ruinous price wars. On the contrary, one would expect the firms in such an industry to consolidate their facilities and retire such capacity as was excess. So suggesting, I do not deny that railroads were overbuilt in the 19th century (partly as a result of governmental subsidy)⁷¹ or that there were rate wars. But apparently the wars were relatively few and sporadic,⁷² and the industry might have shaken down through consolidations had not the Supreme Court held railroad consolidations illegal *per se* under the Sherman Act.⁷³ A costly transition may have been inevitable. What is difficult to accept is that the situation would not eventually have corrected itself without government intervention.

As just implied, however, a possible problem in relying on natural market forces to match the number of firms in a market to the market's cost conditions is the antitrust laws. If as the Supreme Court has intimated economies of scale will not excuse a merger that may substantially lessen competition,⁷⁴ sellers in a natural monopoly market cannot lawfully merge until the brink of failure is reached. This impediment to efficiency could be eliminated by recognizing natural monopoly as a defense in a merger proceeding. Such a defense would complicate merger litigation, perhaps seriously, but only in a few cases. That may be a risk worth taking to avoid the serious and protracted inefficiencies that could result if the sellers in a natural monopoly market (who should be few enough to effectuate a policy of avoiding price competition without detectable collusion) decide not to em-

70. See note 2 *supra* and sources cited therein.

71. See, e.g., M. FAINSD, L. GORDON & J. PALAMOUNTAIN, *GOVERNMENT AND THE AMERICAN ECONOMY* 115-16 (3d ed. 1959).

72. E. TROXEL, *ECONOMICS OF TRANSPORT* 428-32, 656, 726 (1955). Moreover, most of the rate wars apparently were not caused by excess capacity. See P. MACAVOY, *THE ECONOMIC EFFECTS OF REGULATION: THE TRUNK-LINE RAILROAD CARTELS AND THE INTERSTATE COMMERCE COMMISSION BEFORE 1900*, at 195 n.3 (1965). For trenchant criticisms of the ruinous-competition theory see C. KATZEN & D. TURNER, *ANTITRUST POLICY* 196 (1959); Boies, *Experiment in Mercantilism: Minimum Rate Regulation by the Interstate Commerce Commission*, 68 COLUM. L. REV. 599, 660-63 (1968); Reynolds, *Cutthroat Competition*, 30 AM. ECON. REV. 736 (1940).

73. See *United States v. Southern Pac. Co.*, 259 U.S. 214 (1922); *United States v. Union Pac. Ry. Co.*, 226 U.S. 61 (1912); *Northern Sec. Co. v. United States*, 193 U.S. 197, 331 (1904); additional cases cited in M. CONANT, *RAILROAD MERGERS AND ABANDONMENTS* 47-48 (1964). See also *Louisville & N.R.R. v. Kentucky*, 161 U.S. 677 (1896); *Pearsall v. Great Northern Ry.*, 161 U.S. 646 (1896), where state statutes prohibiting railroad consolidations were sustained.

74. See *FTC v. Procter & Gamble Co.*, 386 U.S. 568, 580 (1967) (dictum); *Brown Shoe Co. v. United States*, 370 U.S. 294, 344 (1962) (dictum).

bark on the risky course of determining through price competition who shall survive. On the other hand, it may not be strictly necessary to recognize a formal defense. One can probably rely on the Department of Justice in the exercise of its enforcement discretion not to proceed in such a case, although the Department has not been explicit on the point.⁷⁵

