ECONOMICS OF ANTITRUST: TRANSACTION COST CONSIDERATIONS*

Oliver E. Williamson†

Economic analysis is commonly, though somewhat arbitrarily, divided into macroeconomic and microeconomic categories. The former is concerned with highly aggregative economic issues—such as national income, employment, and inflation—while the latter deals with the behavior of individual consumers, firms, and markets. To the extent that economics is thought to have a bearing on antitrust analysis and policy, the firm and market models of received microtheory are thought, by economists and lawyers alike, to supply the relevant foundations.¹

Although I am in general agreement with this position, I contend that received microtheory sometimes needs to be augmented by introducing transaction cost considerations. Failure or refusal to make allowance for transaction costs, in circumstances where these are arguably nonnegligible, can lead to error. Not only is an understanding of the issues impaired, but incorrect policy prescriptions will sometimes result.

One of the attractive attributes of the transaction cost approach² is that it reduces, essentially, to a study of contract-

---

* Copyright © reserved by the author.
† Professor of Economics, Law and Public Policy, University of Pennsylvania. S.B. 1955, Massachusetts Institute of Technology; M.B.A. 1960, Stanford University; Ph.D. 1963, Carnegie-Mellon. Research on this Article has been supported by a grant from the National Science Foundation. Parts of the Article are based on a lecture given at the University of California at Los Angeles on January 10, 1974 in conjunction with the Major Issue Lecture Series program "Large Scale Enterprise in a Changing Society."

¹ To be sure, these models are sometimes tailored before applying them to particular antitrust problems. Examples of the application of received microtheory to antitrust issues are E. Singer, ANTITRUST ECONOMICS (1968); Baxter, Legal Restrictions on Exploitation of the Patent Monopoly: An Economic Analysis, 76 Yale L.J. 267 (1966); Williamson, Wage Rates as a Barrier to Entry: The Pennington Case in Perspective, 82 Q.J. Econ. 85 (1968).

² More generally, the issues posed involve an assessment of markets and hierarchies. For a discussion of these issues, see Williamson, Markets and Hierarchies: Some Elementary Considerations, 63 Am. Econ. Rev., May 1973, at 316; Williamson, Markets and Hierarchies: Analysis and Antitrust Implications (Aug. 1973) (unpublished paper held by author). This approach is similar to that advocated by J. Commons, INSTITUTIONAL ECONOMICS (1934), who took the position that the transaction
ing—which means that the contracting expertise of lawyers developed in other contexts can be drawn upon. Issues such as the following are addressed: When will a related set of transactions be completed most efficaciously by negotiating contracts between firms (across a market), and when will merger or integration (internal organization) be preferred? In what respects, if any, do pre-existing firm and market structures impede or facilitate the ability of new firms to negotiate the necessary market contracts for labor, capital, materials, and intermediate products to effectuate successful entry? While these types of issues can be addressed in an unconvoluted way using the transaction cost apparatus, the models of received microtheory, in which transaction costs are suppressed, are often ill-suited and sometimes misleading.

I begin with a brief review of received microtheory before setting out the elements of the transaction cost approach. The examination of vertical integration, oligopoly, and conglomerate organization from the transactional point of view suggests antitrust policies somewhat different from those advanced by scholars employing the conventional microtheory approach. Although not exhausting the applications of the transaction cost approach to the study of antitrust issues,3 the Article will hopefully give the reader a sense of the relevance of this approach to the antitrust area.

I. The Basic Approaches

It is widely thought that “the economic background required for understanding antitrust issues seldom requires detailed mastery of economic refinements”4—meaning, presumably, that the standard economic models of firms and markets found in intermediate microtheory textbooks will normally be sufficient for antitrust purposes. I doubt that this is the case. Conventional analysis sometimes needs to be augmented and at other times supplanted by express consideration of transactional problems.

constituted the ultimate unit of investigation. Commons, however, had to fashion many of his transactional concepts himself, while I am able to draw, forty years later, on much more extensive literatures in both economics and organization theory. This is a considerable advantage.

3 For example, the marketing practices of Arnold, Schwinn & Co., to which the Antitrust Division objected, can usefully be examined in transaction cost terms. See United States v. Arnold, Schwinn & Co., 388 U.S. 365 (1967).

A. Received Microtheory

Demand curves (average revenue curves), average cost curves, and the marginal curves of revenue and cost drawn to each of these constitute the basic modeling apparatus for most antitrust treatments of firms and markets. Implicit in this model are efficiency assumptions of two kinds. First, it is assumed that the firm realizes the maximum output of product from each feasible combination of factor inputs (mainly labor and capital). That is, it operates on its production function. Failure to operate on the production function would imply wasteful use of inputs; this is assumed away. Second, given the prices of productive factors, it is assumed that the firm chooses the least-cost combination of factors for each possible level of output. The total cost curve, from which average and marginal cost curves are derived, is constructed in this way.

In circumstances where economies of scale are large in relation to the size of the market, a condition of natural monopoly (or perhaps oligopoly) may be said to exist. The monopolist or the oligopolists who supply goods and services in such a market will be sufficiently large that small percentage changes in their output will perceptibly affect the market price. Price is thus subject to strategic determination. However, in circumstances where economies of scale are exhausted at firm sizes that are small in relation to the market, each firm will regard price as given and a condition of competitive market supply, in which price will be equal to marginal cost, will obtain.

Intermediate types of markets, such as duopoly or oligopoly, are modeled by making appropriate assumptions about the nature of the technology and the interfirm relations which develop. Depending on the underlying technology and the

---

5 In setting out what I think to be the main distinctions between the conventional and transactional approaches, I concede at the outset that my discussion of received microtheory is highly simplified. It is the theory of the firm as that appears in the conventional intermediate price theory textbook. Inasmuch as I often find such a tactic to be a source of considerable irritation when reviewing the work of others who study the behavior of the modern corporation, I resort to it with some reluctance. My defense is that the simplified presentation is an economical way to expose the issues.

6 This assumes that the firms in question behave in an independent (noncollusive) manner.

7 Among the leading types of models for these purposes are Cournot models and their variants, and entry barrier models, which make allowance for potential competition. For an elegant review and extension of Cournot models, see L. Telser,
behavioral assumptions that are employed, the prices and outputs that will be associated with alternative market structures can be succinctly derived. The social welfare implications of each, moreover, can be established by characterizing the benefits and costs resulting from the structure in question in appropriate social welfare terms.\(^8\) The types of tradeoffs that antitrust must contend with in circumstances where monopoly power and production economies both obtain can then be displayed in a relatively straight-forward manner.\(^9\)

Implicit throughout most analyses of this kind is that the nature of the firm—with respect, for example, to what it will make and what it will buy—is simply taken as given. Matters of internal organization (hierarchical structure, internal control processes) are likewise ignored. The firm is thereby reduced to little more than a production function to which a profit maximization objective has been assigned. That many interesting problems of firms and markets are suppressed or neglected as a result should come, perhaps, as no surprise.

B. The Transaction Cost Approach\(^10\)

The transactional approach may be stated compactly as follows: (1) markets and firms are alternative instruments for completing a related set of transactions; (2) whether a set of transactions ought to be executed between firms (across markets) or within a firm depends on the relative efficiency of each mode; (3) the costs of writing and executing complex contracts across a market vary with the characteristics of the human decisionmakers who are involved with the transaction on the one hand, and the objective properties of the market on the other; (4) although the human and transactional factors which impede exchanges between firms (across a market) manifest themselves somewhat differently within the firm,

---


the same set of factors applies to both. A symmetrical analysis of trading, therefore, requires that the transactional limits of internal organization as well as the transactional sources of market failure be acknowledged. Moreover, just as market structure matters in assessing the efficacy of trades in the marketplace, so internal structure matters in assessing internal organization.

The transaction cost approach is interdisciplinary, drawing extensively on contributions from both economics and organization theory. The market failure, contingent claims contracting, and recent organizational design literatures supply the requisite economic background. The administrative man and strategic behavior literatures are the main contributions from organization theory.

With this basis the transaction cost approach attempts to identify a set of market or transactional factors which together with a related set of human factors explain the circumstances under which complex contracts involving contingent claims will be costly to write, execute, and enforce. Faced with such difficulties, and considering the risks that simple, and therefore incomplete, contingent claims contracts pose, the firm may decide to bypass the market and resort to hierarchical modes of organization. Transactions that might otherwise be handled in the market would then be performed internally and governed by administrative processes.

Uncertainty and small numbers exchange relations, in which one party's choice of trading partners is restricted, are the transactional factors to which market failure is ascribed. Unless joined by a related set of human factors, however,

---

15 See, e.g., E. Goffman, Strategic Interaction (1969).
16 This is merely a necessary but not sufficient condition for internal organization to supplant the market. Internal organization also experiences distortion. Shifting a transaction from the market to a firm requires that a net efficiency gain be shown.
such transactional conditions need not impede market exchange. The pairing of uncertainty with bounded rationality and the joining of small numbers with what I will refer to as opportunism are especially important.

Consider first the pairing of bounded rationality with uncertainty. The principle of bounded rationality has been defined by Herbert Simon as follows: "The capacity of the human mind for formulating and solving complex problems is very small compared with the size of the problems whose solution is required for objectively rational behavior in the real world . . . ."\(^{17}\) It refers both to neurophysiological limits on the capacity to receive, store, retrieve, and process information without error\(^{18}\) and to definitional limits inherent in language. If these limits make it very costly or impossible to identify future contingencies and to specify, *ex ante*, appropriate adaptations thereto, long term contracts may be supplanted by internal organization. Recourse

---


\(^{18}\) The implications for contractual purposes of joining bounded rationality with uncertainty are suggested by the following description of the decision process:

For even moderately complex problems . . . the entire decision tree cannot be generated. There are several reasons why this is so: one is the size of the tree. The number of alternative paths in complex decision problems is very large. . . . A second reason is that in most decision situations, unlike chess, neither the alternative paths nor a rule for generating them is available. . . . A third reason is the problem of estimating consequences. . . . For many problems, consequences of alternatives are difficult, if not impossible, to estimate. The comprehensive decision model is not feasible for most interesting decision problems.

Feldman & Kanter *Organizational Decision-Making* in *Handbook of Organizations* 615 (J. March ed. 1965). The infeasibility, or prohibitive cost, of describing the comprehensive decision tree and making *ex ante* optimal choices at every node means that collusive agreements must, except in implausibly simple circumstances, be highly incomplete documents.

A specific illustration of bounded rationality in the large corporation is afforded by the statement of R.H. Davies, President of Electric Autolite Company at the time of the Ford-Autolite merger. He testified on deposition as follows:

. . . Electric Autolite was "concerned" because, when Champion Spark Plug Company "went public" in 1958, "the figures that came out were very large—showing very large profits" and "when Ford saw those figures and saw how much profit there was in it" Electric Autolite "felt" that "the very essence of that much profit going to a supplier would be enough to make Ford think in terms of integration."

Trial Memorandum for Defendant Ford Motor Co. at 14-15, United States v. Ford Motor Co., 286 F. Supp. 407, 435 (E.D. Mich. 1968) (violation of Clayton Act found), 315 F. Supp. 372 (E.D. Mich. 1970) (divestiture ordered), aff'd, 405 U.S. 562 (1972). The example is interesting because it suggests that as large and successful a firm as the Ford Motor Company, with its staff of engineers, cost accountants, and financial analysts, failed to discern the underlying profitability of spark plug manufacture until Champion went public (Champion was the first spark plug firm to go public). In a world of unbounded rationality, such disclosure would be unnecessary to stimulate Ford's interest.
to the internal organization of transactions permits adaptations to uncertainty to be accomplished by administrative processes as each problem arises. Thus, rather than attempt to anticipate all possible contingencies from the outset, the future is permitted to unfold. Internal organization in this way economizes on the bounded rationality attributes of decisionmakers in circumstances where prices are not "sufficient statistics" and uncertainty is substantial.

Rather, however, than resort to internal organization when long term contingent claims contracts are thought to be defective (too costly or perhaps infeasible), why not employ short term contracts instead? Appropriate adaptations to changing market circumstances can then be introduced at the contract renewal interval, thereby avoiding the prohibitive costs of ex ante specification. The pairing of opportunism with small numbers exchange relations, however, creates other obstacles to market transactions.

Developing this set of issues is somewhat involved and the interested reader is referred to discussions elsewhere of the types of contracting problems that give rise to vertical integration. Suffice it to observe here that (1) opportunism refers to a lack of candor or honesty in transactions, to include self-interest seeking with guile; (2) opportunistic inclinations pose little risk to trading partners as long as competitive (large numbers) exchange relations obtain; (3) many transactions which at the outset involve a large number of qualified bidders are transformed in the process of contract execution—often because of economies of scale and accrued cost-advantages attributable to successful bidders learning more about the job as they perform their work (learning by doing)—so that a small numbers supply condition effectively obtains at the contract renewal interval; and (4) short term contracting is costly and risky when opportunism and small numbers relations are joined. The argument will be developed further in other sections of this Article.

In consideration of the problems that both long and short term contracts are subject to—by reason of bounded rationality and uncertainty in the first instance and the pairing of  

19 In circumstances, however, where prices are sufficient statistics, see T. Koopmans, Three Essays on the State of Economics 41-54 (1957), reliance on the price system serves to economize on bounded rationality. See Hayek, The Use of Knowledge in Society, 35 AM. ECON. REV. 519 (1945).

20 Williamson, supra note 10.
opportunism with small numbers relations in the second—internal organization may be used instead. With internal organization, issues are handled as they arise rather than in an exhaustive contingent planning fashion from the outset.\textsuperscript{21} The resulting adaptive, sequential decisionmaking process is the internal organizational counterpart of short term contracting and serves to economize on bounded rationality. That opportunism does not pose the same difficulties for such internal, sequential supply relations that it does when negotiations take place across a market is because (1) internal divisions do not have preemptive claims on profit streams, but act under common ownership and supervision to more nearly maximize joint profits instead, and (2) the internal incentive and control machinery is much more extensive and refined than that which obtains in market exchanges.\textsuperscript{22} The firm is thereby better able to take the long view for investment purposes (and hence is more prepared to put specialized plant and equipment in place) while simultaneously adjusting to changing market circumstances in an adaptive, sequential manner.

