

# The Vertical Integration of Production: Market Failure Considerations\*

By OLIVER E. WILLIAMSON  
*University of Pennsylvania*

The study of vertical integration has presented difficulties at both theoretical and policy levels of analysis. That vertical integration has never enjoyed a secure place in value theory is attributable to the fact that, under conventional assumptions, it is an anomaly: if the costs of operating competitive markets are zero, "as is usually assumed in our theoretical analysis" (Arrow, 1969, p. 48), why integrate?

Policy interest in vertical integration has been concerned mainly with the possibility that integration can be used strategically to achieve anticompetitive effects. In the absence of a more substantial theoretical foundation, vertical integration, as a public policy matter, is typically regarded as having dubious if not outright antisocial properties. Technological interdependencies or, possibly, observational economies, constitute the principal exceptions.

The technological interdependency argument is both the most familiar and the most straight-forward: successive processes which, naturally, follow immediately in time and place dictate certain efficient manufacturing configurations; these, in turn, are believed to have common ownership implications. Such technical complementarity is probably more important in flow process operations (chemicals, metals, etc.) than in separable component

manufacture. The standard example is the integration of iron and steel-making, 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" (Bain, 1968, p. 381).

There is, nevertheless, a distinct unease over the argument. This is attributable, probably, to a suspicion that the firm is more than a simple efficiency instrument, in the usual scale economies and least-cost factor proportions senses of the term, but also possesses coordinating potential that sometimes transcends that of the market. It is the burden of the present argument that this suspicion is warranted. In more numerous respects than are commonly appreciated, the substitution of internal organization for market exchange is attractive less on account of technological economies associated with production but because of what may be referred to broadly as "transactional failures" in the operation of markets for intermediate goods. This substitution of internal organization for market exchange will be referred to as "internalization."

The two principal prior contributions on which the argument relies are Coase's seminal discussion on "The Nature of The Firm" (1937) and Arrow's more recent review of market versus nonmarket allo-

\* Research on this paper has been supported by a grant from the Brookings Institution. It is part of the larger study referred to in footnote 1. Helpful comments from Noel Edelson, Stefano Fenoaltea, Julius Margolis, and Almarin Phillips are gratefully acknowledged.

cation (1969). As will be evident, I agree with Malmgren (1961) that the analysis of transaction costs is uninteresting under fully stationary conditions and that only when the need to make unprogrammed adaptations is introduced does the market versus internal organization issue become engaging.

But while Malmgren finds that the advantage of the firm inheres in its capacity to control information and achieve plan consistency among interdependent activities, which may be regarded as an information processing advantage, I mainly emphasize the differential incentive and control properties of firms in relation to markets. This is not to suggest that information processing considerations are unimportant, but rather that these incompletely characterize the distinctive properties of firms that favor internal organization as a market substitute.

### I. *Internal Organization: Affirmative Aspects*

A complete treatment of vertical integration requires that the limits as well as the powers of internal organization be assessed. As the frictions associated with administrative coordination become progressively more severe, recourse to market exchange becomes more attractive, *ceteris paribus*. It is beyond the scope of this paper, however, to examine the organizational failure aspect of the vertical integration question.<sup>1</sup> Rather it is simply asserted that, mainly on account of bounded rationality and greater confidence in the objectivity of market exchange in comparison with bureaucratic processes, market intermediation is generally to be preferred over internal supply in circumstances in

which markets may be said to "work well."<sup>2</sup>

The properties of the firm that commend internal organization as a market substitute would appear to fall into three categories: incentives, controls, and what may be referred to broadly as "inherent structural advantages." In an incentive sense, internal organization attenuates the aggressive advocacy that epitomizes arms length bargaining. Interests, if not perfectly harmonized, are at least free of representations of a narrowly opportunistic sort; in any viable group, of which the firm is one, the range of admissible intraorganizational behavior is bounded by considerations of alienation. In circumstances, therefore, where protracted bargaining between independent parties to a transaction can reasonably be anticipated, internalization becomes attractive.<sup>3</sup>

Perhaps the most distinctive advantage of the firm, however, is the wider variety and greater sensitivity of control instruments that are available for enforcing intrafirm in comparison with interfirm activities (Williamson, 1970). Not only does the firm have the constitutional authority and low-cost access to the requisite data which permit it to perform more precise own-performance evaluations (of both a contemporaneous and *ex post* variety)

<sup>2</sup> An intermediate market will be said to work well if, both presently and prospectively, prices are non-monopolistic and reflect an acceptable risk premium, and if market exchange experiences low transaction costs and permits the realization of essential economies. To the extent that the stipulated conditions do not hold, internal supply becomes relatively more attractive, *ceteris paribus*.