F. Unfair Competition

One of the oldest complaints against monopoly is that a monopolist will annex a competitive market by using the monopoly profits from his other markets to subsidize a price that his competitors cannot meet because it is below cost. Recent studies, however, have cast doubt on whether "predatory price discrimination" is much of a danger.⁷⁶ Certainly if profit maximization is assumed to be the monopolist's strategy, predatory pricing is a tactic of questionable advantage in most cases. It requires the monopolist to forgo present profits in the hope that he will be able to charge a monopoly price in the competitive market (once he has monopolized it) that will more than recoup his earlier losses. But charging a monopoly price in a market that by definition has a competitive structure will attract new entrants, and the process of predation will have to be repeated indefinitely, with all the losses that the process entails. A possible exception is where entry into the market is difficult. Suppose that, due to economies of scale, a particular market will accommodate only three firms of efficient size. If one of those firms drives out the others, it may be able to raise its price somewhat without attracting entry by a new firm, because of the difficulty involved in large-scale entry. On the other hand, a structural characteristic such as economies of scale that makes entry into a market difficult may also affect the vigor of competition among the existing firms in the market. The three firms in our example might tacitly collude to keep their price as high as was possible without inducing new entry. If so, none of them could anticipate a higher profit rate by driving out the others. Indeed, a firm whose monopoly is not justified by economies of scale may be more vulnerable to the inroads of a new entrant than a cartel of smaller firms, for the single firm may encounter diseconomies of scale.⁷⁷ Furthermore, the very factors that create barriers to new entry may retard the exit of existing firms and make the process of monopolization by below-cost selling protracted and hence exceedingly expensive, while postponing the time at which recoupment can begin. And

75. The Department has publicly indicated that only in "exceptional circumstances" will it accept economies as a justification for a merger. Department of Justice Merger Guidelines, 1 TRADE REG. REP. ¶ 4430, at 6684-85 (1968).

76. See McGee, *Predatory Price Cutting: The Standard Oil (N.J.) Case*, 1 J. LAW & ECON. 137 (1958); Telser, *Cutthroat Competition and the Long Purse*, 9 J. LAW & ECON. 259, 267 (1966); Turner, *Conglomerate Mergers and Section 7 of the Clayton Act*, 78 HARV. L. REV. 4313, 1339-52 (1965).

77. See Williamson, *Hierarchical Control and Optimum Firm Size*, 75 J. POL. ECON. 123 (1967).

Handwritten signature/initials

when that time comes the monopolist may find that the barriers to entry on which it relied to enable it to charge a monopoly price without inducing prompt new entry have disappeared as the result of technological change. Such considerations would appear to make predatory price cutting a dubious tactic in virtually all circumstances.

If the monopolist believes that a monopoly of the competitive market, once achieved, will remain durable without resort to predatory tactics, his wisest course is to buy out the firms in the competitive market. He can well afford to pay the owners something more than the value of the firms as competitive enterprises—to wit, a share of the monopoly profits that he will enjoy when the market is monopolized (appropriately discounted to reflect the fact that they are prospective profits only). The owners will be eager to sell at a premium above the value of their enterprises in the existing competitive setting. What is striking about putting a monopoly together by the merger route—the classical⁷⁸ as well as the sensible way of monopolizing—is that it does not require that the monopolizer have a monopoly in some other market. All he need do is convince the sellers in the market that they will be better off if they eliminate competition by merging, and divide the resulting monopoly profits.

One might object that monopolization by merger would involve blatant violation of the antitrust laws, whereas monopolization by below-cost pricing, although equally an antitrust violation,⁷⁹ is harder to prove. A merger cannot be fudged; cost questions can. Hence, the very fact that we have a strict law against monopolization by merger may create an inducement to engage in predatory price discrimination that would not otherwise exist. However, the extent to which predatory pricing is possible with impunity is easily exaggerated. A course of pricing that led to an actual monopoly would surely provoke a searching investigation, laying the monopolist open to criminal and heavy civil penalties, including dissolution, as well as to treble-damage actions by his victims. The most recent decision of the Supreme Court in this area⁸⁰ has made the law exceedingly severe, possibly too severe. We may now have a set of sanctions that deters not only predatory conduct but some perfectly fair competition as well.

One could argue that the implicit threat of predatory conduct, even if never implemented or implemented so sporadically as to escape detection, should be enough to keep competitors in line and give the firm that monopolizes other markets considerable market power in the competitive

78. See McGee, *supra* note 76; *Investigation of the Telephone Industry in the United States*, H. Doc. No. 340, 76th Cong., 1st Sess. 139-43 (1939).