Having said this, I hasten to add that if internal organization serves frequently to attenuate bounded rationality and opportunism problems, it does not eliminate either condition. Of special relevance in this connection are two propositions: (1) the limitations of internal organization in both bounded rationality and opportunistic respects vary directly with firm size, organization form held constant,\textsuperscript{23} but (2) organization form—that is, the way in which activities in the firm are hierarchically structured—matters.\textsuperscript{24} The import of this latter proposition is developed in the discussion of conglomerates in section IV.

\textsuperscript{21} This is oversimple. Internal organization also provides for contingencies by developing what are referred to as “performance programs,” which are sometimes quite elaborate. Such programs are more easily adapted to unforeseen contingencies than are interfirm contracts, for the reasons given in the text. For a discussion of performance programs, see J. March & H. Simon, Organizations (1958).

\textsuperscript{22} O. Williamson, Corporate Control and Business Behavior 120-35 (1970). Internal organization affords two further benefits: it helps to overcome conditions where one party holds information not available to the other without some expense (information impactedness), because internal audits are more powerful than external, and is sometimes able to reduce uncertainty by promoting convergent expectations. Both of these are important but less basic to the present argument than the effects of internal organization on bounded rationality and opportunism.

\textsuperscript{23} See Williamson, supra note 14.

\textsuperscript{24} See A. Chandler, Strategy and Structure (1962); Williamson, supra note 22.
Moreover the choice between firm and market ought not to be regarded as fixed. Both firms and markets change over time in ways which may render an initial assignment of transactions to firm or market inappropriate. The degree of uncertainty associated with the transactions in question may diminish; market growth may support large numbers of suppliers in competition with one another, and information disparities between the parties often shrink. Also, changes in technology may occur, altering the degree to which bounded rationality limits apply. Thus, the efficacy of completing transactions by hierarchies or markets should be reassessed periodically.

C. An Example: Price Discrimination

The differences between received microtheory and the transaction cost approach can be illustrated by examining the familiar problem of price discrimination. As will be evident, the transaction cost approach does not abandon but rather augments the received microtheory model.

Assume for this illustration that the market in question is one in which economies of scale are large in relation to the size of the market, in which case the average cost curve falls over a considerable output range. Assume, in particular, that demand and cost conditions are as shown in Figure 1.
The unregulated monopolist who both maximizes profits and sells his output at a single, uniform price to all customers will restrict output below the social optimum\(^2\) (shown by \(Q^*\) in Figure 1)\(^2\) at which marginal cost equals price. Instead, the monopolist will produce only to the point \((Q_m)\) at which marginal cost equals marginal revenue so that an excess of price over marginal cost obtains.

It is sometimes argued, however, that price discrimination will correct the allocative efficiency distortion in a monopoly situation. The monopolist who can segregate his market in such a way that each customer is made to pay his full valuation (given by the demand curve) for each unit of output has the incentive to add successive units of output until the price paid for the last item sold just equals the marginal cost. The fully discriminating monopolist will thus be led to expand output from the restricted position of a nondiscriminating monopolist \((Q_m)\) to the social optimum point \((Q^*)\). Although income distribution will be affected in the process (in possibly objectionable ways), the output distortion is removed and an allocative efficiency gain is realized.\(^2\)

Evaluating this allocative efficiency claim gives us our first opportunity to contrast the conventional analysis of received microtheory with a transactions cost approach. Implicit in the above conventional microtheory argument is an assumption that the costs of both discovering true customer valuations for the product and of enforcing restrictions against resale (so that there can be no arbitrage) are negligible and can be disregarded. Such costs vanish, however, only if either (1) customers will honestly self-reveal preferences and self-enforce nonresale promises (no opportunism) or (2) the seller is omniscient (a strong variety of unbounded rationality). Inasmuch as assumptions of both kinds are plainly unrealistic, the question naturally arises: Is there an allocative efficiency gain if nontrivial transaction costs must be incurred to discover true customer valuations and/or to police nonresale restrictions? Unfortunately for received microtheory, the outcome is uncertain.

\(^2\) So-called "second best" issues are assumed away here and throughout the article.

\(^2\) So that a break-even problem will not be posed if output is set at \(Q^*\), I assume that scale economies are exhausted before this output is reached.

\(^2\) If the output of the industry in question is used as an intermediate rather than strictly as a final product, factor distortions at other stages of production may be induced. See McKenzie, Ideal Output and the Interdependence of Firms, 61 Econ. J. 785 (1951). For simplicity, let these be assumed away.
tain when these transaction costs are introduced.

To see this, assume (for simplicity), that the transaction costs of accomplishing full price discrimination are independent of the level of output: the costs are either zero, in which event no effort to price discriminate is made, or T, in which case customer valuations become fully known and enforcement against cheating is complete.\(^2\) Price discrimination will of course be attractive to the monopolist if a net profit gain can be shown—which will obtain if the additional revenues (which are given by the two shaded regions, \(A_1\) and \(A_2\), in Figure 1) exceed the costs of achieving discrimination, \(T\). What is interesting for social welfare evaluation purposes is that an incremental gross welfare gain is realized only on output that exceeds \(Q_m\). This gain is given by the lower triangle \((A_2)\). Consequently the net social welfare effects will be positive only if \(A_2\) exceeds the transaction costs, \(T\). An allocative efficiency loss, occasioned by high transaction costs, but a private monopoly gain, derived from price discrimination applied to the output that would have been produced even without discrimination (this revenue gain being shown by \(A_1\)), is therefore consistent with fully discriminatory pricing in circumstances where non-trivial transaction costs are incurred in reaching the discriminatory result. More precisely, if \(T\) is less than \(A_1 + A_2\) but more than \(A_2\) alone, the monopolist will be prepared to incur the customer information and policing costs necessary to achieve the discriminatory outcome, because his profits will be augmented \((A_1 + A_2 > T)\), but these same expenditures will give rise to a net social welfare loss \((A_2 < T)\).\(^{29}\)

In circumstances where \(T\) is zero or negligible, of course, this contradiction does not arise. But the results of received microtheory rest crucially on such an assumption. If, arguably, the assumption is not satisfied, transaction costs need expressly to be taken into account before a welfare assessment is ventured.

---

\(^{28}\) Generalizing the analysis by expressing the transaction costs of discerning true customer valuations and policing resale restrictions as a continuous function of output is relatively easy but yields little that the simplified assumptions do not. (One difference to be noted is that the price discriminating output will be less than the social optimum, \(Q^*\).) The analysis can likewise be generalized to make the degree of precision of price discrimination a decision variable.

\(^{29}\) The discussion in the text assumes, implicitly, that the uniform pricing monopolist can price at \(P_m\) without inducing entry. If, however, the entry forestalling price (\(P\)) is less than \(P_m\), the initial position to be evaluated is a larger output and lower price than that discussed above. For fixed \(T\), the welfare gains of price discrimination are further reduced. (In all likelihood an entry threat will attenuate the private gains as well.)
II. Vertical Integration

The discussion of market exchange versus internal organization in the preceding section suggests that internal organization has attractive properties in circumstances where long term contracts are not feasible, because contractual contingencies overwhelm the limited planning capacities of parties subject to bounded rationality, and where short term contracts pose hazards, because of the conjunction of opportunism with a small numbers exchange condition. Prospective interfirm contracting difficulties are thus responsible for the decision to integrate. The details of such a transactional approach to vertical integration have been worked out elsewhere.30

It should be appreciated, however, that this has not been the prevailing rationale for vertical integration among economists. More often the argument runs in terms of technological considerations. Two such arguments are examined below and are rejected in favor of the transactional approach. The possibility that vertical integration might inhibit potential entry is then explored and the incentive to integrate as a means by which to circumvent government controls (taxes, quotas) is briefly treated. I conclude this section with a statement of the antitrust enforcement implications of the argument.

A. A Life Cycle Analysis

George Stigler deduces, from his explication of Adam Smith's theorem that the division of labor is limited by the extent of the market, that vertical integration is related to an industry’s life cycle: vertical integration will be extensive in firms in young industries; disintegration will be observed as an industry grows; and reintegration will occur as an industry passes into decline.31 These life cycle effects are illustrated by reference to a multiprocess product in which each process involves a separable technology and hence its own distinct cost function.32 Some of the processes display falling cost curves, others curves that rise continuously, and still others U-shaped cost curves.

Stigler then asks, why does the firm not exploit the decreasing cost activities by expanding them to become a monopoly?

30 Williamson, supra note 10.
31 Stigler, The Division of Labor is Limited by the Extent of the Market, 59 J. Pol. Econ. 185 (1951).
32 Stigler employs the separability assumption for convenience; relaxing it complicates but does not alter the general argument.
He answers by observing that, at the outset, the decreasing cost functions may be “too small to support a specialized firm or firms.” But, unless the argument is meant to be restricted to global or local monopolies, for which there is no indication, resort to a specialized firm does not exhaust the possibilities. Assuming that there are at least several firms in the business, why does one of them not exploit the available economies, to the mutual benefit of all the parties, by producing the entire requirement for the group? The reasons, I submit, turn on transaction costs inherent in interfim rivalry.

If, for example, the exchange of specialized information between the parties is involved (Stigler specifically refers to “market information” as one of the decreasing cost possibilities) strategic misrepresentation issues are posed. The risk here is that the specialist firm will disclose information to its rivals in an incomplete and distorted manner. Because the party buying the information can establish its accuracy only at great cost, possibly by collecting the original data itself, the exchange fails to go through. If, however, rivals were not given to opportunism, the risk of strategic distortion would vanish and the technologically efficient specialization of information could proceed.

The exchange of physical components that experience decreasing costs is likewise discouraged where both long term and short term contracts incur prospective transactional difficulties. Long term contracts are principally impeded by bounded rationality considerations: the extent to which uncertain future events can be expressly taken into account—in the sense that the cost of appropriate adaptations can be estimated and contractually specified—is simply limited. Since, given opportunism, incomplete long-term contracts predictably pose interest conflicts between the parties, other arrangements are apt to be sought.

Spot market (short term) contracting is an obvious alternative. Such contracts, however, are hazardous if there are only a small number of suppliers, which (by assumption) holds true for the circumstances described by Stigler. The buyer then incurs the risk that the purchased product or service will, at some later time, be supplied under monopolistic terms. Industry growth, moreover, need not eliminate the tension of small numbers bargaining if the item in question is one for

33Stigler, supra note 31, at 188.
which learning by doing is important and if the market for human capital is imperfect.³⁴ Delaying own-production until own-requirements are sufficient to exhaust scale economies would, considering the learning costs of undertaking own-production at this later time, incur substantial transition costs. It may, under these conditions, be more attractive from the outset for each firm to produce its own requirements—or, alternatively, for the affected firms to merge.³⁵ Absent present or prospective transaction costs of the sorts described, however, specialization by one of the firms (monopoly supply), to the mutual benefit of all, would presumably occur. Put differently, technology is no bar to contracting; it is transactional considerations that are decisive.

Aspects of the above argument can be illustrated with the help of Figure 2. The average costs of supplying the item in question by a specialized outside supplier at time 1 are shown by the curve \( AC^1 \). Firms that are already in the industry can supply the same item at the average costs shown by \( AC^2 \). The

³⁴ For a discussion of learning by doing, see P. Doeringer & M. Piore, Internal Labor Markets and Manpower Analysis (1971).

³⁵ Mergers would permit the firms involved to realize economies of scale with respect to the decreasing cost activity in question. Such mergers might also, however, result in market power. That such mergers are attractive in a private benefit sense is clear, but social net benefits need not obtain. See Williamson, supra note 9.
curve $AC_5$ is everywhere above the curve $AC^x$ because firms already in the industry avoid the setup costs which a specialized outside supplier would incur. Each of the firms in the industry generates requirements for the item at time 1 of $Q_i$. The total industry requirement at time 1 is $Q_T$.

The implicit comparison that Stigler makes in his explanation for vertical integration is point A versus point B. Thus although having a specialized supplier service the whole industry (produce $Q_T^t$) would permit economies of scale to be more fully exploited, the declining cost advantage is more than offset by the setup costs. Therefore, the average costs of the specialized supplier (at B) exceed the average costs that each individual firm would incur by supplying its own requirements (at A). My argument, however, is that point A should also be compared with point C—where point C shows the average costs of supplying the requirements for the entire industry by one of the firms that is already in the industry. Such a firm does not incur those setup costs which disadvantage the outside specialist supplier. Given the decreasing cost technology that Stigler assumes, the average costs at C are necessarily less than those at A. Why then not have one of the firms already in the industry supply both itself and all others? The impediments, I submit, are the hazards of interfirmer contracting (of both long term and spot market types) that have been described above.

The comparison, moreover, can be extended to include a consideration of the curve $AC_2$, which represents the average costs that will be incurred by a firm at time 2 that has been supplying continuously during the interval from time 1 to time 2. The curve $AC_2$ is everywhere lower than $AC_1^x$ by reason of advantages gained from learning by doing. To the extent that such learning advantages are not or cannot be shared with others, they will accrue only to firms that have undertaken own-production during the period in question. Thus if one of the firms in the industry becomes the monopoly supplier to all others at time 1, and if at time 2 the other firms become dissatisfied with the monopoly supplier's terms, the buying firms cannot undertake to supply their own requirements at a later date on cost parity terms because they have not had the benefit of learning by doing.