<sup>3</sup> Common ownership by itself, of course, does not guarantee goal consistency. A holding company form of organization in which purchaser and supplier are independent divisions, each maximizing individual profits, is no solution. Moreover, merely to stipulate joint profit maximization is not by itself apt to be sufficient. The goal needs to be operationalized, which involves both rulemaking (with respect, for example, to transfer pricing) and the design of efficacious internal incentives. For a discussion, see Williamson (1970).

<sup>1</sup> I discuss the organizational failure dimension of this issue in *Aspects of Monopoly Theory and Policy* (forthcoming). Policy implications of the argument are also examined there.

than can a buyer, but its reward and penalty instruments (which include selective use of employment, promotion, remuneration, and internal resource allocation processes) are more refined.

Especially relevant in this connection is that, when conflicts develop, the firm possesses a comparatively efficient conflict resolution machinery. To illustrate, fiat is frequently a more efficient way to settle minor conflicts (say differences of interpretation) than is haggling or litigation. *Interorganizational* conflict can be settled by fiat only rarely, if at all. For one thing, it would require the parties to agree on an impartial arbitrator, which agreement itself may be costly to secure. It would also require that rules of evidence and procedure be established. If, moreover, the occasion for such interorganizational settlements were to be common, the form of organization converges in effect to vertical integration, with the arbiter becoming a manager in fact if not in name. By contrast, *intraorganizational* settlements by fiat are common (Whinston, 1964, pp. 410–14).

The firm may also resort to internalization on account of economies of information exchange. Some of these may be due to structural differences between firms and markets. Others, however, reduce ultimately to incentive and control differences between internal and market organization. It is widely accepted, for example, that communication with respect to complex matters is facilitated by a common training and experience and if a compact code has developed in the process. Repeated interpersonal interactions may permit even further economies of communication; subtle nuances may come through in familiar circumstances which in an unfamiliar relationship could be achieved only with great effort. Still, the drawing of an organizational boundary need not, by itself, prevent intensely fa-

miliar relations from developing between organizations. Put differently, but for the goal and control differences described above, the informational advantages of internal over market organization are not, in this respect, apparent. Claims of informational economies thus should distinguish between economies that are attributable to information flows *per se* (structure) and those which obtain on account of differential veracity effects (see Part D, Section II).

## II. *Market Failure Considerations*

What are referred to here as market failures are failures only in the limited sense that they involve transaction costs that can be attenuated by substituting internal organization for market exchange. The argument proceeds in five stages. The first three are concerned with characterizing a successively more complex bargaining environment in which small numbers relations obtain. The last two deal with the special structural advantages which, either naturally or because of prevailing institutional rules, the firm enjoys in relation to the market.

### A. *Static Markets*

Consider an industry that produces a multicomponent product, assume that some of these components are specialized (industry specific), and assume further that among these there are components for which the economies of scale in production are large in relation to the market. The market, then, will support only a few efficient sized producers for certain components.

A monopolistic excess of price over cost under market procurement is commonly anticipated in these circumstances—although, as Demsetz (1968) has noted, this need not obtain if there are large numbers of suppliers willing and able to bid at the initial contract award stage. As-

sume, however, that large numbers bidding is not feasible. The postulated conditions then afford an "apparent" incentive for assemblers to integrate backward or suppliers to integrate forward. Two different cases can be distinguished: bilateral monopoly (oligopoly) and competitive assembly with monopolistic supply. The former is considered here; the latter is treated in Part C.

