DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration

Citation
Grossman, Sanford J., and Oliver D. Hart. 1986. The costs and benefits of ownership: A theory of vertical and lateral integration. Journal of Political Economy 94(4): 691-719.

Published Version
doi:10.1086/261404

Permanent link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:3450060

Terms of Use
This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story .

Accessibility

The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration

Sanford J. Grossman

Princeton University

Oliver D. Hart

Massachusetts Institute of Technology

Our theory of costly contracts emphasizes that contractual rights can be of two types: specific rights and residual rights. When it is costly to list all specific rights over assets in the contract, it may be optimal to let one party purchase all residual rights. Ownership is the purchase of these residual rights. When residual rights are purchased by one party, they are lost by a second party, and this inevitably creates distortions. Firm 1 purchases firm 2 when firm 1's control increases the productivity of its management more than the loss of control decreases the productivity of firm 2's management.

I. Introduction

A. *General Introduction*

What is a firm? What are the determinants of how vertically or laterally integrated the activities of the firm are? This paper builds on the foundations laid by Coase (1937), Klein, Crawford, and Alchian (1978), and Williamson (1979), which emphasize the benefits of "con-

Research was supported by the National Science Foundation, the Alfred P. Sloan Foundation, and the International Centre for Economics and Related Disciplines at the London School of Economics. This is an extensively revised version of an earlier paper (Grossman and Hart 1984). We gratefully acknowledge helpful comments from Debra Aron, Peter Diamond, Richard Epstein, Naava Grossman, Paul Joskow, John Minahan, Jim Poterba, Andrew Postlewaite, Peter Thistle, Martin Weitzman, Oliver Williamson, Charles Wilson, Toni Zabalza, and two referees.

[*Journal of Political Economy*, 1986, vol. 94, no. 4]
© 1986 by The University of Chicago. All rights reserved. 0022-3808/86/9404-0004$01.50

trol" in response to situations in which there are difficulties in writing or enforcing complete contracts (see also Williamson 1971, 1983; Williamson, Wachter, and Harris 1975; Teece 1980). We define the firm as being composed of the assets (e.g., machines, inventories) that it owns. We present a theory of costly contracts that emphasizes that contractual rights can be of two types: specific rights and residual rights. When it is too costly for one party to specify a long list of the particular rights it desires over another party's assets, it may be optimal for that party to purchase all the rights except those specifically mentioned in the contract. Ownership is the purchase of these residual rights of control. We show that there can be harmful effects associated with the wrong allocation of residual rights. In particular, a firm that purchases its supplier, thereby removing residual rights of control from the manager of the supplying company, can distort the manager's incentives sufficiently to make common ownership harmful. We develop a theory of integration based on the attempt of parties in writing a contract to allocate efficiently the residual rights of control between themselves.

We begin by reviewing some transactions cost–based arguments for integration. Coase (1937) suggested that transactions will be organized in the firm when the cost of doing this is lower than the cost of using the market. He added some content to this idea by proposing that the costs of constant recontracting with an outside firm or manager can be high relative to those of signing a long-term contract with an employee in which the employee agrees to carry out the commands of the employer. Klein et al. (1978) and Williamson (1979) added further content by arguing that a contractual relationship between a separately owned buyer and seller will be plagued by opportunistic and inefficient behavior in situations in which there are large amounts of surplus to be divided ex post and in which, because of the impossibility of writing a complete, contingent contract, the ex ante contract does not specify a clear division of this surplus. Such situations in turn are likely to arise when either the buyer or seller must make investments that have a smaller value in a use outside their own relationship than within the relationship (i.e., there exist "asset specificities").