---

36 Again, this is because the market for human capital is imperfect. Firm X cannot simply hire firm Y's experienced employees away without incurring very considerable transfer costs. The learning by doing knowledge is thus impacted in firm Y.
Note finally the arrow that points away from point A toward point D. If the industry is expected to grow (plainly the case for the circumstances described by Stigler) and if each of the firms in the industry can be expected to grow with it, then each firm, if it supplies its own requirements \((Q_t)\) at time \(t\) and incurs average costs of \(A\), can, by reason of both growth and learning by doing, anticipate declining own-supply costs —perhaps to the extent that each substantially exhausts the economies of scale that are available. Since supplying its own requirements avoids the transactional hazards of procuring its supply from a market with only a few trading partners, vertical integration of the items with a decreasing cost technology is all the more to be expected.

B. Technological Interdependency

Of the various rationales for vertical integration that have been advanced, the technological interdependency argument is both the most familiar and straightforward: successive processes which naturally follow immediately in time and place dictate certain efficient manufacturing configurations; these, in turn, are held to have common ownership implications. Such technical complementarity is probably more important in flow process operations, such as chemicals and metals, than in separable component manufacture. The standard example is the integration of the making of iron and steel, where thermal economies are said to be available through integration. It is commonly held that where "integration does not have this physical or technical aspect—as it does not, for example, in integrating the production of assorted components with the assembly of those components—the case for cost savings from integration is generally much less clear."\(^{37}\)

I submit, however, that such technological interdependency is neither essential for cost savings to be realized by integration nor typical of most integrated activities. Consider Adam Smith's pinmaking example.\(^{38}\) Pin manufacture involved a series of technologically distinct operations such as wire straightening, cutting, pointing, and grinding. In principle, each of these activities could be performed by an independent specialist and work passed from station to station by contract. The introduction of buffer inventories at each station, moreover, would


decrease the coordination requirements and thereby reduce contractual complexity. Each worker could then proceed at his own pace, subject only to the condition that he maintain his buffer inventory at some minimum level. A series of independent entrepreneurs rather than a group of employees, each subject to an authority relation, would thus perform the tasks in question.

Transaction costs militate against such an organization of tasks, however. For one thing, it may be possible to economize on buffer inventories by having the entire group act as a unit, under common direction, with respect to such matters as work breaks and variable rates of production. Although rules could be worked out in advance and made explicit in the contract, or the authority to make such decisions could be rotated among the members of the group, coordination might usefully be assigned to a "boss", who oversees the entire operation and can more easily judge the fatigue and related work attitudes in the group.

The more pressing reasons for replacing autonomous contracting by an employment relation, however, turn on adaptability considerations. Suppose one of the individuals becomes ill (real or feigned) or becomes injured. Who nominates and chooses a replacement, or otherwise arranges to pick up the slack, and how is compensation determined? Reaching agreement on such matters is apt to be relatively costly compared to having a boss reassign the work among the members of the group or make other ad hoc arrangements on the group's behalf. Similarly, what is to be done if an individual declines to deliver the requisite quantity or quality to the next station? How are penalties determined? Litigation is apt to be costly and time consuming, and to what avail if the individual lacks the requisite assets to compensate for the losses attributable to his deviant behavior? Again, remedies and adaptations under an employment relation, where an individual has much weaker property claims to a work station, are likely to be quicker and less costly to effectuate.

The problem, more generally, is that autonomous contracting in small numbers circumstances is fraught with difficulties if unforeseen events requiring adaptation frequently appear, especially if the parties are given to opportunism. Rather than endure the costs that can be expected to arise when a series of bilateral contracts are negotiated among a group of individuals each of whom enjoys, in the short run at
least, a monopoly position, a firm will integrate such related activities instead. Central ownership of the work stations and an employment relation between the workers and entrepreneur will facilitate adaptation.\textsuperscript{39}

C. The Condition of Entry

Stigler observes that "it is possible that vertical integration increases the difficulty of entry by new firms, by increasing the capital and knowledge necessary to conduct several types of operation rather than depend on rivals for supplies or markets."\textsuperscript{40} Others, however, take exception to this argument. Robert Bork, for example, contends that "In general, if greater than competitive profits are to be made in an industry, entry should occur whether the entrant has to come in at both levels or not. I know of no theory of imperfections in the capital market which would lead suppliers of capital to avoid areas of higher return to seek areas of lower return."\textsuperscript{41} Similarly, Ward Bowman observes that "difficulties of access to the capital market that enable X to offer a one dollar inducement (it has a bankroll) and prevent its rivals from responding (they have no bankroll and, though the offering of the inducement is a responsible business tactic, for some reason cannot borrow the money) . . . [have] yet to be demonstrated."\textsuperscript{42} As I hope to make apparent, these and related arguments of the received microtheory variety go through only if transaction cost considerations are suppressed. The pairing of bounded rationality with uncertainty and the joining of opportunism with a condition of "information impactedness" (where one party to a transaction has access to information that the other party can obtain only at some expense, if at all) are the neglected factors.\textsuperscript{43}

The phenomenon to be explained is not merely an increase of the financial requirements, as Stigler indicates, but an ad-

\textsuperscript{39}The specialization of riskbearing and strategic decisionmaking may also favor common ownership and the replacement of autonomous contracting by an employment relation.

\textsuperscript{40}Stigler, supra note 31, at 191.

\textsuperscript{41}Bork, Vertical Integration and Competitive Processes, in Public Policy Toward Mergers 139, 148 (1969).


\textsuperscript{43}Examples of information impactedness are given in this section and the sections that follow. For a specific illustration, see especially the text on implementation under uncertainty in Section III subsection infra.
verse alteration of the terms under which capital becomes available. Borrowing by the firm to finance additional plant and equipment is, of course, unlike borrowing by the consumer to purchase a house. The firm borrows funds in anticipation of realizing a prospective stream of earnings. These prospective earnings, as well as the resale value of the assets in question, are used to support the loan in question. The homeowner, by contrast, is not ordinarily able to augment his earnings by purchasing a house. Thus, whereas the householder who successively increases the size of his mortgage eventually incurs adverse capital costs, because the risks of default are greater, the firm need not likewise be impeded. Why then, if at all, does vertical integration by established firms disadvantage prospective entrants on account of capital market “defects”?

An assessment of the issues will be facilitated by setting out the specific alternatives. Suppose that two distinct stages of production can be identified in the industry in question (designated I and II respectively). Assume further that stage I in the industry is essentially monopolized while stage II may or may not be integrated. The question now is whether a potential entrant who has developed a technologically satisfactory stage I substitute and has an established reputation in activity related to stage I will be unaffected by the integrated condition of stage II. Consider, in particular, the following contrasting conditions: (1) the monopolistic stage I producer is not integrated, in which case the prospective new entrant can come into stage I only and sell his product to stage II producers (suitably expanded, if that is necessary for absorption of the additional stage I production), and (2) the monopolistic stage I producer is integrated into stage II so that either (a) the new entrant himself must come in at both stages or (b) independent new entrants appear simultaneously at both stages. If Bork and Bowman are correct, the cost of capital ought to be independent of these conditions.

To contend that the terms of finance are the same under condition 2(a) as they are under condition 1 implies that the capital market has equal confidence in the new entrant’s qualifications to perform stage II activities as it does in firms that are already experienced in the business. Except in circumstances where experienced firms are plainly inept, this is tantamount to saying that experience counts for nought. This, however, is implausible for transactions that involve large, discrete investments rather than small but recurring commit-
ments of funds. Thus, although a series of small, recurring transactions can be monitored reasonably effectively on the basis of \textit{ex post} experience, this is much less easy for transactions of the large, discrete variety—which are the kind under consideration here. Reputation, which is to say prior experience, is of special importance in establishing the terms of finance for transactions that involve large, discrete commitments of funds.

The significance that lenders and investors attach to reputation can be traced in part to the incompleteness of information regarding the qualifications of applicants for financing. Faced with incomplete information, suppliers of capital are vulnerable to opportunist representations. Unable to distinguish between those unknown candidates who have the capacity and the will to execute the project successfully from opportunists who assert that they are similarly qualified, when objectively (omnisciently) they are not, the terms of finance are adjusted adversely against the entire group. Furthermore, and of special relevance to the issue at hand, if lenders are not omniscient then, as between two candidates for financing, both of whom would be judged by an omniscient assessor to have identical capacities and wills to execute the project, but only one of whom has a favorable and widely known performance record, the unknown candidate will find that he is disadvantaged.\footnote{\textquoteleft\textquoteleft As H. B. Malmgren has noted, in a related context, some firms will see opportunities, but be unable to communicate their own information and expectations favorably to bankers, and thus be unable to acquire finance, or need to pay a higher charge for the capital borrowed. Bankers and investors of funds in turn will be attracted to those firms which have shown in the past an ability to perceive and exploit effectively new opportunities, as against new firms which can only give their word that what they think is good is in fact good. Malmgren, \textit{Information, Expectations and the Theory of the Firm}, 75 \textit{Q. J. Econ.} 417 (1961).}

Moreover, where both candidates are equally suspect, but one has access to internal sources of financing while the other does not, the candidate requiring outside financing may be unable to proceed. In this connection, timing can be of critical significance. If one firm moves to the integrated structure gradually and finances the undertaking out of internal funds, while the second firm perceives the market opportunity later but, to be viable, must move immediately to a comparably integrated structure, the second firm may have to contend with adverse capital market rates.
The learning by doing conditions referred to earlier are also germane to an assessment of the earnings opportunities of an integrated versus nonintegrated new entrant. By assumption, the prospective entrant is well qualified in stage I processes. If learning by doing yields significant cost advantages and if the prospective entrant has no special qualifications in stage II processes, will his incentive to enter be any the less keen if, by reason of the integration of his competitors, he must now come in at both stages? I submit that if the knowledge gleaned from experience is deeply impacted, which is to say that it is not generally known or easily made knowable to those who lack experience, and if it is very costly to hire away the requisite experienced personnel from the integrated firm, the prospective entrant is plainly at a disadvantage. Information impactedness and imperfect labor markets thus combine to explain the cost disadvantage of the otherwise qualified new entrant in relation to the experienced firm. Were the monopolistic stage I producer not to have integrated into stage II, so that the prospective entrant could come in at stage I only and could rely on already experienced stage II firms to acquire the necessary capital to expand appropriately and service his stage II needs, capital costs would be lower and the prospect of entry thereby enhanced.

The problems, moreover, do not vanish if the new entrant comes in at stage I only and relies on independent entry into stage II to occur (condition 2(b)). Not only is the cost of capital adjusted adversely against potential new processors in stage II, by reason of the lack of experience referred to above, but simultaneous yet independent entry into both stages may be impeded because of "nonconvergent expectations" such that interdependent decisions between stages will fail to be made in a compatible way. Lack of common information among independent stage I and stage II specialists with respect to the market opportunities which they confront and doubts regarding the true investment intentions and contractual reliability of other parties are the apparent impediments to effec-

---

45 See text accompanying notes 34 & 36 supra.
46 If the knowledge advantage of the experienced firm is dispersed among a team of individuals, negotiations to hire away the team are likely to be prohibitively expensive.
47 This assumes that the cost of capital varies directly with the perceived risk of the incremental investment, ceteris paribus.
48 Malmgren, supra note 44, at 401, 405.
tive coordination. Ultimately, however, the problems are to be attributed to the human and transactional factors described in section I.

To be sure, the argument has no special significance to analysis of monopoly power unless the industry in question is already very concentrated\(^4\) or, in less concentrated markets, conditions of effective collusion, which include collective refusal to deal, obtain. In such circumstances, however, interfirm rivalry, by itself, cannot be expected to self-policing the market in a way that reliably assures the competitive outcome. Accordingly, potential competition has an important market policing role to play. If potential entrants regard limitation of prevailing vertical structures as contributing importantly to the prospect of successful entry (as they may in highly concentrated industries), vertical restrictions that require funds to be raised by less, rather than more, experienced firms can impede entry.

The financing issue, then, is not that capital markets perversely avoid earnings opportunities, the test proposed by Bork, or that financing cannot be arranged under any terms whatsoever, the condition referred to by Bowman. Rather, the cost of capital is at issue. If a prospective new entrant has the self-financing to come in at one stage (or can raise the capital at reasonable terms, perhaps because of a proven capability at this stage of operations) but lacks the self-financing (and incurs adverse terms should he attempt to raise the capital) to come in at the second stage, the condition of entry can clearly be affected by pre-existing vertical restrictions.\(^5\)

D. Circumventing Regulation

As Ronald Coase\(^5\) and George Stigler\(^5\) have both pointed out, vertical integration is sometimes employed as a device by which to evade sales taxes, quota schemes, and other methods of nonprice rationing. Since such efforts by the government to interfere with the price mechanism typically apply to market-mediated but not to internal transactions, the shift of such transactions from the market to the firm serves to circumvent

---

\(^4\) Provisionally, I define a very concentrated industry to be one where the four-firm concentration ratio exceeds eighty percent.

\(^5\) Economies of scale at stage II can also serve as an impediment to entry if the monopolist has integrated into stage II. See Williamson, supra note 42, at 656.


these regulatory schemes.\textsuperscript{53} This is perfectly straightforward and is derived from received microtheory without appeal to transaction cost considerations.

Conventional microtheory can also be made to address the following issue: can a regulated firm that is permitted only a "fair" rate of return in supplying a final good or service effectively evade the regulatory restraint by integrating backward into supply of its own equipment? As David Dayan\textsuperscript{54} has shown, such backward integration will permit the regulated industry to earn monopoly profits if either equipment transfer prices or the rate of return at the equipment supply stage is unregulated.

While I do not wish to minimize the importance of such considerations in individual industries, I submit that these are rather special cases and that the main incentive for vertical integration is that integration serves to economize on transaction costs and/or is undertaken for the strategic purpose of impeding entry. The types of issues raised in the discussion of the transactional approach in section I and in the earlier parts of this section are the root causes for integration.