Bilateral monopoly requires that both price and quantity be negotiated. Both parties stand to benefit, naturally, by operating on rather than off the contract curve—which here corresponds to the joint profit maximizing quantity (Fellner, 1947). But this merely establishes the amount to be exchanged. The terms at which this quantity will be traded still need to be determined. Any price consistent with nonnegative profits to both parties is feasible. Bargaining can be expected to ensue. Haggling will presumably continue until the marginal private net benefits are perceived by one of the parties to be zero. Although this haggling is jointly (and socially) unproductive, it constitutes a source of private pecuniary gain. Being, nevertheless, a joint profit drain, an incentive to avoid these costs, if somehow this could be arranged, is set up.

One possible adaptation is to internalize the transaction through vertical integration; but a once-for-all contract might also be negotiated. In a perfectly static environment (one that is free of disturbances of all kinds), these may be regarded with indifference: the former involves settlement on component supply price while merger requires agreement on asset valuation. Bargaining skills will presumably be equally important in each instance (indeed, a component price can be interpreted in asset valuation terms and conversely). Thus, although vertical integration may occur under these conditions, there is nothing in the nature of the prob-

lem that requires such an outcome.

A similar argument in these circumstances also applies to adaptation against externalities: joint profit considerations dictate that the affected parties reach an accommodation, but integration holds no advantage over once-for-all contracts in a perfectly static environment.

Transforming the relationship from one of bilateral monopoly to one of bilateral oligopoly broadens the range of bargaining alternatives, but the case for negotiating a merger agreement in relation to a once-for-all contract is not differentially affected on this account. The static characterization of the problem, apparently, will have to be relaxed if a different result is to be reached.

#### B. *Contractual Incompleteness*

Let the above conditions be enriched to include the stipulation that the product in question is technically complex and that periodic redesign and/or volume changes are made in response to changing environmental conditions. Also relax the assumption that large numbers bidding at the initial contract award stage is infeasible. Three alternative supply arrangements can be considered: a once-for-all contract, a series of short-term contracts, and vertical integration.

The dilemma posed by once-for-all contracts is this: lest independent parties interpret contractual ambiguities to their own advantage, which differences can be resolved only by haggling or, ultimately, litigation, contingent supply relations ought exhaustively to be stipulated. But exhaustive stipulation, assuming that it is feasible, is itself costly. Thus although, if production functions were known, appropriate responses to final demand or factor price changes might be deduced, the very costliness of specifying the functions and securing agreement discourages the effort. The problem is made even more severe

where a changing technology poses product redesign issues. Here it is doubtful that, despite great effort and expense, contractual efforts reasonably to comprehend the range of possible outcomes will be successful. An adaptive, sequential decision process is thus indicated. If, however, contractual revisions or amendments are regarded as an occasion to bargain opportunistically, which predictably they will be, the purchaser will defer and accumulate adaptations, if by packaging them in complex combinations their true value can better be disguised; some adaptations may be foregone altogether. The optimal sequential decision-making process can in these respects be distorted.

Short-term contracts, which would facilitate adaptive, sequential decision-making, might therefore be preferred. These pose problems, however, if either (1) efficient supply requires investment in special-purpose, long-life equipment, or (2) the winner of the original contract acquires a cost advantage, say by reason of "first mover" advantages (such as unique location or learning, including the acquisition of undisclosed or proprietary technical and managerial procedures and task-specific labor skills).

The problem with condition (1) is that optimal investment considerations favor the award of a long-term contract so as to permit the supplier confidently to amortize his investment. But, as indicated, long term contracts pose adaptive, sequential decision-making problems. Thus optimal investment and optimal sequential adaptation processes are in conflict in this instance.

It might be argued that condition (2) poses no problems since initial bidders will fully reflect in their original bids all relevant factors. Thus, although anticipated downstream cost advantages (where downstream is used both here and subsequently in the sense of time rather than

place) will give rise to small numbers competition for downstream supply, competition at the initial award stage is sufficient to assure that only competitive returns will be realized over the entire supply interval. One might expect, therefore, that the low bidder would come in at a price below cost in the first period, set price at the level of alternative supply price in later periods, and earn normal returns over-all. Appropriate changes can be introduced easily at the recontracting interval.