While these statements help us understand when the costs of contracting between separately owned firms may be high, they do not elucidate what the benefits are of "organizing the transaction within the firm." In particular, given that it is difficult to write a complete contract between a buyer and seller and this creates room for opportunistic behavior, the transactions cost–based arguments for integration do not explain how the scope for such behavior changes when one of the self-interested owners becomes an equally self-interested employee of the other owner. Furthermore, if vertical integration

always reduces transaction costs, any buyer A and seller B that have a contractual relationship should be able to make themselves better off as follows: (i) A buys B and makes the previous owner of B the manager of a new subsidiary; (ii) A sets a transfer price between the subsidiary and itself equal to the contract price that existed when the firms were separate enterprises; and (iii) A gives the manager of B a compensation package equal to the profit of the subsidiary. Given this, however, how can integration ever be strictly worse than nonintegration; that is, what limits the size of the firm?[1]

A second question raised by the transactions cost–based arguments concerns the definition of integration itself. In particular, what does it mean for one firm to be more integrated than another? For example, is a firm that calls its retail force "employees" more integrated than one that calls its retail force "independent but exclusive sales agents"?

Existing theories cannot answer these questions because they do not give a sufficiently clear definition of integration for its costs and benefits to be assessed. It is not clear whether these theories are designed to explain the types of people called employees or instead the types of assets under the control of a single ownership unit. We define integration in terms of the ownership of assets and develop a model to explain when one firm will desire to acquire the assets of another firm. We will argue that, if one party gets rights of control, then this diminishes the rights of the other party to have control. To the extent that there are benefits of control, there will always be potential costs associated with removing control (i.e., ownership) from those who manage productive activities.

B. *What Is Integration?*

We define a firm to consist of those assets that it owns or over which it has control; we do not distinguish between ownership and control

[1] See Evans and Grossman (1983) for an elaboration of the critique of the transactions cost–based arguments for integration. Coase (1937) states that the size of the firm is limited by the managerial capacity of the single owner to manage many activities. As noted in the text, this is unconvincing since the owner could always hire another manager. The other authors do not give any clear statement as to what limits the size of the firm but appear to accept Coase's view that integration transforms a hostile supplier into a docile employee; thus the contracting problems associated with independent ownership are greatly diminished. However, there are some references to increased bureaucracy and its associated cost. See Williamson (1967), Rosen (1982), Keren and Levhari (1983), and Waldman (1984) for specific models of how the number of people involved in production affects the overall cost of production. None of these papers makes any distinction between the activities carried out via contract between separate owners and the activities carried out in a single ownership unit. That is, the theories are equally valid descriptions of how a firm can use hierarchies of outside contractors as they are theories of employment within the firm.

and virtually define ownership as the power to exercise control. In a corporation the shareholders as a group have control and delegate this control to the board of directors (i.e., management). Of course, control or ownership is never absolute. For example, a firm that owns a machine may not be able to sell it without the permission of the lenders for which the machine serves as collateral; more generally, a firm may give another firm specific authority over its machines. However, ownership gives the owner all rights to use the machine that he has not voluntarily given away or that the government or some other party has not taken by force. We believe that this terminology is roughly consistent with standard usage.[2]

In our attempt to explain asset ownership, we do not distinguish between employees and outside contractors in the case in which the firm provides all the tools and other assets used by the contractor. For example, in insurance retailing a firm may use its own employees as commissioned agents or use independent agents. The important difference between the two forms of retailing is that the employee-agent does not own the list of his clients, while the independent agent does own the list. If the firm owned the list and all the other important assets of the independent agents, then we would say that such a company had the same degree of integration as a company in which the retail sales force was composed of "employees." (A detailed discussion of the insurance industry may be found in Sec. IV.) As another example, consider vertical integration in shoe manufacturing. In the eighteenth century much of the manufacturing of shoes switched from the "putting out" system, in which the worker sewed the upper and lower halves of the shoe at home, to factory work, in which the factory owner's machines were used by the worker to put the shoes together (see Chandler 1977, p. 54). Even if workers are paid by the piece in both cases, the firm is more integrated in the latter case because it owns more of the machines used in production.