79. See Sherman Antitrust Act § 2, 15 U.S.C. § 2 (1964); Clayton Act § 2(a), *as amended*, 15 U.S.C. § 13(a) (1964); Robinson-Patman Antidiscrimination Act § 3, 15 U.S.C. § 13a (1964); Moore v. Mead's Fine Bread Co., 348 U.S. 115 (1954).

80. Utah Pie Co. v. Continental Baking Co., 386 U.S. 685 (1967).

market, although not a complete monopoly. If the previous analysis is correct, however, the threat would lack credibility. The competitive firms would know that a rational monopolist would not pursue predation seriously. Why should their conduct be affected by a bluff?

Thus far, we have assumed that the monopolist is a strict profit maximizer. If not, he has unexploited monopoly power that he could use to support below-cost selling in competitive markets without impairing his chosen return. But a firm that chooses not to maximize profits is presumably maximizing some other preference that it would have to forgo in order to engage in predatory price cutting (unless the firm is completely passive, which hardly seems compatible with aggression on any front). A significant exception would be a firm seeking to maximize sales or growth.⁸¹ However, of all the ways to build sales, one would imagine predatory price discrimination to be among the least attractive to a monopolist able and inclined (as the sales-maximization hypothesis assumes) to subordinate short-run profit maximization to more strategic goals of corporate and personal gain. Selling below cost is bound to be unpopular, to say the least, among the firms inhabiting the market; it can only draw attention to the monopolist's size and power; and, to repeat, it is both a civil and (if the requisite intent is proved) a criminal violation of the antitrust laws.

At all events, it is striking how few substantiated incidents of predatory pricing have turned up in the annals of American business.⁸² With the empirical and theoretical foundations of the fear of predatory price discrimination so thin, and given that we already have strict laws on the books forbidding the practice, it hardly seems to warrant additional regulation.

Another weapon of a monopolist bent on aggrandizement, besides monopoly profits, is buying power. If a monopolist is also a monopsonist (the sole consumer of a product), and if he can establish his own supply facilities, then he is in a position to monopolize the supply market. But it is unclear why he would want a monopoly of supply, when by hypothesis the only customer who could be exploited by such a monopoly would be himself. A wiser tactic for him would be to use his buying power to drive highly favorable terms with the existing suppliers, to ensure, in other words, that he is not exploited. In some cases, indeed, the monopsonist may be able to obtain inputs below the competitive price. This would happen if the monopsony arose after the suppliers had committed resources to the market that were not readily transferable to other uses. It is possible to demonstrate

81. See text accompanying note 14 *supra*.

82. For a review of the evidence see Telser, *supra* note 76, at 268-70. To be sure, difficulty of detection may have something to do with this. Moreover, before there was settled law against monopolization by merger, firms could be expected to follow that, the less costly, route. See text accompanying note 78 *supra*.

that a monopsony price, the obverse of the monopoly price, is equally undesirable from the standpoint of efficient allocation. But it does not follow that monopsony pricing is a general problem of monopoly. A local electrical company is a monopolist, but it competes with many other firms both inside and outside the electric-utility industry in the purchase of its supplies. In general, monopoly does not confer monopsony power, and monopsony power may exist independently of monopoly. They are separate problems.

Exploitation of buying power and predatory price discrimination are only two examples of a wide variety of unfair business practices commonly associated with monopoly. Examples of others are vertical integration, patent abuses, tying arrangements, and refusals to deal with potential competitors. Such practices have frequently cropped up in the regulated industries. Examples are the railroads' refusal to carry piggyback vans tendered by motor carriers at the same rates as those tendered by ordinary shippers and the telephone carriers' refusal to permit the attachment to their lines of terminal or interconnection equipment not supplied by them.⁸³ But none of these practices is uniquely associated with monopolists, and all are within the conventional scope of general antitrust and trade-regulation law.⁸⁴