A number of potential problems are posed, however. For one thing, unless the total supply requirements are stipulated, "buying in" strategies are risky. Also, and related, the alternative supply price is not independent of the terms that the buyer may subsequently offer to rivals. Moreover, alternative supply price is merely an upper bound; an aggressive buyer may attempt to obtain a price at the level of current costs on each round. Haggling could be expected to ensue. Short-term contracts thus experience what may be serious limitations in circumstances where nontrivial first-mover advantages obtain.

In consideration, therefore, of the problems that both long and short-term contracts are subject to, vertical integration may well be indicated. The conflict between efficient investment and efficient sequential decision-making is thereby avoided. Sequential adaptations become an occasion for cooperative adjustment rather than opportunistic bargaining; risks may be attenuated; differences between successive stages can be resolved more easily by the internal control machinery.

It is relevant to note that the technological interdependency condition involving flow process economies between otherwise separable stages of production is really a special case of the contractual incompleteness argument. The contractual dilemma

is this: On the one hand, it may be prohibitively costly, if not infeasible, to specify contractually the full range of contingencies and stipulate appropriate responses between stages. On the other hand, if the contract is seriously incomplete in these respects but, once the original negotiations are settled, the contracting parties are locked into a bilateral exchange, the divergent interests between the parties will predictably lead to individually opportunistic behavior and joint losses. The advantages of integration thus are not that technological (flow process) economies are unavailable to nonintegrated firms, but that integration harmonizes interests (or reconciles differences, often by fiat) and permits an efficient (adaptive, sequential) decision process to be utilized. More generally, arguments favorable to integration that turn on "supply reliability" considerations commonly reduce to the contractual incompleteness issue.<sup>4</sup>

### C. Strategic Misrepresentation Risk

Contractual incompleteness problems develop where there is *ex ante* but not necessarily *ex post* uncertainty. Strategic misrepresentation risks are serious where there is uncertainty in both respects. Not only is the future uncertain but it may not be possible, except at great cost, for an outside agency to establish accurately what has transpired after the fact. The advantages of internalization reside in the

facts that the firm's *ex post* access to the relevant data is superior, it attenuates the incentives to exploit uncertainty opportunistically, and the control machinery that the firm is able to activate is more selective.

1. AFFIRMATIVE OCCASIONS for INTEGRATION. Three affirmative occasions to integrate on account of strategic misrepresentation risk and two potentially anticompetitive consequences of integration can be identified.

(a) MORAL HAZARD. The problem here arises because of the conjoining of inharmonious incentives with uncertainty—or, as Arrow puts it (1969, p. 55), it is due to the "confounding of risks and decisions." To illustrate, consider the problem of contracting for an item the final cost and/or performance of which is subject to uncertainty. One possibility is for the supplier to bear the uncertainty. But, he will undertake a fixed price contract to deliver a specified result the costs of which are highly uncertain only after attaching a risk premium to the price. Assume that the buyer regards this premium as excessive and is prepared on this account to bear the risk himself. The risk can easily be shifted by offering a cost-plus contract. But this impairs the incentives of the supplier to achieve least-cost performance; the supplier may reallocate his assets in such a way as to favor other work to the disadvantage of the cost-plus contract.

Thus, although, if commitments were self-enforcing, it might often be institutionally most efficient to divide the functions of risk bearing and contract execution (that is, cost-plus contracts would have ideal properties), specialization is discouraged by interest disparities. At a minimum, the buyer may insist on monitoring the supplier's work. In contrast therefore to a fixed-price contract, where it is sufficient to evaluate end-product performance, cost-plus contracts, because

<sup>4</sup> It is sometimes suggested that breach of contract risk affords an additional reason for integration: the small supplier of a critical component whose assets are insufficient to cover a total damage claim leaves the purchaser vulnerable. But this is an argument against small suppliers, not contracting quite generally; the large, diversified supplier might well have superior risk pooling capability to that of the integrated firm. The risks of contractual incompleteness, however, remain and may discourage purchasing from large, diversified organizations. For a discussion of "ideal" contracts in this connection, see Arrow (1965, pp. 52-53).

they expose the buyer to risks of inefficient (high cost) contract execution, require that *both* inputs and outputs be evaluated.