E. Policy Implications

Vertical integration raises serious antitrust issues only in those circumstances where otherwise qualified actual or potential rivals can be said to be disadvantaged by it. The two situations in which disadvantage to rivals may arise are dominant firm (or otherwise very concentrated) industries and moderately concentrated industries where collusion has been successfully effected. For the reasons given in section III, such collusion is usually difficult to achieve. Accordingly, very concentrated industries in which the bulk of production is accounted for by integrated firms constitute the subset of principal interest for antitrust policy.

Even in concentrated industries, vertical integration can not be held to be objectionable per se. Two cases can be distinguished. The easiest to deal with is the case where, but for vertical integration by the leading firms in stage I of the


industry, stage II would be competitively organized. The objection to vertical integration here is twofold. For one thing, the residual (nonintegrated) sector of the stage II market is so reduced that only a few firms of efficient size can service the stage II market. Firms that would otherwise be prepared to enter stage I may therefore be discouraged from coming in by the prospect of having to engage in small numbers bargaining, with all the hazards that this entails, with these few nonintegrated stage II firms. Moreover, integrated entry may be rendered unattractive if prospective stage I entrants lack experience in stage II-related activity, and therefore would incur high capital costs were they to enter both stages themselves. The integration of stages I and II by leading firms is anticompetitive then, in entry aspects at least, if severing the vertical connection would permit competitive (large numbers) stage II activity to develop without loss of scale economies.

The second case is that where economies of scale at both stages are large in relation to the size of the total (not merely the residual) market. The advantage of severing the vertical connection in these circumstances is that potential entrants into one of the stages will be less deterred from entering that stage, because they will not also have to incur the adverse capital costs attached to entry at the unfamiliar stage. Whether a welfare gain will thereby result depends, however, on offsetting factors of two kinds. First, with only a small number of firms at each stage, frequent haggling over contractual terms, imposing preparatory and negotiating costs on each side, may be expected between stage I and stage II firms. But even if these transaction costs do not obtain, the prospect is that goods and services will not be transferred between the stages at marginal cost prices. Assuming that the technologies in question are of the variable proportions types, inefficient factor utilization results. The question then is whether costs of these two types are more than offset by the gains of facilitating entry accomplished by severing a vertical connection.

Vertical integration in industries with low to moderate degrees of concentration does not, however, pose these same problems. Here a firm entering either stage can expect to strike competitive bargains with firms in the other stage, whether

55 If, however, the condition of entry into the supply stage is easy, small numbers by itself will not occasion monopolistic prices. See Stigler, supra note 31, at 188.
56 McKenzie, supra note 27.
they are integrated or nonintegrated. The reasons are that no single integrated firm enjoys a strategic advantage with respect to such transactions and that collusion by the collection of integrated firms (in supply or demand respects) is difficult to effectuate. Vertical integration rarely poses an antitrust issue, therefore, except when the industry in question is highly concentrated or, in less concentrated industries, collective refusals to deal are observed. But for such circumstances, vertical integration is apt to be of the efficiency promoting kind.

III. Oligopoly

The treatment of oligopoly in this section is less an analysis of oligopoly as such than an explication of why oligopoly can be expected to differ in nontrivial ways from monopoly. Although this difference may seem obvious, it has not always been so; the view that dissolution into oligopoly is no remedy for monopoly is widely held. Don Patinkin contends that unless there are “enough independent firms resulting from the dissolution to make the operation of competition possible . . . we will replace monopoly with some oligopolistic situation, and it is quite possible that we would be as badly off as under monopoly.”

I take exception to that position here. It fails to make allowance for the advantages of internal organization as compared with contracting in adaptational respects, and it gives

57 That a firm can expect to strike competitive bargains does not, of course, guarantee that it will earn “normal” profits. This depends on supply and demand conditions. In a growing industry, however, the nonintegrated but otherwise qualified entrant should be able to secure a niche for itself without difficulty, although its profit rate may vary over a business cycle more than do the rates of integrated firms.

58 This assumes that stage II entry is not easy.

59 Vertical integration within a stage, I take it, poses no problems for anyone. The rationale here is that supplied above in the context of the pinmaking example. Whether economies of vertical integration are realized between stages in what appears to be an unconcentrated industry is apt to turn on product differentiation considerations. Some of the components required by firms producing differentiated products may well be firm-specific, in which event a genuine large numbers supply condition may not be feasible. Where, however, competitive supply terms (both presently and prospectively) can be anticipated, own-supply has little to commend it. (On this, see Williamson, supra note 15).

60 E.g., J.K. Galbraith, American Capitalism 58 (1952). The view that tight oligopoly and monopoly are equivalent is especially prevalent among non-industrial organization specialists.

61 Patinkin, Multiple-Plant Firms, Cartels, and Imperfect Competition, 61 Q.J. Econ. 184 (1947).
insufficient standing to the different incentives, and the related propensity to cheat, that distinguish internal from interfirm organization.

A. Economic Antecedents of the Transactional Approach

1. Fellner on Qualified Joint Profit Maximization

William Fellner contends that it is impossible to deduce determinate prices and outputs for oligopoly markets on the basis of demand and supply functions that are derived from technological data and utility functions. Rather, fewness carries with it a range of indeterminacy. Thus, although received price theory is useful for establishing the region of indeterminacy, notions of "conjectural interdependence" are needed to ascertain how choice is made within these limits. As he sees it, "all problems of conjectural interdependence are essentially problems of bargaining—provided we interpret bargaining in the broader sense, including the 'implicit' variety."

Within the range of indeterminacy, Fellner identifies four factors which determine relative bargaining power. The first two are concerned with social and political limits on bargaining and need not detain us here. The second two are more situation specific: the ability of the parties to take and to inflict losses during stalemates; and toughness, in the sense of unwillingness to yield.

He notes that quasi-agreements (bargains) will change in response to shifts in relative strength among the parties, and that changing market circumstances make it necessary for oligopolistic rivals to adapt their behavior appropriately. Such quasi-agreements, moreover, "do not usually handle all economic variables entering into the determination of aggregate gains." Although this is partly because of "administrative circumstances," where these are left undefined, "it is

---

62 Two important treatments of the oligopoly problem to which I would call attention, but do not discuss here, are L. Telser, supra note 7 and Shubik, Information, Duopoly, and Competitive Markets: A Sensitivity Analysis, 26 Kyklos 736 (1973). Both are somewhat in the spirit of the transactional approach that I propose, and both develop a useful modeling apparatus to help evaluate the issues.

63 Fellner, Competition Among the Few 9-11 (1949).

64 Id. 16.

65 Id. 27-28.

66 Id. 34.

67 Id.
TRANSACTION COST IN ANTITRUST POLICY

largely a consequence also of uncertainty due to which various persons and organizations discount their own future possibilities. . . . This is especially true of those variables that require skill and ingenuity in handling (such as those directly connected with advertising, product variation, technological change, and so forth)."68 Later he indicates that the use of strategic variables of these kinds requires inventiveness,69 and indicates that "the present value of this future flow of inventiveness cannot be calculated with sufficient accuracy" for the relative strength of the parties to be established.70 This in turn prevents the corresponding quasi-agreement from being reached. As an industry "matures," however, and particularly if new entrants do not appear, the degree of competition with respect to nonprice variables may be attenuated.71

Fellner indicates that profit pooling would not be necessary to reach a full-blown joint profit maximization result in those oligopolies where (1) the product is undifferentiated and (2) all firms have identical horizontal cost curves.72 In these circumstances a simple market sharing agreement will suffice to achieve this result. Such conditions, however, represent a very special case. Even here, moreover, there is the need to reach agreement on what adjustments to make to changing demand conditions: who decides? how are differences reconciled?

In the more usual case where cost differences and/or product differentiation exist, joint profit maximization requires interfirm cash flows. Complete pooling in these circumstances implies that "no attention is paid to how much profit each participant earns directly on the market but only to how much the aggregate of the participants earns. Each participant is compensated from the pool of earnings according to his share."73 Profit pooling, however, is held to be hazardous both for antitrust reasons and, even more, because some firms will be at a "substantial disadvantage if the agreement is terminated and aggressive competition is resumed."74 Consequently, only qualified joint profit maximization among oligopolists is to be expected.

68 Id. 34-35.
69 Id. 183-84.
70 Id. 185.
71 Id. 188-89.
72 Id. 129.
73 Id. 135.
74 Id. 133, 196.
2. Stigler on Oligopoly

Stigler takes as given that oligopolists wish, through collusion, to maximize joint profits and attempts to establish the factors which affect the efficacy of such aspirations. While he admits that "colluding firms must agree upon the price structure appropriate to the transaction classes which they are prepared to recognize," his analysis is focused entirely on the problem of policing such a collusive agreement. "A price structure of some complexity," one which makes "appropriate" provision for heterogeneity among products and buyers and for the hazard of activating potential entrants, is simply imputed to oligopolists.

Stigler notes that since secret violations of such agreements commonly permit individual members of an oligopoly to gain larger profits than they would gain by strict adherence to the agreement, a mechanism to enforce agreements is needed. Enforcement for Stigler "consists basically of detecting significant deviations from the agreed-upon prices. Once detected, the deviations will tend to disappear because they are no longer secret and will be matched by fellow conspirators if they are not withdrawn." Accordingly, a weak conspiracy is one in which "price cutting is detected only slowly and incompletely."

Since an audit of transaction prices reported by sellers is unlawful, and in any event may be unreliable, transaction prices paid by buyers are needed to detect secret price cutting. Stigler contends, in this connection, that statistical inference techniques are the usual way in which such price cutting is discovered. In particular, the basic method of detecting a price cutter is that he is getting business that he would not otherwise obtain. Among the implications of this statistical inference approach to oligopoly are that (1) collusion is more effective in markets where buyers correctly report prices tendered (as in government bidding), (2) collusion is limited if the identity

---

76 Id. 45.
77 Id. 45.
78 Stigler simply assumes "that the collusion has been effected, and a price structure agreed upon." Id. 46.
79 Profits here are expressed as expected, discounted values.
80 Stigler, supra note 75, at 46.
81 Id.
82 Id. 47.
83 Id. 48.
84 Id.
of buyers is continuously changing (as in the construction indus-
tries), and (3) elsewhere the efficacy of collusion varies
inversely with the number of sellers, the number of buyers,
and the proportion of new buyers, but directly with the degree
of inequality of firm size among sellers.

B. Legal Antecedents of Transactional Analysis

1. Turner on Conscious Parallelism

Donald Turner's basic position on conscious parallelism is
that such behavior, by itself, does not imply agreement. It
needs to be buttressed by additional evidence that the observed
parallelism is not simply "identical but unrelated responses
by a group of similarly situated competitors to the same set
economic facts." He illustrates the argument by posing
an "extreme hypothetical" in which there are only two or three
suppliers—each of identical size, producing an identical prod-
and markets are static. He contends,
in these circumstances, that "the 'best' price for each seller
would be precisely the same, would be known to be the same
by all, and would be charged without hesitation in absolute
certainty that the others would price likewise." Although
he is not explicit on this, the price that he appears to have in
mind is the joint profit maximizing (monopoly) price.

Turner then goes on to note that the hypothetical is rather
unrealistic. Products are rarely fully homogeneous, cost dif-
ences will ordinarily exist, and adaptations will need to be
made to changing market circumstances. He accordingly
holds that "for a pattern of noncompetitive pricing to emerge

\[ q = \left[ \frac{1}{n+1} \right] Q \],
where $n$ is the number of firms in the industry
and $Q$ is the competitive output. The joint profit maximizing output ($q^*$) for each
such firm, by contrast, is

\[ q^* = \frac{1}{2n} Q \].
Plainly, $q^* < q$ for $n > 1$. (For $n = 1$, both
formulae yield the monopoly output.)

Turner, supra note 87, at 664.
... requires something which we could, not unreasonably, call a 'meeting of the minds'.\textsuperscript{92} He declines, however, to regard this as unlawful. Absent explicit collusion, this is merely rational price making in the light of all market facts: \textsuperscript{93} "If monopoly and monopoly pricing are not unlawful per se, neither should oligopoly and oligopoly pricing, absent agreement of the usual sort, be unlawful per se."\textsuperscript{94}

Because the behavior in question cannot be rectified by injunction ("What specifically is to be enjoined?"),\textsuperscript{95} relief would presumably have to take the form of dissolution or divestiture.\textsuperscript{96} This, however, is to admit that the fundamental issue is structure, not remediable conduct. Unless structural monopoly is to be subject to dissolution, structural oligopoly ought presumably to be permitted to stand. Although Turner declined in 1962 to propose a structural remedy for either condition, he has since altered his position on both.\textsuperscript{97}

2. Posner on Oligopoly

Richard Posner takes exception to Turner's position that oligopolistic interdependence of a natural and noncollusive sort explains the price excesses in oligopolistic industries.\textsuperscript{98} Rather, a small numbers condition is held to be merely a necessary but not a sufficient condition for such price excesses to appear.\textsuperscript{99} Because "interdependence theory does not explain... how oligopolistic sellers establish a supracompetitive price,"\textsuperscript{100} including adjustment to changing market conditions, Posner suggests that the study of oligopoly proceed in terms of cartel theory instead.\textsuperscript{101}

\textsuperscript{92} Id. Note again, as pointed out in note 90 \textit{supra}, that independently operating Cournot duopolists do not charge competitive prices yet are not colluding in any usual sense either. Turner seems implicitly to hold that independent pricing will yield the competitive solution. Hence, any price that exceeds the competitive price is regarded as an indication of interdependence realized. Posner appears also to be of this view. See text accompanying notes 98-107 \textit{infra}.
\textsuperscript{93} Turner, \textit{supra} note 87, at 666.
\textsuperscript{94} Id. 667-68.
\textsuperscript{95} Id. 669.
\textsuperscript{96} Id. 671.
\textsuperscript{99} Id. 1571.
\textsuperscript{100} Id. 1568, 1578.
\textsuperscript{101} Id. 1568-69.
Posner's basic argument is that "voluntary actions by the sellers are necessary to translate the rare condition of an oligopoly market into a situation of noncompetitive pricing." Effective cartel behavior is, moreover, costly to effectuate; costs of bargaining, adaptation, and enforcement must all be incurred. The upshot is that because "tacit collusion or non-competitive pricing is not inherent in an oligopolistic market structure but, like conventional cartelizing, requires additional, voluntary behavior by sellers," a conduct remedy under section one of the Sherman Act is held to be appropriate. Once the oligopolist is faced with the prospect of severe penalties for collusion, tacit or otherwise, Posner concludes that the rational oligopolist will commonly decide not to collude but will expand his output until competitive returns are realized.