The examples above illustrate that the issue of ownership can be separated from the issue of contractual compensation. A firm may pay another firm or person by the piece or a fixed amount (salary), irrespective of the ownership of the machines. As Coase points out, the benefits of integration must surely be more than the ability to choose a new payment method. We assume that a payment method,

[2] Richard Posner, whose opinion on the legal definition of ownership we solicited, has referred us to the following statement by Oliver Wendell Holmes (1881/1946, p. 246): "But what are the rights of ownership? They are substantially the same as those incident to possession. Within the limits prescribed by policy, the owner is allowed to exercise his natural powers over the subject-matter uninterfered with, and is more or less protected in excluding other people from such interference. The owner is allowed to exclude all, and is accountable to no one but him."

whether it be salary compensation to an employee in the integrated company or a price for goods to be delivered between companies, is some function of the observable states of nature and the observable performance of the parties to the contract. We further assume that integration in itself does not make any new variable observable to both parties. Any audits that an employer can have done of his subsidiary are also feasible when the subsidiary is a separate company.[3]

It may be extremely costly to write a contract that specifies unambiguously the payments and actions of all parties in every observable state of nature. We assume that integration in itself does not change the cost of writing down a particular contractual provision.[4] What it does change is who has control over those provisions not included in the contract. Consider, for example, a contract between a publisher and a printer for a particular number of copies of a book. If the contract has no provision for an additional print run but the publisher receives some new information that makes it profitable for another run, then it is obvious that the right to decide whether or not to have the run belongs to the owner of the printing press. This is the simplest possible illustration of our assumption that the owner of an asset has the residual rights of control of that asset, that is, the right to control all aspects of the asset that have not been explicitly given away by contract.

C. *Introduction to the Model*

In order to be more specific about the costs and benefits of integration, it is necessary to set up a formal model of the relationship between two firms. This is done in Section II. For simplicity, the relationship, which may be either vertical or lateral, is assumed to last 2 periods.[5] In the first (i.e., the ex ante) period, the manager of each

[3] Arrow (1975) has analyzed the benefits of vertical integration based on the assumption that without integration it is more costly for one firm to communicate information to another than with integration. We do not see why any new method of communication becomes feasible under integration. The incentives of people to lie may change if their incentive structure changes, but Arrow does not explain how integration changes the set of feasible incentive structures. However, it might be the case that the right to audit is sometimes a residual right rather than a contractible right, in which case the theory developed below can explain the dependence of information on ownership patterns.

[4] Williamson (1983 pp. 523–24) gives an example of a contract written between nonintegrated firms in which there is no penalty for cancellation. He assumes that under vertical integration, or via the use of hostages, it is possible to extract a penalty from the buyer when he fails to take delivery of the seller's product. We shall ignore the possibility that there are artificial legal barriers to cancellation penalties and that integration is used by the parties as a way of getting around these.

[5] We model the relationship as a "once and for all" event. To the extent that the relationship is repeated, the incentives for vertical integration may be different from

firm makes relationship-specific investments, while in the second (i.e., the ex post) period, some further production decisions are taken and the benefits from the relationship are realized. A basic assumption of the model is that the production decisions, represented by **q**, are sufficiently complex that they cannot be specified completely in an initial contract between the firms. We have in mind a situation in which it is prohibitively difficult to think about and describe unambiguously in advance how all the potentially relevant aspects of the production allocation should be chosen as a function of the many states of the world. To simplify, we suppose that *no* aspect of **q** is ex ante contractible.[6] The noncontractibility of **q** creates the need to allocate residual rights of control since, if it is not specified how **q** will be chosen, there must be some implicit or explicit default that allows some party to choose the relevant components of **q** in the second period. We assume that the owner of each asset has the right to control that asset in the case of a missing provision.

Although **q** is ex ante noncontractible, we suppose that, once the state of the world is determined, the (small number of) relevant aspects of the production allocation become clear and the parties can negotiate or recontract over these (costlessly). That is, **q** is ex post contractible. Since the parties are assumed to have symmetric information, costless recontracting will always lead to an ex post efficient allocation, whatever is the initial allocation of ownership rights.[7] The *distribution* of ex post surplus, however, will be sensitive to ownership rights. For example, in the case of the printer and the publisher, while it may be efficient to have another print run, the printer will extract more surplus if he owns the printing plant and can therefore refuse to have the additional printing if negotiations fail.