Internalization does not eliminate the need for input evaluation. Rather, the advantage of internalization, for input monitoring purposes, resides in the differential ease with which controls are exercised. An external agency, by design, lacks recourse to the internal control machinery: proposed remedies require the consent of the contractor and then are highly circumscribed; unrestricted access by the buyer to the contractor's internal control machinery (including selective use of employment, promotion, remuneration, and internal resource allocation processes) is apt to be denied. In consideration of the costs and limitations of input monitoring by outsiders, the buyer may choose instead to bear the risk and perform the work himself. The buyer thus internalizes, through backward vertical integration, a transaction which, but for uncertainty, would move through the market. A cost-type contract for *internal procurement* is arranged.

(b) EXTERNALITIES/IMPUTATION. The externality issue can be examined in two parts. First, has a secure, unambiguous, and "appropriate" assignment of property rights been made? Second, are the accounting costs of imputing costs and benefits substantial? If answers to these questions are affirmative and negative respectively, appropriability problems will not become an occasion for vertical integration. Where these conditions are not satisfied, however, integration may be indicated.

The assignment aspect of this matter is considered in Part E below. Here it is assumed that an efficacious assignment of property rights has been made and that only the expense of imputing costs and benefits is at issue. But indeed this is apt often to be the more serious problem.

High imputation expenses which discourage accurate metering introduce ambiguity into transactions. Did party A affect party B and if so in what degree? In the absence of objective, low cost standards, opposed interests can be expected to evaluate these effects differently. Internalization, which permits protracted (and costly) disputes over these issues to be avoided, may on this account be indicated.

(c) VARIABLE PROPORTIONS DISTORTIONS. Consider the case where the assembly stage will support large numbers; fewness appears only in component supply. Whether monopolistic supply prices provide an occasion for vertical integration in these circumstances depends both on production technology and policing expense. Variable proportions at the assembly stage afford opportunities for nonintegrated assemblers to adapt against monopolistically priced components by substituting competitively priced factors (McKenzie, 1951). Although conceivably the monopolistic component supplier could stipulate, as a condition of sale, that fixed proportions in assembly should prevail, the effectiveness of such stipulations is to be questioned—since, ordinarily, the implied enforcement costs will be great. Where substitution occurs, inefficient factor proportions, with consequent welfare losses, will result. The private (and social) incentives to integrate so as to reduce total costs by restoring efficient factor combinations are evident.

2. ANTICOMPETITIVE CONSEQUENCES. Anticompetitive effects of two types are commonly attributed to integration: price discrimination and barriers to entry (cf. Stigler, 1968, p. 303).

(a) PRICE DISCRIMINATION. The problem here is first to discover differential demand elasticities, and secondly to arrange for sale in such a way as to preclude reselling. Users with highly elastic demands which purchase the item at a low price

must not be able to service inelastic demand customers by acting as a middleman; all sales must be final. Although vertical integration may facilitate the discovery of differential elasticities, it is mainly with respect to the non-resale condition that it is regarded as especially efficacious.

Integration, nevertheless, is a relatively extreme response. Moreover, price discrimination is clearly practiced in some commodities without recourse to vertical integration (witness electricity and telephone service). What are the distinguishing factors? Legality considerations aside, presumably it is the cost of enforcing (policing) terms of the contract that are at issue. Some commodities apparently have self-enforcing properties—which may obtain on account of high storage and repacking costs or because reselling can not be arranged inconspicuously. The absence of self-enforcing (policing) properties is what makes vertical integration attractive as a means of accomplishing discrimination.

(b) ENTRY BARRIER EFFECTS. That the vertical integration of production might be used effectively to bar entry is widely disputed. Bork (1969, p. 148) argues 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 at once 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." But the issue is not one of profit avoidance but rather involves cost incidence. If borrowers are confronted by increasingly adverse rates as they increase their finance requirements, which Hirshleifer suggests is a distinct possibility (1970, pp. 200–1), cost may not be independent of vertical structure.