C. The Transaction Cost Approach

To focus attention on what I believe to be the critical issues, I will assume, initially, that oligopolistic agreements are lawful, in that there is no legal bar to collusion, but that oligopolists cannot appeal to the courts for assistance in enforcing the terms of an oligopolistic agreement. The oligopolists themselves, however, can take punitive actions to bring deviant members into line, provided that laws such as those prohibiting libel or the destruction of property are respected. Entry is assumed to be difficult; also, I will assume that profit pooling is permitted but that horizontal mergers between the firms are disallowed.

I will argue that oligopolists will commonly have difficulty in reaching, implementing, and enforcing agreements under these circumstances, but this argument does not mean that laws regarding oligopoly are of no account. The stipulations that horizontal mergers are disallowed and that collusive agreements are unenforceable in the courts are both important in this connection. If, however, it can be shown that monopolistic outcomes are difficult to effectuate even when the law permits

---

102 Id. 1575.
103 Id. 1570.
104 Id. 1578 (emphasis added).
106 Posner, supra note 98, at 1578-93.
107 Id. 1591. This conclusion appears, however, to be unwarranted because independently operating Cournot oligopolists do not produce competitive outputs. See note 90 supra.
108 Telser, supra note 7, does not make this last assumption. His analysis differs from mine partly for this reason.
collusion, then the performance differences between monop-
olies (dominant firm markets) and oligopolies are not to be
attributed principally to the unlawfulness of collusion among
oligopolists.\footnote{109} It follows, of course, that if express and law-
ful agreements are difficult for oligopolists to reach and im-
plement, tacit agreements are even less reliable instruments
for achieving collusion.

An agreement between two or more parties will be attrac-
tive in the degree to which (1) the good, service, or behavior
in question is amenable to specification in writing; (2) joint
gains from collective action are potentially available; (3) imple-
mentation in the face of uncertainty does not occasion costly
haggling; (4) monitoring the agreement is not costly; and (5)
detected noncompliance carries commensurate penalties at
low enforcement expense. Consider the application of the
transaction cost approach proposed in section I to each of
these conditions in an oligopolistic agreement.

1. Specification of Terms

Recall that oligopolistic collusion is assumed to be lawful.
The parties to the collusive arrangement can therefore nego-
tiate openly and express the details of the agreement in writ-
ing without exposing themselves to prosecution. The question
to be assessed here is whether the latitude thus afforded will
permit a comprehensive collusive agreement to be specified.

I submit that, except in rather special and unlikely cir-
cumstances, a comprehensive agreement to maximize joint
profits (but not entailing merger) will rarely be feasible be-
cause of transaction costs. A comprehensive statement of this
kind would require an inordinate amount of knowledge about
the cost and product characteristics of each firm, the inter-
action effects between the decision variables within each firm,
and the interaction effects of decision variables between firms.
Not only is the relevant information costly to come by, to say
nothing of digesting it and devising the appropriate adapta-
tion for each of the firms to make, but, if anything approxi-
mating a complete agreement is to be written, this information
gathering and analysis needs to be done \textit{ex ante} for a whole
series of contingent future events, most of which will never
materialize.

\footnote{109} I do not, however, mean to suggest that the antitrust statutes prohibiting
collusion are without purpose. They certainly compound the typical oligopolist's
problems.
The point is that joint profit maximization, even as an abstract exercise, is very difficult to accomplish once one departs from the simplest sort of textbook exercise. Homogeneous products, identical linear and horizontal cost curves, and static markets constitute the "ideal." Maintaining these product and cost assumptions in the face of changing demand does not greatly complicate the abstract analysis, in that the conditions of joint profit maximization are easy to display, but the operational problems become somewhat more difficult in the face of uncertainty, which will be discussed below.\textsuperscript{110}

In more realistic circumstances, involving differentiated products, product and process innovations, organization form\textsuperscript{111} changes, and revisions in selling expense and financial strategies, the resulting complexity becomes impossibly great in relation to the bounded rationality of planners. When, in addition, the optimization problem is cast in a multiperiod framework under conditions of uncertainty, abstract analysis breaks down.\textsuperscript{112} One concludes, accordingly, that the absence of legal prohibitions to collusive agreements is not what prevents comprehensive collusion.\textsuperscript{113} Rather, it is prevented by elementary considerations of bounded rationality.\textsuperscript{114}

2. Joint Gains

Suppose, arguendo, that it were possible to specify the joint profit maximizing strategy. Would the parties then be prepared to make such an agreement? I submit that, but for

\textsuperscript{110} See subsection 3 infra.

\textsuperscript{111} In the sense of WILLIAMSON, supra note 22, at 109-81.

\textsuperscript{112} For an operational treatment of the problem of joint profit maximization in a multiproduct firm where (1) product lines are independent, (2) only heuristic rather than full-blown optimization methods are attempted, and (3) only the financial decision is considered, see Hamilton & Moses, An Optimization Model for Corporate Financial Planning, 21 OPERATIONS RESEARCH 677 (1972). Their model contains approximately 1000 variables and 750 constraints, id. 686, and tests not one but various configurations of the strategic variables. Replicating such an arrangement by interfirm agreement boggles the mind. Complicating the analysis further to include interdependent products (which, of course, is the case in oligopoly) and the full range of decision variables discussed in the text reveals the manifest impossibility of attempting comprehensively to maximize joint profits—even by heuristic simulation methods, much less by determinate written agreements.

\textsuperscript{113} Again, however, the view expressed in note 109 supra applies.

\textsuperscript{114} It is possible, of course, that oligopolists could reach agreement on some aspects of the market more easily than on others. Faced with diminishing marginal returns to their efforts to obtain an agreement (transaction costs), they would probably settle on an agreement of less than comprehensive scope. It is also possible, however, that the inability to agree on some matters would frustrate any agreement whatsoever, even on matters which might be settled if they could be considered in isolation.
the simple textbook cases referred to above, the parties would commonly decline to accept comprehensive joint profit maximization of the profit pooling kind.

Partly disagreement might arise, as Fellner suggests, on account of differences between the parties concerning the appropriate discount rates to be used in evaluating future prospects. Surely more fundamental, however, are the risks and monitoring expenses that profit pooling entails. As Fellner notes, some of the parties must accede to reductions in relative output and to contractions in relative firm size if the joint profit maximizing result is to be realized. This, however, is hazardous. Firms which are authorized to expand relatively as a result of the agreement will be powerfully situated to demand a renegotiated settlement at a later date. Wary of such opportunism, firms for which retrenchment is indicated will decline from the outset to accept a full-blown profit pooling arrangement. Moreover, even setting such concerns aside, monitoring the profit pooling agreement will be costly because of the pairing of opportunism with information impactedness. This will be discussed below.\(^\text{115}\)

3. Implementation under Uncertainty\(^\text{116}\)

Implementing an agreement under conditions of uncertainty requires that the parties agree, when changes in the environment occur, on what new state of the world obtains. Problems can arise if, for any true description of the state of the world, (1) some parties would realize benefits if a false state were to be declared, and either (2a) information regarding the state of the world is dispersed among the parties and must be pooled or (2b), despite the possession of identical information by all the parties, definitive agreement must still be reached.

Consider information condition 2b. Even though all parties have identical information with respect to the true condition of the environment, they need not agree on what state of the world has actually been realized. Unless the parties have fully stipulated how observations are to be interpreted as state of the world descriptions, differences in opinion can be anticipated. If some parties stand to benefit from having one state

\(^{115}\) See subsection 4 infra.

\(^{116}\) Unlike the preceding and succeeding subsections, the argument here assumes that joint profit maximization is not attempted.
declared, but others would benefit if another state were declared, and if each side can make a plausible case for its position, opportunistic representations in support of each outcome can be expected. Costly haggling may then ensue.

To illustrate, suppose that demand on day t is known to be a function only of the mean temperature on day t-1: if the mean temperature on day t-1 exceeds T_0, demand on day t is of type D_1; otherwise it is of type D_2. Suppose also that all firms have free access to temperature readings on day t-1 at 4:00 a.m., 12:00 noon, and 8:00 p.m. If on date t the unit weighted average of the temperatures on the preceding day is well above or well below T_0, the declaration of demand types is made without difficulty. Suppose, however, that the unit weighted average of day t-1 temperatures just slightly exceeds T_0 while weights of 0.95, 1.10, and 0.95 would reduce it to below T_0. If some parties benefit if demand is declared to be of type D_2, even though it is actually D_1, they may then assert that “everyone knows” that the noontime temperature deserves to be assigned a greater weight in computing the daily mean. Protracted haggling could ensue. Moreover, in the usual circumstances where the state of the world is multidimensional, the occasion for such disputes naturally increases.\footnote{If the state of the world is described by a vector of n components, each of which can take on only one of two values, the number of possible states is 2^n. For n = 8, which hardly constitutes a complex description of the state of the world, the number of possible states is 256, which is impressively large.}

The problems are compounded if the 2a condition obtains. Here the necessary information to ascertain the true state of the world is dispersed and pooling of the data is required. An agreement upon how to interpret the data is to little purpose if the parties selectively disclose or distort the information to which they have preferred access. The pairing of opportunism with information impactedness thus poses serious implementation problems to the oligopolists.

4. Monitoring Execution of the Agreement

As Stigler points out, and as is widely recognized, oligopolists have an incentive to cheat on price fixing agreements if they believe that cheating will go, for a time at least, undetected. Given that information about individual sales is impacted in that the seller knows exactly what the terms were but, given uncertainty, his rivals do not and can establish the terms only at some cost, the individual seller can often cut...
prices below the agreed level to the disadvantage of the other parties to the conspiracy. The pairing of opportunism (here manifested as cheating) with information impactedness makes oligopolistic agreements difficult to police.

This argument, moreover, applies to oligopolistic collusion with respect to considerations other than price as well. If anything, agreeing to collude with respect to marketing expense, research and development efforts, and similar business practices is even more hazardous than price collusion for nonopportunistic parties who are prepared to abide by the agreements. Although it is easy to establish after the fact that a rival has made significant design changes or introduced a new product in violation of the agreement, such information may come too late. If recovery from a large shift in market share, attributable to, for example, an "illicit" innovation is inordinately expensive, the detection of such a violation is to little avail—unless, of course, all firms have maintained a defensive posture against such contingencies, in which case collusion in these nonprice respects can scarcely be said to be operative.

As mentioned above, profit pooling is also subject to problems of monitoring. Even if firms were prepared to enter into agreements in which all profits are pooled and each participant is assigned a share of the total, there is still the problem of determining what the contribution of each firm to the pool should be. Individual firms have an incentive to understate true profits in these circumstances.

Moreover, merely auditing the earnings of each firm, even to the extent that all sources of revenue and cost are fully disclosed, is not sufficient to avoid distortion. An assessment of individual expense items must also be made. The problems facing the auditor here are akin to those facing the defense agencies in monitoring cost-plus (or, more generally, cost-sharing) defense contracts. Unless it can be established that certain types or amounts of actual costs are unwarranted, and hence will be disallowed, each firm has an incentive to incur excessive costs.

Expense excesses can take any of several forms. Perhaps the simplest is to allow some operations to run slack so that the management and workers in the firm take part of their

---

118 See subsection 2 supra.
rewards as on-the-job leisure. A second way is to allow emoluments to escalate, in which case corporate personal consumption expenditures exceed levels which, from a profit maximizing standpoint, would be incurred. Third, and most important, firms may incur current costs which place them at a strategic advantage in future periods. Developing new and improved technology or training the work force are examples of this sort of cost. Evaluating individual firm performance in these several respects is at least an order of magnitude more difficult than simple audits of revenue and cost streams. Profit pooling, therefore, poses severe enforcement problems, even assuming that the agreement itself were legal.

5. Penalizing Violations

Recall that it was assumed that while collusive agreements are not unlawful, the participants in such agreements cannot call upon the courts to help enforce the agreement. Instead, violators must be determined and penalties must be administered by the parties to the contract. Problems of two types arise in connection with penalties. First, do the penalties, if implemented, constitute an effective deterrent to the would-be violator? Second, even if penalties can be devised that would be efficacious, will the parties to the conspiracy be prepared to impose them?

Because the conspirators lack legal standing, conventional penalties such as fines and jail sentences are presumably unavailable. Rather, penalties are *exacted in the market place* by confronting the violator with unusually adverse circumstances. Price reductions are matched and perhaps even undercut. Normal types of interfirm cooperation (e.g., supply of components) is suspended. Key employees may be raided. Except, however, as deviant firms are highly dependent on rivals for vital supplies, such market reactions may well be ones that the deviant is prepared to risk.