Through their influence on the distribution of ex post surplus, ownership rights will affect ex ante investment decisions. That is, although ex post efficiency (relative to investment decisions) is guaranteed under any ownership structure, each ownership structure

---

those we give here. See Telser (1980) and Kreps (1984) for the role of reputation in long-term relationships as an enforcement device and Williamson (1979) for arguments on the role of repetitive idiosyncratic purchases in providing a cost to nonvertical integration. None of these papers deals with the influence of reputation on the ownership of assets. To the extent that reputation helps to enforce implicit agreements, repetition of the relationship is likely to increase the parties' surplus whether they are separate firms or part of the same firm. It is therefore unclear why reputation should have any particular implications concerning the ownership of assets.

[6] See Grossman and Hart (1984) for models in which some components of **q** are contractible while others are not.

[7] In a more complex model ex post inefficiencies will also appear in conjunction with costs of renegotiation. See Grossman and Hart (1984, sec. 2) for a model of ownership in which ex post inefficiencies rather than ex ante inefficiencies are analyzed.

will lead to a (different) distortion in ex ante investment. The ex ante investments that we are referring to are those that cannot be specified in the contract either because they are too complex to be described or because they stand for nonverifiable managerial effort decisions. We suppose that the parties allocate ownership rights in such a way that the ex ante investment distortions are minimized. The implications this has for the desirability of integration are the main focus of the paper and are analyzed in Section III.

It is worth asking why, in the context of our model, the usual argument that the feasible set can only become larger under integration fails. Given the existence of residual rights of control, if firm 1 buys firm 2, the owner of firm 1 will have the power to intervene in firm 2 in ways that may distort the incentives of firm 1's manager. Moreover, the owner cannot commit himself to intervene selectively in his subsidiary's operations since by their very definition residual rights refer to powers that cannot be specified in advance (at least in the detail required to make them part of an enforceable contract). It follows that integration can impose costs as well as benefits.

Since there are features of our theory that lack quantitative completeness, in Section IV we show how the theory can be applied to a particular industry, the insurance industry. Finally, Section V contains conclusions.

II. The Model

Consider two firms, 1 and 2, engaged in a relationship, which for simplicity we suppose lasts 2 periods. We assume that each firm is run by a manager who receives the full return from his firm's activities (the reason for this extreme assumption will become clear below). The firms sign a contract at date 0, and soon after managers 1 and 2 make relationship-specific investments, denoted by $\mathbf{a}_1$ and $\mathbf{a}_2$, respectively. At date 1, some further actions $\mathbf{q}_1$ and $\mathbf{q}_2$ are taken and the gains from trade are realized. We write the benefit of firm i's manager from the relationship at date 1, net of investment costs, as

$$B_i[\mathbf{a}_i, \phi_i(\mathbf{q}_1, \mathbf{q}_2)]. \tag{1}$$

All costs and benefits are measured in date 1 dollars. We will often interpret the relationship as a vertical one in which upstream firm 2 supplies downstream firm 1 with an input. In this case $B_2 < 0$ may be a cost. However, another interpretation is that the relationship is a lateral one, for example, between two retail stores with adjacent locations. For technical reasons, we have assumed that B_i depends on some function ϕ_i of $\mathbf{q}_1$ and $\mathbf{q}_2$ and is increasing in ϕ_i. We shall be

interested in cases in which there is a conflict of interest in the $\mathbf{q}$'s; for example, B_i might be increasing in $\mathbf{q}_i$ and decreasing in $\mathbf{q}_j$.[8]

The $\mathbf{q}_i$'s represent rights of control over firm i's assets, which are assumed to be ex ante noncontractible (as of date 0) but ex post contractible (as of date 1). As noted in the Introduction, we have in mind a situation in which it is extremely difficult to think about and describe in advance how the production allocation should depend on the "state of the world" but in which it is relatively easy to specify production decisions ex post once the state of the world is realized (a more detailed discussion of this may be found in n. 14). Since $\mathbf{q}_i$ is ex ante noncontractible, it qualifies as a residual right of control, and our assumption is that the owner of firm i has the right to choose it at date 1. Given that $\mathbf{q}_i$ is ex post contractible, however, firm i's owner may be prepared to give up this right in exchange for a side payment as part of renegotiation of the contract at date 1.