Assuming that vertical integration has the effect of increasing capital requirements, the critical issues are to what ex-

tent and for what reasons the supply curve of finance behaves in the way postulated. The following conjecture is offered as a partial explanation: unable to monitor the performance of *large, complex* organizations in any but the crudest way or to effect management displacement easily except on evidence of seriously discreditable error, investors demand larger returns as finance requirements become progressively greater, *ceteris paribus*. Thus the costs of policing against the contingency that managers will operate a rival enterprise opportunistically are, on this argument, at least partly responsible for the reputed behavior of the supply curve of capital. In consideration of this state of affairs, established firms may use vertical integration strategically to increase finance requirements and thereby to discourage entry if potential entrants feel compelled, as a condition of successful entry, to adopt the prevailing structure—as they may if the industry is highly concentrated.

#### D. Information Processing Effects

As indicated in Section I, one of the advantages of the firm is that it realizes economies of information exchange. These may manifest themselves as information impactedness, observational economies, or what Malmgren (1961) refers to as the "convergence of expectations."

1. INFORMATION IMPACTEDNESS. Richardson illustrates the problems of information impactedness by reference to an entrepreneur who was willing to offer long-term contracts (at normal rates of return, presumably) but which contracts others were unprepared to accept because they were not convinced that he had "the ability, as well as the will, to fulfill them. He may have information sufficient to convince himself that this is the case, but others may not" (Richardson, 1960, p. 83). He goes on to observe that the perceived



risks of the two parties may be such as to make it difficult to negotiate a contract that offers commensurate returns to each; objective risks are augmented by contractual risks in these circumstances. Integration undertaken for this reason is akin to self-insurance by individuals who know themselves to be good risks but are priced out of the insurance market because of their inability, at low cost, to "reveal" this condition to insurers.

2. OBSERVATIONAL ECONOMIES. As Radner indicates, "the acquisition of information often involves a 'set-up cost'; i.e., the resources needed to obtain the information may be independent of the scale of the production process in which the information is used" (Radner, 1970, p. 457). Although Radner apparently had horizontal firm size implications in mind, the argument also has relevance for vertical integration. If a single set of observations can be made that is of relevance to a related series of production stages, vertical integration may be efficient.

Still, the question might be raised, why common ownership? Why not an independent observational agency that sells information to all comers? Or, if the needed information is highly specialized, why not a joint venture? Alternatively, what inhibits efficient information exchange between successive stages of production according to contract? In relation, certainly, to the range of intermediate options potentially available, common ownership appears to be an extreme response. What are the factors which favor this outcome?

One of the problems with contracts is that of specifying terms. But even if terms could be reached, there is still a problem of policing the agreement. To illustrate, suppose that the common information collection responsibilities are assigned by contract to one of the parties. The purchasing party then runs a *veracity risk*: information may be filtered and possibly

distorted to the advantage of the firm that has assumed the information collection responsibility. If checks are costly and proof of contractual violation difficult, contractual sharing arrangements manifestly experience short-run limitations. If, in addition, small numbers prevail so that options are restricted, contractual sharing is subject to long-run risks as well. On this argument, observational economies are mainly to be attributed to strategic misrepresentation risks rather than to indivisibilities.

3. CONVERGENCE OF EXPECTATIONS. The issue to which the convergence of expectations argument is addressed is that, if there is a high degree of interdependence among successive stages of production and if occasions for adaptation are unpredictable yet common, coordinated responses may be difficult to secure if the separate stages are operated independently. March and Simon (1958, p. 159) characterize the problem in the following terms:

Interdependence by itself does not cause difficulty if the pattern of interdependence is stable and fixed. For, in this case, each subprogram can be designed to take account of all the subprograms with which it interacts. Difficulties arise only if program execution rests on contingencies that cannot be predicted perfectly in advance. In this case, coordinating activity is required to secure agreement about the estimates that will be used as the basis for action, or to provide information to each subprogram unit about the activities of the others.

This reduces, in some respects, to a contractual incompleteness argument. Were it feasible exhaustively to stipulate the appropriate conditional responses, coordination could proceed by contract. This is ambitious, however; in the face of a highly variable and uncertain environment, the attempt to program responses is apt to be inefficient. To the extent that an unprogrammed (adaptive, sequential)

decision process is employed instead, and in consideration of the severe incentive and control limitations that long-term contracts experience in these circumstances (See Part B above), vertical integration may be indicated.