For one thing, the contract violator is not the only firm to be adversely affected by exacting these penalties in the market place. The firms meting out the penalties also incur costs.¹²⁰ Second, and related, securing the collective action needed to punish the violator may be difficult. Thus, although

¹²⁰ Punitive market responses require firms to incur short run profit sacrifices in the hope of discouraging future chiselers and returning the current chiseler to the fold.
all firms may agree both that a violation has taken place and that the violator deserves to be punished, not all may be prepared to participate in administering it. Defectors (e.g., those willing to supply the deviant with the essential component, perhaps at a premium price), which is to say opportunists, who refuse to incur the costs of punishing the violator, naturally reduce the costs of being detected in violation of the agreement. Where such defection is deemed likely, collusive agreements are all the less probable.\footnote{121}

D. Policy Implications: Dominant Firms versus Oligopolistic Interdependence

The monopolist (or dominant firm) enjoys an advantage over oligopolists in adaptational respects because he does not have to write a contract in which future contingencies are identified and appropriate adaptations thereto are devised. Instead, he can face the contingencies when they arise; each bridge can be crossed when it is reached, rather than having to decide \textit{ex ante} how to cross all bridges that one might conceivably face.\footnote{122} Put differently, the monopolist can employ an adaptive, sequential decision-making procedure, which greatly economizes on bounded rationality demands, without exposing himself to the risks of contractual incompleteness which confront a group of conspiring oligopolists. Adaptation within a firm (in contrast to that between firms) is also promoted by the more complete development of efficient, albeit often informal, communication codes and an associated trust relationship between the parties.\footnote{123} Thus, while I do not mean

\footnote{121}Although the opportunistic behavior described mainly reflects an aggressive effort to realize shortrun individual gains, to the disadvantage of the group, firms may also engage in such behavior for defensive reasons. Defensive opportunism reflects a lack of confidence in the trustworthiness of other members of the group and an unwillingness to risk being put to a strategic disadvantage.

While aggressive or assertive opportunism is to be expected whenever the viability of any particular firm is threatened, whatever the degree of “maturity” of the firms in an industry, defensive opportunism will vary inversely with maturity. Because defensive opportunism, if widely practiced, is mutually disadvantageous, and because this is self-evident to the parties, organizational learning is normally to be expected. Among other things, ways of announcing or signaling intentions in ways that will not be misinterpreted as aggressive, when no such intention exists, are apt to develop. Unless, therefore, the industry is one in which new entrants regularly appear, with obviously disruptive consequences for interfirm learning and accommodation, occasions for defensive opportunism are likely to decline as an industry matures.

\footnote{122}For a discussion of adaptive, sequential decisionmaking, see H. CHERNOFF & L. MOSES, \textit{Elementary Decision Theory} 166-94 (1959).

to suggest that there are no costs whatsoever to dissolution,\textsuperscript{124} and, accordingly, do not propose it as an automatic remedy, to suggest that oligopolists will be able easily to replicate the (joint) profit maximization strategy of a monopolist is simply unwarranted. Even if cheating on a specific agreement were not a problem, there is still the need among oligopolists to reach the specific agreement. The high cost of exhaustively complete specification of agreements discourages efforts toward comprehensiveness—in which case, because actual oligopolistic contracts are of the incomplete coordination kind, competition of a nonprice sort predictably obtains.

To assume, moreover, that oligopolists will voluntarily adhere to whatever limited agreements they reach is plainly unreasonable. Cheating is a predictable consequence of oligopolistic conspiracy; the record is replete with examples.\textsuperscript{125} The pairing of opportunism with information impactedness explains this condition.

The monopolist, by contrast, does not face the same need to attenuate opportunism. Even within the monopoly firm in which semi-autonomous operating divisions have been created, with each operated as a profit center, interdivisional cheating on agreements will be less than interfirm cheating because (1) the gains can be less fully appropriated by the defector division, (2) the difficulty of detecting cheating is much less, and (3) the penalties to which internal organization has access (including dismissing opportunist division managers) are more efficacious. Unlike independently owned oligopoly firms, the operating divisions do not have fully pre-emptive claims on their profit streams (so the inclination to cheat is less) and, unlike oligopolies, they are subject to detailed audits, including an assessment of internal efficiency performance. Also, where oligopolists can usually penalize defectors only by incurring losses themselves (e.g., by matching or over-matching price cuts), the monopoly firm has access to a powerful and delicately conceived internal incentive system that does not require it to incur market penalties of a price cutting sort.\textsuperscript{126} It can mete out penalties to groups and individuals in the firm in a quasijudicial fashion and in this way it assumes

\textsuperscript{124} See Williamson, supra note 97, at 1528-30.

\textsuperscript{125} For some discussions and examples of cheating and the breakdown of oligopolistic collusion, see Patinkin, supra note 61, at 200-04; Posner, supra note 98, at 1576; R. Smith, Corporations in Crisis 113-66 (1966).

\textsuperscript{126} See Williamson, supra note 22, at 54-73, 109-19.
some of the functions of a legal system. Altogether, the opportu-

nism which threatens agreements among oligopolists is a less

severe problem for the dominant firm.

More generally, my argument comes down to this: it is

naive to regard oligopolists as shared monopolists in any com-

prehensive sense—especially if they have differentiated products,

have different cost experiences, are differently situated with

respect to the market by virtue of size, and plainly lack the

machinery by which oligopolistic coordination, except of the

most primitive variety, is accomplished and enforced. Except,

therefore, in highly concentrated industries producing homogo-

neous products, with nontrivial barriers to entry, and at a

mature stage of development oligopoly is unlikely to pose anti-

trust issues for which dissolution is an appropriate remedy.

In the usual oligopoly situation efforts to achieve collusion

are unlikely to be successful or, if they are, will require suffi-

cient explicit communication that normal remedies against

price fixing, including injunctions not to collude, will suffice.

Where, however, the industry is of the special type just

described, recognized interdependency may be sufficiently

extensive to permit tacit collusion to succeed. Injunctive rem-

edies, as Turner noted, are unsatisfactory in such circum-

stances.\textsuperscript{127} Accordingly, dissolution ought to be actively con-

sidered. The recent case brought by the Antitrust Division

against the major firms in the gypsum industry affords a cur-

rent example of a case in which, assuming the charges can be

proved, dissolution would appear to be warranted.\textsuperscript{128} By con-

trast, the cereal case brought by the Federal Trade Commis-

sion is not one for which comprehensive collusion seems

likely.\textsuperscript{129}

This does not, however, imply that the cereal industry

poses no public policy problems whatsoever. Simply because

the shared monopoly model does not fit well does not mean

that public policy concerns vanish. But I would urge that at-

tention be focused on those specific practices in the industry

which are thought to be objectionable. If, for example, exces-

sive advertising in the cereal industry can be reasonably estab-

\textsuperscript{127} Turner, \textit{supra} note 95.

\textsuperscript{128} United States v. United States Gypsum Co., Crim. No. 73-347 (W.D. Pa.,


\textsuperscript{129} \textit{In re} Kellogg Co., No. 8883 (FTC, filed Jan. 24, 1972).
lished, this can be dealt with directly. Selective attention to specific wasteful practices, rather than grand conspiracy theories, are called for.

A related implication of the argument is that dissolution of dominant firms is not an idle economic exercise, done to reduce large aggregations of corporate power for political or social purposes alone but unlikely to have significant economic performance consequences. For all the reasons developed above, several independent entities cannot realize the same degree of coordination between their policies in price and nonprice respects as can a single firm. Moreover, the price and nonprice differences that predictably arise will typically redound to the consumer’s benefit. Accordingly, a more assertive antitrust policy with regard to the dissolution of dominant firms is indicated.

130 See note 112 supra.


132 Wasteful selling or product development expenditures among differentiated product oligopolists are sometimes, however, observed. Specific steps might properly be taken to restrict this were a dominant firm to be split into independent, differentiated parts.

133 See sources cited at note 97 supra.

The reader may find a summary of the transactional approach and its antecedents useful at this point. The approaches often coincide, but there are many contrasts. The problem of oligopoly under the transaction cost approach, as under Fellner’s approach, is treated as a problem of interdependence recognized. Also, as with Fellner, the multidimensional nature of the interdependence issue is emphasized; price coordination is only a part of the problem, especially in industries with differentiated products. But whereas Fellner attributes the problems of interdependence to the complexities of discounting uncertain future values and in pooling risks, I put the issue in terms of “contracting” about contingent claims. While these approaches are not unrelated, the latter highlights the issues of coming to an agreement and enforcing it which permits us to draw expressly on the transaction cost framework sketched out in section I. A more complete assessment of the problems of oligopolistic collusion is thereby afforded.

Stigler’s analysis runs almost entirely in terms of prices. Moreover, he takes the collusive agreement itself as given, focusing attention instead on cheating and on statistical inference techniques for detecting cheaters. While this last is very useful, and calls attention in an interesting way to aspects of the oligopoly problem that others have rather neglected, it is also highly incomplete. The discussion in subsection C reveals that monitoring is only one of a series of steps in the oligopolistic collusion, and not plainly the one that warrants greatest attention.

Both Turner and Posner also give primary attention to prices in their discussions of oligopoly. But their similarity to each other ends there. Whereas Turner emphasizes tacit collusion of the recognized interdependence sort and finds injunctive relief to be inefficacious, Posner regards interdependence theory as unsatisfactory, and discusses oligopoly instead as a cartel problem, concluding that injunctive relief is appropriate.

The spirit of my discussion is somewhat akin to Posner’s cartel analysis, but the specifics plainly differ. I restate the problem in terms of what a lawful cartel could
IV. CONGLOMERATE ORGANIZATION

A. Received Microtheory versus Transactional Interpretations

As the remarks of Bork and Bowman cited earlier make clear, received microtheory is loath to concede that capital markets may fail to operate frictionlessly. Partly for this reason, the fiction that managers operate firms in fully profit maximizing ways is maintained. It is argued that any attempt by opportunist managers to promote their own goals at the expense of corporate profitability would occasion intervention through the capital market. Effective control of the corporation would be transferred to those parties who perceived the lapse; profit maximizing behavior would then be quickly restored.

Parties responsible for the detection and correction of deviant behavior in the firm would, of course, participate in the greater profits which the reconstituted management would realize. This participation would not, however, be large. One reason is that incumbent managements, by assumption, have little opportunity for inefficiency or malfeasance because any tendency toward waywardness would be quickly detected and costlessly extinguished. Accordingly, the incremental profit gain occasioned by takeover would be small. Moreover, the market for corporate control is presumably one in which large numbers of qualified takeover agents are noncollusively organized. Competitive offers assure that the takeover gains mainly redound to the stockholders.

Shorey Peterson's sanguine views on corporate behavior are roughly of this kind. He characterizes the latitude to disregard the profit goal as "small" and goes on to observe that "[f]ar from being an ordinary election, a proxy battle is a catastrophic event whose mere possibility is a threat, and one not remote when affairs are in conspicuous disarray." Indeed, even "stockholder suits . . . may be provoked by evidence of serious self-dealing." On the principle that the efficacy of legal prohibitions is to be judged "not by guilt discovered but accomplish. Also, I am much more concerned than Posner with the details of and impediments to successful interfirm agreements. Finally, I agree with Turner that injunctive relief in highly concentrated, homogeneous product, entry impeded, mature industries is unlikely to be effective. Structural relief is indicated here instead.

134 Text accompanying notes 41-42 supra.
135 Peterson, Corporate Control and Capitalism, 79 Q.J. Econ. 1, 11 (1965).
136 Id. 21 (emphasis added).
137 Id. (emphasis added).
by guilt discouraged," he concludes that such suits, albeit rare, may have accomplished much in helping to police the corporate system.\textsuperscript{138}

While I do not mean to suggest that such deterrence has not been important, Peterson's observations appear to me to be consistent with the proposition that traditional capital markets are beset by serious problems of information impactedness and incur nontrivial displacement costs if the incumbent management is disposed to resist the takeover effort. Why else the reference to catastrophic events, conspicuous disarray, and serious self-dealing? Systems that are described in these terms are not ones for which a delicately conceived control system can be said to be operating. As recent military history makes clear, controls that involve a large and discrete shock to the system are appropriate only when an offense reaches egregious proportions. The scope for opportunism, accordingly, is wider than Peterson seems prepared to concede.

The reasons why traditional control of management performance by the capital market is relatively crude are that internal conditions in the firm are not widely known or easy to discover (information impactedness) and that those seeking to gain control of the firm (takeover agents) might well take opportunistic advantage of the shareholders' bounded rationality. Information impactedness means that outsiders cannot make confident judgments that the firm has departed from profit maximizing standards, except with difficulty. The firm is a complex organization and its performance is a joint consequence of exogenous economic events, rival behavior, and internal decisions. Causal inferences are correspondingly difficult to make and, hence, opportunism is costly to detect. Moreover, once detected, convincing interested stockholders that a displacement effort ought to be supported encounters problems. Inasmuch as time and the analytical capacity of stockholders are not free goods (which is to say that the limits imposed by bounded rationality must be respected) the would-be takeover agent cannot simply display all of his evidence and expect stockholders to evaluate it and reach the "appropriate" conclusion. Rather, any appeal to the stockholders must be made in terms of highly digested interpretations of the facts. Although this helps to overcome the stockholder's bounded rationality problem, it poses another: How is the

\textsuperscript{138} Id.
interested stockholder (or his agent) to distinguish between bona fide and opportunistic takeover agents?

The upshot of these remarks is that the transaction costs associated with traditional capital market processes for policing management, of the sort described by Peterson, are considerable. Correspondingly, the range of discretionary behavior open to incumbent managements is rather wider than Peterson and other supporters of the fiction of the frictionless capital market concede.\textsuperscript{139}

One of the more attractive attributes of the conglomerate form of organization (of the appropriate kind)\textsuperscript{140} is that it serves to overcome certain of these limitations of traditional capital markets. The argument, which I will develop below, essentially reduces to the proposition that conglomerate firms (of the appropriate kind) function as miniature capital markets with consequences for resource allocation which are, on balance, beneficial.

This poses, however, the following paradox: under conventional assumptions that more choices are always preferred to fewer, the banking system ought to have superior resource allocation properties to any miniature imitation thereof. Put differently, why should a miniature capital market ever be preferred to the real thing? As might be anticipated, transaction cost considerations supply the resolution. If decision-makers could be easily apprised of an ever wider range of alternatives and choose intelligently among them, there would be no occasion to supplant the traditional market. But it is elementary that, where complex events have to be evaluated, information processing capacities are quickly reached. As a result, expanding the range of choice may not only be without purpose but can have net detrimental effects. A trade-off between breadth of information, in which respect the banking

\textsuperscript{139} Smiley estimates that "per share transaction costs are approximately 14% of the market value of the shares after a successful [tender] offer" and suggests that such a cost level warrants "skepticism about the efficacy of the tender offer in constraining managers to act in the best interests of their shareholders." R. Smiley, The Economics of Tender Offers 124-25, July 1973 (unpublished Ph.D dissertation, Stanford University).