The ex ante investments $\mathbf{a}_i$ are also supposed to be noncontractible either because they are too complex to be described (they are multidimensional, not just dollar amounts) or because they stand for managerial effort decisions that are not verifiable (to third parties, such as the courts); for example, $\mathbf{a}_i$ might be manager i's effort in setting up a well-functioning firm. Investment decisions are assumed to be made independently and noncooperatively by the two managers just after the contract is signed at date 0. We shall suppose that each manager observes the other's investment decision after it has been made; in this model, there will be no asymmetries of information between the managers.

After investment decisions are made ex ante and ϕ_i is determined ex post through the choice of $\mathbf{q}_1$ and $\mathbf{q}_2$, manager i receives the benefit B_i. This benefit is again supposed to be nonverifiable and hence noncontractible. That is, B_i is a private benefit, accruing directly to firm i's manager, that does not show up in firm i's accounts. For example, B_i might stand for managerial perquisites or effort. A consequence of B_1 and B_2's not being verifiable is that it is impossible to write in the date 0 contract that firm 1, say, should transfer its benefit B_1 to firm 2.

We can summarize our assumptions so far as follows: (1) None of the variables $\mathbf{a}_i$, $\mathbf{q}_i$, and B_i is ex ante contractible, although the managers have symmetric information about these variables. Hence all the date 0 contract can do is to allocate ownership rights or residual rights of control to the two managers. (2) After the contract is signed, $\mathbf{a}_1$ and $\mathbf{a}_2$ are chosen simultaneously and noncooperatively by managers 1 and 2. (3) At date 1, the owner of firm i (i.e., the manager who has

[8] Here $\mathbf{a}_i$ and $\mathbf{q}_i$ are vectors in compact subsets of Euclidean spaces A_i and Q_i, respectively, and B_i and ϕ_i are continuous functions.

been given ownership rights in the date 0 contract) has the right and power to choose $\mathbf{q}_i$.[9] If there is no further negotiation, the choices of different owners are made simultaneously and noncooperatively. Given that the $\mathbf{q}$'s become contractible at date 1, however, the contract may be *renegotiated* (costlessly). Then B_1 and B_2 are realized.

It should be stressed that we assume that separate managers are needed to choose $\mathbf{a}_i$ and $\mathbf{a}_j$ under any ownership structure (but see Sec. III, remark 1).[10]

Finally, we assume that there is a competitive market in identical potential trading partners at date 0, which determines the ex ante division of the surplus between the two managers. Given this ex ante division, an optimal contract simply maximizes one manager's benefit subject to the other manager's receiving his reservation utility (note that there is no uncertainty). We make the standard assumption that the functions B_1 and B_2, as well as the domains of the variables $\mathbf{q}_i$ and $\mathbf{a}_i$, are common knowledge at date 0.

An example may be useful. Imagine that firm 1 is an electricity generating plant that is located next to a coal mine in order to use the mine's coal to make electricity (for a detailed analysis of long-term contracts between mine-mouth electricity generating plants and coal mines, see Joskow [1985]). Let $\phi_1(\mathbf{q}_1, \mathbf{q}_2)$ represent the quality of the coal delivered. Suppose that the boiler firm 1 installs to burn coal does not function well if the coal supplied is impure. Ex ante there may be many potential impurities, and it may be impossible to allow for each of these in the contract. Ex post, however, it may be clear what the relevant impurity is—high ash content, say. Our supposition is that, if firm 1 owns firm 2, it can, ex post, exercise its rights of control over firm 2's assets to direct that the coal should be taken from a deposit with low ash content (i.e., firm 1 chooses a subvector of $\mathbf{q}_2$). In contrast, if firm 2 owns firm 1, it can exercise its right of control over firm 1's assets to direct that the boiler should be modified to accept coal with high ash content.