But what of the possibility of short-term contracts? It is here that the convergence of expectations argument is of special importance. Thus assume that short-term contracts are not defective on account either of investment disincentives or first-mover advantages. It is Malmgren's (1961) contention that such contracts may nevertheless be vitiated by the absence of structural constraints. The costs of negotiations and the time required to bring the system into adjustment by exclusive reliance on market (price) signals are apt to be great in relation to that which would obtain if successive states were integrated and administrative processes employed as well or instead.

#### *E. Institutional Adaptations*

Institutional adaptations of two types are distinguished: simple economic and extra-economic.

1. **SIMPLE ECONOMIC.** As has been noted by others, vertical integration may be a device by which sales taxes on intermediate products are avoided, or a means by which to circumvent quota schemes and price controls (Coase, 1937, pp. 338-39; Stigler, 1968, pp. 136-37). But vertical intergration may also be undertaken because of the defective specification of property rights.

Although the appropriate assignment of property rights is a complex question, it reduces (equity considerations aside) to a simple criterion: What assignment yields maximum total product (Coase, 1960, p. 34)? This depends jointly on imputation and negotiation expenses and on the incentives of the compensated party. So as to focus on the negotiation expense aspect,

assume that imputation expenses are negligible and set the incentive question aside for the moment.<sup>5</sup> An "appropriate" assignment of property rights will here be defined as one which automatically yields compensation in the amount of the external benefit or cost involved, while an "inappropriate" assignment is one that requires bargaining to bring the parties into adjustment. Thus if A and B are two parties and A's activity imposes costs on B, the appropriate assignment of property rights is to require A to compensate B. If instead property rights were defined such that A is not required to compensate B, and assuming that the externality holds at the margin, efficient adaptation would occur only if B were to bribe A to bring his activity into adjustment—which entails bargaining. Only if the costs of such bargaining are neglected can the alternative specifications of property rights be said to be equivalent. For similar reasons, if A's activity generates benefits for B, the appropriate specification of property rights will be to require B fully to compensate A. Harmonizing the otherwise divergent interests of the two parties by internalizing the transaction through vertical merger

<sup>5</sup> As Coase has emphasized (1960, pp. 32-33, 41), compensation can impair the incentives of the compensated party that experiences an external cost to take appropriate protective measures. Parties that are assured of compensation will be content to conduct business as usual. Such a practice easily contributes to greater social cost than would obtain were compensation denied. A sensitivity to what, in a broad sense, might be regarded as contributory negligence is thus required if the system is to be brought fully into adjustment. Clairvoyance with respect to contributory negligence would of course permit the courts to supply those who experience the external cost with requisite incentives to adapt appropriately. Since, however, such clairvoyance (or even unbiasedness) cannot routinely be presumed, internalizing the transaction through vertical integration may be indicated for this reason as well. (Interestingly, a symmetrical problem is not faced where the externality is a benefit. Stipulating that compensation shall be paid induces Meade's (1952) orchard grower not merely to extend his production appropriately, but also to shift from apples to peaches if this is socially advantageous.)

promises to overcome the haggling costs which result when property rights are left either undefined or inappropriately specified.

2. OTHER. Risk aversion refers to the degree of concavity in the utility valuation of pecuniary outcomes. Decision-makers who are risk averse will be concerned not merely with the expected value, but also with the dispersion in outcomes associated with alternative proposals: the greater the dispersion, the lower the utility valuation. *Ceteris paribus*, decision-makers who are the less risk averse will presumably assume the risk bearing function. Even, however, if attitudes toward risk were identical—in the sense that every individual (for any given set of initial endowments) would evaluate a proposal similarly—differing initial asset positions among the members of a population could warrant a specialization of the risk-bearing function, with possible firm and market structure effects (Knight, 1965).

Arrow calls attention to norms of social behavior, including ethical and moral codes. He observes in this connection that "It is useful for individuals to have some trust in each other's word. In the absence of trust, it would become very costly to arrange for alternative sanctions and guarantees, and many opportunities for mutually beneficial cooperation would have to be foregone" (1969, p. 62). One would expect, accordingly, that vertical integration would be more complete in a low-trust than a high-trust culture, *ceteris paribus*.