\textsuperscript{140} This assumes that the hierarchic structure and internal control processes of the conglomerate satisfy the requirements that I have stipulated elsewhere (Williamson, supra note 22, at 120-53). Although there are certainly other types of conglomerates, those which lack for an underlying efficiency rationale (as contrasted with a temporary financial rationale) will presumably be sorted out in the long run. Those which pose financial problems are best dealt with by the SEC. My discussion sidesteps these and focuses on antitrust issues.
system may be presumed to have the advantage, and depth of
information, which is the advantage of the specialized firm,\footnote{141} is involved. The conglomerate can be regarded as an inter-
mediate form that, ideally, optimizes with respect to the
breadth-depth trade-off.\footnote{142} Although the number of alter-
natives considered by a conglomerate's management is limited,
its knowledge (\textit{ex post} and \textit{ex ante}) with respect to each remains
relatively deep. Operating as it does as an internal control
agent, its auditing powers are more extensive and its control
instruments are more selective than an external control agent
can employ. Information impactedness is reduced as a result
and opportunism is attenuated in the process.

\textbf{B. Objections to the Conglomerate}

The failure on the part of received microtheory to regard
the internal organization of the firm as interesting is, I be-
lieve, responsible for what Posner has called "the puzzle of
the conglomerate corporation."\footnote{143} This puzzle has not,
however, deterred those who most rely on received microtheory
from venturing the opinion that the conglomerate is innocent
of anticompetitive purpose or potential and ought not to be
an object of antitrust prosecution.\footnote{144} But an affirmative ra-
tionale for the conglomerate, based on received microtheory,
has yet to appear.\footnote{145}

The populist critics of the conglomerate have not allowed
this lapse to go unnoticed. Robert Solo's views are perhaps

\footnote{141} Depth of information problems can, however, appear as the specialized firm
becomes very large. \textit{See} O. \textit{Williamson, supra} note 22, at 14-40.
\footnote{142} For a somewhat similar interpretation of the conglomerate, see Alchian &
Demsetz, \textit{Production, Information Costs, and Economic Organization}, 62 \textit{AM. ECON. REV.}
Dec. 1972, at 777-95. For a study of the use of the computer to extend the firm's
capacity to deal effectively with a wider set of investment alternatives, see Hamilton
& Moses, \textit{supra} note 112. For a cross-sectional study of conglomerates (which, how-
ever, does not make organization form distinctions), see Weston & Mansinghka, \textit{Tests
of the Efficiency Performance of Conglomerate Firms}, 26 \textit{J. FINANCE} 419 (1971).
\footnote{143} R. \textit{Posner, Economic Analysis of Law} 204 (1972).
\footnote{144} For a report that approaches this position, see U.S. \textit{President's Task Force
on Productivity and Competition}.
\footnote{145} Some contend that reciprocity has attractive efficiency properties, in that
it facilitates priceshading in otherwise rigid price circumstances. While I concede
that reciprocity can be used in this way, I do not find it an especially compelling
economic rationale for the conglomerate. Surely the entire conglomerate movement is
not to be explained in these terms. Also, I think it useful to appreciate that reciprocity
can have inefficiency consequences. Once begun, perhaps as a price shading tech-
nique, it may be continued because it suits the bureaucratic preferences of the sales
staff.
representative. He contends that "when faced with a truly dangerous phenomenon, such as the conglomerate mergers of the 1960's, produced by financial manipulators making grist for their security mills, the professional antitrust economists were silent. Like other realities of a modern enterprise, this phenomenon, which will probably subvert management effectiveness and organizational rationale for generations, is outside their conceptual framework."  

Several things should be said in this connection. First, in defense of antitrust economists, I would point out that financial manipulation is not their main concern. This is the principal business of the Securities and Exchange Commission rather than the Antitrust Division. Although Solo might object, with cause, that economists are excessively narrow, nevertheless, as matters are divided up currently, it is the security specialists who are presumably at fault. Second, and more important, Solo's sweeping charges leave the particular dangers of the conglomerate phenomenon completely unspecified. Third, I agree that an understanding of the conglomerate requires an extension of the conventional framework. Nevertheless I think it noteworthy that populist critics of the conglomerate and received microtheorists alike pay little heed to the resource allocation consequences, in the form of capital market substitution effects, of internal organization. Finally, conglomerates come in a variety of forms and have a variety of purposes. Accordingly, any attack on conglomerates should be selective rather than broadside.

Responses to organizational innovations vary. The initial response of rival firms and financial analysts is typically to disregard such changes. Partly this is because "reorganization" is a common reaction by firms that are experiencing adversity. Discerning whether the response is intended to eliminate accumulated bureaucratic deadwood or to buy time from the stockholders by giving the impression that corrective action has been taken, or whether (instead or in addition) it represents a really fundamental change in structure that warrants more widespread attention is initially quite difficult.  

Expressed in transaction cost terms, the problem is that oppor-

---

146 Solo, New Maths and Old Sterilities, SATURDAY REVIEW, Jan. 22, 1972, at 47-48.  
147 It is interesting in this connection to note that General Motors' executives went to considerable effort in the 1920's to apprise the business community at large of the character and importance of the multidivisional structure which they had devised, but to little avail.
tunistic structural changes cannot easily be distinguished from fundamental ones on account of information impactedness and bounded rationality. Given the incapacity (or high cost) of communicating about and abstractly assessing the importance of organizational changes, the tendency is to wait and see how organizational changes manifest themselves in performance consequences. Inasmuch as performance is a function of many factors other than organizational structure alone, sorting out the effect of organizational changes is difficult. Therefore, a long recognition lag between fundamental innovation and widespread imitation is common.148

Public policy analysts of populist persuasions are prone to regard organizational innovations as having anticompetitive purposes. Rarely are such innovations thought to have possible efficiency consequences, mainly because efficiency is thought to reside in technological rather than transactional factors. Harlan Blake's widely admired assessment of the conglomerate and its policy implications is in this technological tradition.149 Like Solo's, his treatment tends to be global rather than selective. References to "mergers whose anticompetitive potential is so widespread that it might appropriately be described as having an effect upon the economic system as a whole—in every line of commerce in every section of the country"150 is unguarded. An understanding of the conglomerate phenomenon will be better promoted by delimiting the attack.151

For one thing, organization form distinctions, of which Blake makes none, ought to be made. Size considerations aside, he treats all conglomerates as an undifferentiated group. But there are indications that even come courts may be more discriminating than this.152 More generally, the point is this: just as the structure of markets influences the performance of markets, so likewise ought allowance to be made for the pos-

---

150 Id.
151 For an attempt to delimit the attack, see the discussion in subsection C infra.
sibility that internal organization influences firm performance.\textsuperscript{153}

Although Blake recognizes that the conglomerate may have had invigorating effects on the market for corporate control,\textsuperscript{154} he does not regard its ability to reallocate assets internally from lower yield to higher yield uses as an affirmative factor. If anything, he seems to suggest that internal resource reallocations are undesirable as compared to reallocations in the capital market.\textsuperscript{155} In an economy, however, where returning funds to and reallocating funds by the capital market incurs nontrivial transaction costs and/or where managers of specialized firms have an opportunistic preference to retain earnings, the internal reallocation of resources to uses returning a higher yield is what most commends the conglomerate as compared with similarly constituted specialized firms.\textsuperscript{156} The conglomerate in these circumstances assumes miniature capital market responsibilities of an energizing kind. That Blake is unimpressed with such consequences is explained by his assessment (which he shares with conventional microtheory)\textsuperscript{157} that only economies having technological origins are deserving of consideration and his conviction that the supplanting of “competitive market forces”, however feeble these forces may be, by internal organization is anti-competitive.\textsuperscript{158}

\textsuperscript{153} For an interpretation of the transformation of “inside contracting,” which was practiced by New England manufacturing firms in the late 1900’s, to vertical integration for transaction cost reasons, see Williamson, supra note 2, at 322-24.

\textsuperscript{154} Blake, supra note 149, at 562-63, 572-73.

\textsuperscript{155} Id. 571-72. Blake observes in this connection that “[O]ne objective of antitrust policy is to preserve a competitive system—a structure of the economy in which all economic units in the unregulated sector are subject to the continuing discipline of competitive market forces. The creation of vast conglomerate enclaves in which decisions with respect to resource use are insulated from these forces is inconsistent with the basic tenets of antitrust policy.” Id. 574.

I submit that, subject to the condition that the internal resource reallocations result in higher social as well as private yields—which is normally to be expected when investments are shifted from activities with lower to higher marginal profitability, one of the leading objectives of antitrust policy is being served.

\textsuperscript{156} For a discussion, see O. Williamson, supra note 22, at 143-44. For a fascinating study of the internal resource allocation process at work in one major corporation, see Hamilton & Moses, supra note 112. I concede that the system developed for and used by the International Utilities Corporation represents the leading edge of internal resource allocation capabilities in a conglomerate firm, but it is not an isolated instance. Firms such as ITT have had a similar, albeit less formal, approach to the internal resource allocation problem for years. See Address, Management Must Manage, by Harold Geneen, before the Investment Group of Hartford, Conn., Feb. 15, 1968.

\textsuperscript{157} Blake, supra note 149, at 566, 578.

\textsuperscript{158} Id. 574, 579. See note 155 supra.
Blake also finds conglomerates objectionable because of "hard evidence to support the no longer novel theory—and widely held belief in the business community—that large conglomerates facing each other in several markets tend to be less competitive in price than regional or smaller firms." There are two problems with the argument. First, I would scarcely characterize the evidence on which Blake relies as "hard." Part of the evidence cited by Blake is Scherer's discussion of the "spheres of influence hypothesis." But Scherer is very careful to characterize the evidence quite differently, noting that even with respect to the pre-war international chemical industry, which aside from marine cartels is his only Western example, the evidence is fragmentary. With respect to other industries he concludes that "there is a dearth of evidence on spheres of influence accords."

Second, the definition of a conglomerate requires attention. Are all specialized firms (such as National Tea, to which Blake earlier refers) that operate similar plants or stores in geographically dispersed markets really to be regarded as conglomerates? Stretching the definition of a conglomerate to include geographically dispersed, but otherwise specialized, enterprises, shrinks the number of nonconglomerate large firms to insignificance. If "conglomerate" is defined in terms of product diversification, Blake (and the Federal Trade Commission) ought to be expected to generate examples of abuse of conglomerate structure from the universe of product-diversified firms. If instead all large multimarket firms, whatever their product specialization ratios, are the objectionable subset, the suspect firms ought to be expressly identified in this way rather than by designating them as "conglomerates."

Although I share Blake's suspicions with respect to the behavior of very large product-diversified firms (which is the narrower definition of the conglomerate), the facts have yet to be assembled. As things stand now, the price-competitive-ness of such firms can not be adversely distinguished from that of other large multimarket organizations.

The data are somewhat better with respect to reciprocity. Blake conjectures in this connection that "empirical research,

---

159 Id. 570.
161 Id. 279.
162 Blake, supra note 149, at 557 n.13.
if it could be carried out, would show that reciprocity is as inevitable a result of widespread conglomerate structure as price rigidity is a consequence of oligopoly structure—where, apparently, the latter, and hence the former, is believed to be extensive. Jesse Markham's recent study of conglomerates, which was unavailable to Blake, suggests otherwise: "highly diversified companies are no more, and may be even less, given to reciprocity than large corporations generally."

Blake's principal policy proposal is that conglomerate acquisitions by firms above a specified size (the subset of firms that are to be restricted is not explicitly identified, but Blake makes several references to the top 200 firms) be accompanied by a spin-off of comparable assets. He further stipulates that no exception be permitted for acquiring a toehold in the new market. His argument against the toehold exception is that small, independent firms are more apt to engage in price competition than large conglomerates—relying a second time on the purportedly "hard" evidence referred to above—and contends that "a size based presumption would help restore the idea that internal growth is the normal, and usually the most socially efficient, means of industrial expansion, by making it the only means available to the largest corporations absent a special showing of procompetitive effect or of efficiencies."

As already indicated, however, the evidence on which Blake relies is rather limited. Moreover, the basis for his refusal to admit a toehold exception is really unclear. By itself the acquisition of a very small firm scarcely contributes much to the growth of the large firm. Correspondingly, requiring the large firm to release assets in an equivalent amount whenever a toehold acquisition is made is scarcely more than a nuisance. Furthermore, toehold acquisitions made for the pur-

---

163 Id. 569.
165 Blake, supra note 149, at 559-69.
166 Id. 590.
167 Id. 590-91.
168 For the purpose of size control, a large firm that engages in a series of toehold acquisitions within a specified time interval might be required to spin off assets comparable to the aggregate of those acquired if some absolute value is exceeded. Even small percentage positions in some industries (e.g., petroleum) can represent quite large absolute asset values. Individual toehold acquisitions in these circumstances might exceed the absolute asset value threshold of, say, $100 million. A spin-off might be indicated.
pose of securing a position that will subsequently be expanded is internal growth of the sort Blake favors. Either there is little point to Blake’s toehold argument, or he regards expansion by small firms as socially preferable to similar investments by large firms.

Assuming, arguendo, that the same investments will be made whether the small firm is acquired or not, it is easy to agree with Blake—though I repeat that the evidence on the competitive behavior of small firms, as compared with product divisions in diversified large firms, is scarcely dispositive. But it is doubtful that the same investments will actually occur. This raises transfer process issues.