G. The Political Dimension of the Monopoly Problem

Opposition to monopoly is frequently premised on political grounds. Private economic power, epitomized by the monopolist, is thought to endanger democratic processes. Basically, however, the objection is to large firms rather than to monopolists as such. General Dynamics has more power in any sense relevant to the political process than the independent telephone company that serves Rochester, New York, but the latter is a monopolist and the former is not. Scale—the number of workers, managers, shareholders, suppliers, distributors, creditors, and other dependents or potential allies that a firm has—would appear to be a far more important determinant of the firm's political weight than whether it enjoys a monopoly, that is, whether its stockholders (and perhaps managers) receive unjustified returns or its retained earnings are abnormal. The Bell System is the classic instance of a firm that both is a monopolist and is so large in absolute size⁸⁵ as to raise the question whether it may not enjoy undue influence in legislative and other political arenas. But one imagines that

83. Both restrictions were recently voided. See *American Trucking Ass'n v. Atchison, T. & S.F. Ry.*, 387 U.S. 397 (1967); *Use of the Carterfone Device*, 13 F.C.C.2d 420, on reconsideration, 14 F.C.C.2d 571 (1968).

84. See Clayton Act §§ 2, 3, 7, 8, as amended, 15 U.S.C. §§ 13, 14, 18, 19 (1964); Federal Trade Commission Act § 5, as amended, 15 U.S.C. § 45 (1964); Sherman Antitrust Act §§ 1-8, as amended, 15 U.S.C. §§ 1-7 (1964).

85. The Bell System is the largest private corporation in the world, with annual revenues amounting to \$13 billion in 1967. *FORTUNE*, June 15, 1968, at 217.

such political power as it may possess resides in its scale of operations and thus that profit controls would not make much difference.

H. Managerial Incompetence

A fair summation of the discussion thus far is that if the management of a firm that enjoys a natural monopoly is reasonably competent, one cannot assert with any confidence that performance is likely to fall greatly short of our economic or social objectives.⁸⁶ The traditional economic objection to monopoly—that it leads to suboptimal output of the monopolized product—has a core of validity as applied to the natural monopolist, but there are, as we saw, a good many reasons for questioning whether the allocative effects of unregulated natural monopoly are in fact likely to be serious. In addition, the management of such a monopoly will have strong incentives to press cost reduction and to innovate. And contrary to popular myth a monopolist is not likely to abuse the public, project its monopoly into competitive markets, or enjoy disproportionate political power. The “stick” of competition, as well as the diversity of approach that the existence of competitors would assure, will be lacking; but there is no convincing basis for the view that performance will be markedly affected thereby, although internal inefficiency may be something of a problem. Nor are the distributive effects of natural monopoly profits demonstrably a cause for serious practical concern or moral condemnation.

It is always possible, however, that the management of a monopolist will be incompetent—that it will make foolish mistakes harmful both to the consumer and to the stockholder (such as selling above the monopoly price), but not so obviously foolish as to invite a proxy fight. Under competition, managerial incompetence is not a social problem. The firm with persistently inferior management will simply fail. There is no such automatic corrective in the case of the monopolist. Substitute services eventually may make such inroads as to awaken the owners to the existence of a managerial problem, but even then it may be difficult to determine whether poor management or exogenous factors were responsible since there are no exactly comparable firms as there would be under competition. Thus, the steady decline of telegraph service led to charges that the management of the Western Union Telegraph Company was incompetent; the reply was that telegraphy had simply been outdistanced by telephony and that the trend could not have been reversed by the shrewdest of managements.⁸⁷

86. It is interesting to note that a careful full-length study of Alcoa's prewar monopoly failed to uncover sufficient evidence of suboptimal performance to justify, in the author's view, any recommendation for governmental action. D. WALLACE, *MARKET CONTROL IN THE ALUMINUM INDUSTRY* 352-53, 365 (1937). A detailed study of United Shoe Machinery Corporation's monopoly found considerable price discrimination but no persuasive indication that the company was less progressive than it would have been under competition. C. KATZEN, *UNITED STATES v. UNITED SHOE MACHINERY CORPORATION: AN ECONOMIC ANALYSIS OF AN ANTI-TRUST CASE* 207-08 (1956).

87. See text accompanying notes 143-44 *infra*.

Handwritten signature and initials