### III. Conclusions

That product markets have remarkable coordinating properties is, among economists at least, a secure proposition. That product markets are subject to failure in various respects and that internal organization may be substituted against the

market in these circumstances is, if somewhat less familiar, scarcely novel. A systematic treatment of market failure as it bears on vertical integration, however, has not emerged.

Partly this is attributable to inattention to internal organization: the remarkable properties of firms that distinguish internal from market coordination have been neglected. But the fragmented nature of the market failure literature as it bears on vertical integration has also contributed to this condition; the extensive variety of circumstances in which internalization is attractive tends not to be fully appreciated.

The present effort attempts both to address the internal organization issue and to organize the market failure literature as it relates to vertical integration in a systematic way. The argument, however, by no means exhausts the issues that vertical integration raises. For one thing, the discussion of market failures may be incomplete in certain respects. For another, a parallel treatment of the sources and consequences of the failures of internal organization as they relate to vertical integration is needed. Third, the argument applies strictly to the vertical integration of production; although much of it may have equal relevance to backward vertical integration into raw materials and forward integration into distribution, it may have to be delimited in significant respects. Fourth, game theoretic considerations, which may permit the indicated indeterminacy of small numbers bargaining situations to be bounded, have been neglected. Finally, nothing in the present analysis establishes that observed degrees of vertical integration are not, from a social welfare standpoint, excessive. It should nevertheless be apparent that a broader a priori case for the vertical integration of production exists than is commonly acknowledged.

## REFERENCES

- Kenneth J. Arrow, *Aspects of the Theory of Risk-Bearing*, Helsinki, 1965.
- , "The Organization of Economic Activity: Issues Pertinent to the Choice of Market versus Nonmarket Allocation," in *The Analysis and Evaluation of Public Expenditures: The PPB System*, Vol. 1, Joint Economic Committee, Washington, D.C., 1969 pp. 47-64.
- Joe S. Bain, *Industrial Organization*, New York, 1968.
- Robert H. Bork, "Vertical Integration and Competitive Processes," in J. Fred Weston and Sam Peltzman, eds., *Public Policy Toward Mergers*, Pacific Palisades, Calif., 1969, pp. 139-49.
- Ronald H. Coase, "The Nature of the Firm," *Economica*, Nov. 1937, 4, 386-405; reprinted in George J. Stigler and Kenneth E. Boulding, eds., *Readings in Price Theory*, Homewood, Ill., 1952, pp. 331-51.
- , "The Problem of Social Cost," *Journal of Law and Economics*, 1960, 3, 1-44.
- Harold Demsetz, "Why Regulate Utilities?" *Journal of Law and Economics*, April 1968, 11, 55-66.
- William Fellner, "Prices and Wages under Bilateral Oligopoly," *Quarterly Journal of Economics*, August 1947, 61, 503-32.
- J. Hirshleifer, *Investment, Interest and Capital*, Englewood Cliffs, N.J., 1970.
- Frank H. Knight, *Risk, Uncertainty and Profit*, New York, 1965.
- Lionel McKenzie, "Ideal Output and the Interdependence of Firms," *Economic Journal*, December 1951, 61, 785-803.
- H. B. Malmgren, "Information, Expectations and the Theory of the Firm," *Quarterly Journal of Economics*, August 1961, 75, 399-421.
- James E. Meade, "External Economies and Diseconomies in a Competitive Situation," *Economic Journal*, March 1952, 62, 54-67.
- Roy Radner, "Problems in the Theory of Markets under Uncertainty," *American Economic Review*, May 1970, 60, 454-60.
- G. B. Richardson, *Information and Investment*, London, 1960.
- George J. Stigler, *The Organization of Industry*, Homewood, Ill., 1968.
- Andrew Whinston, "Price Guides in Decentralized Organizations," in W. W. Cooper, H. J. Leavitt, and M. W. Shelly, II, eds., *New Perspectives in Organization Research*, New York, 1964, pp. 405-48.
- Oliver E. Williamson, *Corporate Control and Business Behavior*, Englewood Cliffs, N.J., 1970.