An examination of these matters suggests that small firms apparently enjoy a comparative advantage at early and developmental stages of the technical innovation process. Large, established firms, by contrast, display comparative advantages at large scale commercial production and distribution stages. Not only may the management of the small firm lack the financial resources to move to the commercial stage in any but a gradualist manner because its credit standing does not permit it to raise significant blocs of capital except at adverse rates, but the management of the small firm may be poorly suited to make the transition. Different management skills and knowledge are required to bring a project successfully to large scale commercial development than may have been needed at earlier stages. If, because of management experience and team considerations similar to those described in section II above, the talents needed to facilitate internal expansion cannot be costlessly identified and assembled, transferring the project to an established firm that already possesses the requisite talents may be the more economical alternative.

169 However, see the qualification in note 168 supra.
171 Though it varies somewhat with organizational structure, projects for which only small scale commercialization is anticipated are not ones for which large firms are typically well suited. For a novel organizational “solution”, see Sabin, At Nuclepore They Don’t Work for G.E. Anymore, 88 FORTUNE, Dec. 1973, at 145.
172 Moving from a prototype to a commercial stage commonly involves a substantial investment in organizational infrastructure, much of which has no value should the enterprise fail. Lacking a known performance record and tangible assets to secure the investment, lenders are apprehensive to invest except on a sequential basis. The risks of opportunism, given information impactedness, are perceived as too great.
Again, it is transactions, not technology, which dictate this result. Put in these terms, it is unclear that the no toehold position survives.

I am nevertheless sympathetic with the proposition that the acquisition of already large firms by other large firms ought to be accompanied by a divestiture of equivalent assets. As Richard Hofstadter has observed, the support for antitrust enforcement rests less on a consensus among economists about its efficiency enhancing properties than it does on a political and moral judgment that power in the American economy should be diffused. The wisdom of such populist social and political attitudes is illustrated by the misadventures of the ITT Corporation in domestic and foreign affairs. Much of Blake’s disenchantment with conglomerates appears to be attributable to a concern that giant size and political abuse are positively correlated, and I would urge that the case be made expressly in these terms. If giant firms rather than all conglomerates are what is objectionable, attention ought properly to be restricted to these firms.

A requirement that very large firms divest themselves of equivalent assets when larger than toehold acquisitions are made is also favored by the prospect that this policy will help curb bureaucratic abuses associated with very large size. Although such divestitures sometimes occur voluntarily, such efforts predictably encounter bureaucratic resistance. If, however, such divestiture commonly has beneficial effects of an organizational self-renewal sort, making divestiture mandatory is scarcely objectionable. It would merely strengthen the hand of those in the firm who are anxious to forestall bureaucratic stagnation. Absent such a rule, internal agreement on divestiture may be difficult to secure; parties with vested interests will make partisan (opportunistic) representations that will be difficult to reject. Given such a rule, however, the general office can simply plead that it has no choice but to divest (assuming that a large acquisition is to be made). The preferences


175 Blake, supra note 149, at 574, 576, 578, 579, & 591. That giant size procures political favors does not imply that atomistic organization (e.g., farmers) is the favored economic alternative. Often with the latter, however, the favors are more likely to be transparent.

of the general office, reflecting efficiency considerations for the entire conglomerate enterprise, are thus made to prevail more fully.

C. Policy Implications

A transactional interpretation of the conglomerate, which emphasizes the limitations of capital markets with respect to policing corporate management, reveals that conglomerate firms (of the appropriate kind) are not altogether lacking in social purpose. If maintaining the market for corporate control\textsuperscript{177} is thought to be generally beneficial, if reallocating resources away from projects with lower returns to favor those with higher net private returns also generally yields social net benefits as well, and if the antitrust enforcement agencies are to maintain a tough policy with respect to horizontal and vertical mergers, a policy of moderation with respect to conglomerate mergers is in order. In particular, public policy with respect to conglomerate acquisitions should focus on (1) mergers where potential competition is meaningfully impaired, and (2) mergers by giant firms that are not accompanied by a spin-off (or other disposition) of comparable assets. Acquisitions of the second kind have been discussed above.\textsuperscript{178} Consider therefore the potential competition issue.

As I have already indicated, Blake’s views on potential competition are rather broad.\textsuperscript{179} The law, however, appears to be moving in the direction of interpreting the potential competition issue more narrowly. Commissioner Dennison, speaking for a unanimous Commission in the recent FTC decision \textit{Beatrice Foods Co.}, discussed the factual proof required to show that potential competition has been or probably will be reduced:

Complaint Counsel in essence attempt to rest their case on the existence of concentration ratios

\textsuperscript{177} For a discussion of the market for corporate control, see Manne, \textit{Mergers and the Market for Corporate Control}, 73 J. POL. ECON. 110 (1965).

\textsuperscript{178} Notes 152-59 \textit{supra} & accompanying text. Although I suspect that there is little real cost advantage that an already giant sized firm can confer on an acquired firm that could not be as (or more) effectively conferred by a somewhat less gargantuan enterprise, it may be useful not to prohibit such acquisitions altogether, so as to preserve the market for corporate control. If requiring the giant sized firm to divest itself of comparable assets tends to forestall bureaucratic stagnation in the firm and has beneficial political consequences, a reasonable result would seem to have been reached.

\textsuperscript{179} See text accompanying notes 150-57 \textit{supra}.
alone. The test for finding injury due to elimination of a potential competitor is not simple. Additional factors enter into any analysis of the loss of a potential competitor. Among these are: trends toward concentration in the market; extensive entry barriers; high probability that the lost potential competitor would have actually entered the market; whether the lost potential competitor was one of only a few such potential competitors and whether, if he had entered the market, his new competition would have had a significant impact on price and quality. Although the number of competing firms or trends toward concentration may be enough without more to condemn many horizontal mergers between existing rivals in a market, the condition of entry by new firms as well as these other factors mentioned above must be considered when dealing with elimination of a potential competitor.  

This reference to the condition of entry warrants additional development.

As Turner has argued forcefully, potential competition is apt to be impaired if one of a few most likely potential entrants acquires a firm that exceeds toehold proportions. If the industry in question is highly concentrated, so that, but for the threat of potential competition, competitive results will not reliably obtain, the quality of competition is degraded by the loss of one of a few "most likely potential entrants." I would like to urge that the appellation "most likely potential entrant" has genuine economic significance, as contrasted with transitory business significance, only to the extent that nontrivial barriers to entry into the industry in question can be said to exist.

The antitrust distinction to be made is between firms which (for transitory reasons) may have demonstrated an acquisition interest in the industry and firms which, despite entry barriers (nontransitory considerations), are strategically situated to enter. Because the interest of firms of the first kind is unlikely to persist, being dependent on such factors as the current interests of the chief executive, temporary cash balances, and immediate income statement considerations,

---


181 Turner, Conglomerate Mergers and Section 7 of the Clayton Act, 78 Harv. L. Rev. 1313 (1965).
prohibiting entry by acquisition to such firms is of little affirmative economic purpose. No long term benefit to potential competition is thereby secured. Rather, the principle effect is to shrink the acquisition market, thereby impairing both the market for corporate control and the incentives for entrepreneurs to invest in new enterprises.

The situation is quite different, however, if the industry in question has nontrivial barriers to entry and the firm evidencing an acquisition interest is one of only a few firms for which de novo or toehold entry would be very easy. Consider in this connection the entry barrier conditions identified by Joe Bain, namely, economies of scale that are large in relation to the size of the market, absolute cost advantages, and product differentiation. Although Bain describes these barriers without reference to specific firms, plainly the height of the barrier varies among possible entrants. Thus, though economies of scale may be large in relation to the size of the market, this impediment to entry is apt to be less severe for those few firms which have closely complementary production processes and sales organizations. Similarly, a few firms may be well-situated with respect to absolute cost advantages. Although patents may constitute a severe impediment to entry, high grade ore deposits may be in limited supply, or specialized labor skills may be required, a few firms are apt to stand out from all the rest by reason of a complementary technology, which facilitates inventing around the established patents, because they possess medium grade ore deposits, or because their labor force has acquired, in a learning by doing fashion, the requisite specialized skills. Product differentiation advantages are likewise attenuated for those firms that market related types of consumer goods and themselves enjoy brand recognition. Ceteris paribus, those firms for which the barriers are least are the firms that are usefully designated most likely potential entrants.

In circumstances, however, where all such barriers to entry are negligible (economies of scale are not great; patents and specialized or otherwise scarce resources are unimportant; product differentiation is insubstantial), no small subset of firms can be said to enjoy a strategic advantage. In that case, it is fatuous to attempt to identify a group of most likely

---

182 J. Bain, Barriers to New Competition (1956).
potential entrants the loss by acquisition of any of which would significantly impair the quality of potential competition.\(^{183}\)

V. Concluding Remarks

Received microtheory provides the analyst with some very powerful tools, but it is also incomplete. Among other things, as Peter Diamond has noted, standard "economic models . . . [treat] individuals as playing a game with fixed rules which they obey. They do not buy more than they know they can pay for, they do not embezzle funds, they do not rob banks."\(^{184}\) Expressed in terms of the language introduced in section I, individuals are not opportunists. Standard models also, as Simon has repeatedly emphasized, impute considerable power of computation and analysis to economic actors\(^{185}\)—which is to say that bounded rationality is rarely thought to pose a problem. The transaction cost approach relaxes both of these behavioral assumptions.

Although there is no necessary connection, those who rely exclusively on the received microtheory model of the firm are prone to express considerable confidence in the efficacy of competition. Problems of small numbers supply and of adapting efficiently to uncertainty are apt to be dismissed or settled in a rather artificial fashion. The upshot is that many of the interesting problems of economic organization are either finessed or dealt with in a dogmatic way.

The transaction cost approach is concerned with the costs of running the economic system, especially the costs of adapting efficiently to uncertainty. It expressly makes allowance for elementary attributes of human decision makers—in particular, bounded rationality and opportunism—and permits the implications of these conditions to be explored in a way that received microtheory does not.

This does not require that received microtheory be rejected, however. Transaction cost analysis is more a complement to than a substitute for received microtheory. It is appro-

\(^{183}\) One might, however, wish to prevent entry by acquisition by "dominant firms," the presence of which discourages rivalry (for deep pocket reasons) and otherwise transforms the market in uncertain ways. The Procter & Gamble acquisition of Clorox has been characterized by Justice Marshall in these terms. United States v. Falstaff Brewing Corp., 410 U.S. 526, 558-59 (1973) (Marshall, J., concurring).

\(^{184}\) Diamond, Comment, in FRONTIERS OF QUANTITATIVE ECONOMICS 29, 31 (M. Intriligator ed. 1971).

\(^{185}\) H. Simon, supra note 14, at 198-99.
priate for studying the frictions in the system which may prevent the implications of received microtheory from going through. This focus makes it especially wellsuited to help delimit the public policy issues with which the antitrust enforcement agencies are concerned. Moreover, transaction cost analysis is comparatively valuefree: it is biased neither for nor against the modes of organization associated with an unfettered market.

Perhaps the simplest application of the transaction cost approach is to price discrimination. Not only does transaction cost analysis call attention to the fact that price discrimination is costly to effectuate, which has been apparent to any student who has given serious consideration to the issue, but it identifies the reasons for this and permits additional efficiency implications to be derived. The usual proposition that allocative efficiency is improved by fully discriminating monopoly, as compared with uniform price monopoly, is challenged. A private net gain but social net loss can plainly obtain when transaction costs are expressly introduced into the net benefit calculus.

With respect to vertical integration, the transaction cost approach counsels caution. The more strident claims of those who proclaim vertical integration (and, more generally, vertical market restrictions of all kinds) to be altogether innocent of anticompetitive potential are shown to be exaggerated. Vertical integration can have entry impeding consequences in highly concentrated industries if capital markets do not operate frictionlessly—which in this context means omnisciently. Where, however, the industry in question is not highly concentrated, this same anticompetitive potential is much less severe. Absent collusion, the presumption that vertical integration is innocent or beneficial is appropriate.

The transaction cost approach also reveals that the oligopoly problem should not be uncritically equated with the dominant firm problem. It is much more difficult to negotiate a comprehensive collusive agreement, and there are many more problems to effecting a joint profit maximizing outcome, than is commonly suggested. Accordingly, theories of "shared monopoly" ought to be regarded with skepticism. An economically rational antitrust policy would presumably first address the industries with dominant firms and, where feasible, effect

dissolution here before going on to attack oligopolies. Contrary to what is sometimes said, there are prospective benefits from converting a dominant firm industry into an oligopolistic one.

The broadside attack that some have leveled against conglomerates appears to be overdrawn. Again, frictions in the capital market turn out to be of fundamental importance. Absent capital market frictions impeding takeover or proclivities of incumbent managements to reinvest earnings (or otherwise behave in opportunistic ways), the conglomerate appears to lack compelling economic purpose of a socially redeeming kind. Since enthusiasts of received microtheory have been reluctant to concede that a corporate control problem has even existed, they have had little to offer in the way of a rationale for the conglomerate firm. Once such frictions are admitted, however, there is plainly a case for encouraging, or at least not impeding, organizational innovations which have the potential to attenuate internal organizational distortions of a managerial discretion kind. Subject to the qualifications about organization form, which I have repeatedly emphasized, the conglomerate has attractive properties both because it makes the market for corporate control more credible, thereby inducing self-policing among otherwise opportunistic managements, and because it promotes the reallocation of resources to high yield uses. Except, therefore, among giant sized firms, where the risk of offsetting political distortions is seriously posed, a more sympathetic posture on the part of the antitrust enforcement agencies towards conglomerates would seem warranted.

Donald Dewey has described the role of economists in antitrust as follows: "The important issues in the control of monopoly are 'economic' in the sense that judges and administrators are compelled to make decisions in the light of what they think the business world is 'really' like, and it is the task of economists through research and reflection to provide them with an increasingly accurate picture."\(^{187}\) To the consternation of administrators and judges alike, the picture provided by received microtheory is sometimes vague and at other times simplistic. Transaction cost analysis is intended to supplement received microtheory in such circumstances.