An alternative to ownership in this example is a contract that gives firm 1, say, the specific right to direct the areas of the mine in which coal is dug out. This would clearly be reasonable for any one particular right of control. However, we have in mind a situation in which there are many aspects of a firm's operations, each of which may be important in a different contingency, and thus the costs of assigning

[9] We suppose that no special skills are required to choose $\mathbf{q}_i$. This means that the owner of firm i can contract with a subordinate to implement the choice of $\mathbf{q}_i$; moreover, since there are many subordinates available, none is in a position to refuse to carry out the owner's wishes or to argue about terms.

[10] The contrary assumption that integration is useful because it substitutes one manager for two has been advanced by Aron (1984) and Mann and Wissink (1984).

specific rights of control ex ante are much higher than the costs of assigning generalized control.

It may be useful if we comment briefly on the motivation for our assumption that $\mathbf{a}_i$, $\mathbf{q}_i$, and B_i are all ex ante noncontractible. We shall see in the next section that, if either the $\mathbf{a}_i$'s or the $\mathbf{q}_i$'s are ex ante contractible, the first-best can be achieved under any ownership structure, and so the degree of integration of the firms is irrelevant. The same is true if the B_i's are contractible since in this case the parties can always write a contract that transfers firm i's benefit to firm j, thus removing all conflicts of interest. Hence, in order to develop an interesting theory of ownership, it is necessary to assume that the $\mathbf{a}_i$'s, $\mathbf{q}_i$'s, and B_i's are all at least partly noncontractible. It is nonetheless very strong to assume that *no* aspects of these variables are contractible. In any realistic situation, some parts of a firm's performance will be reflected in verifiable shareholders' profit, even if other parts, such as managerial well-being, are not. Similarly, in a vertical relationship, while the parties may have difficulty in specifying the quality of input to be exchanged in advance, they can surely at least contract on the *quantity* of input. While we are confident that some version of our results will continue to hold when $\mathbf{a}_i$, $\mathbf{q}_i$, and B_i are partly contractible, the formal extension of our analysis to this case is by no means straightforward. As a first step, it therefore seems reasonable to study the case in which no date 1 variables are contractible at date 0.[11]

III. Analysis of the Optimal Contract, Including the Allocation of Ownership Rights

An optimal contract maximizes one manager's benefit subject to the other manager's receiving his reservation utility. Given that there is no uncertainty and that monetary transfers are available, it follows that an optimal contract must maximize the total ex ante net benefits or surplus of the two managers,

$$B_1[\mathbf{a}_1, \phi_1(\mathbf{q}_1, \mathbf{q}_2)] + B_2[\mathbf{a}_2, \phi_2(\mathbf{q}_1, \mathbf{q}_2)]. \tag{2}$$

[11] Elsewhere we have considered the effect of date 0 contractibles for the special case in which no revisions of the date 0 contract are permitted at date 1 (see Grossman and Hart 1984). In the present model, however, in which revisions are allowed, the introduction of contractibles complicates matters greatly. With a contractible, not only can the parties agree on a schedule relating the payment from firm i to firm j to the contractible, but they can also agree on a way of revising this price schedule at date 1 according to messages manager i and manager j send reflecting the choice of the sunk investments $\mathbf{a}_1$, $\mathbf{a}_2$ (for an analysis of this in a special case, see Hart and Moore [1985]). With no contractibles, the payment from firm i to firm j at date 1 is just a constant, and any attempt to make it sensitive to the environment will fail since price revisions are a zero-sum game from the point of view of the buyer and seller.

It is useful to consider as a benchmark the first-best, where contrary to our assumptions above $\mathbf{a}_1$ and $\mathbf{a}_2$ are verifiable and $\mathbf{q}_1$ and $\mathbf{q}_2$ are ex ante contractible.

DEFINITION. Let $\mathbf{a}_1^*$, $\mathbf{a}_2^*$, $\mathbf{q}_1^*$, and $\mathbf{q}_2^*$ be the (assumed unique) maximizers of $B_1 + B_2$ subject to $\mathbf{a}_i \in A_i$, $\mathbf{q}_i \in Q_i$ $(i = 1, 2)$.

The first-best contract would state that manager i must choose $\mathbf{a}_i^*$ at date 0 and $\mathbf{q}_i^*$ at date 1 (if not he must pay manager j a large penalty) and would specify a monetary transfer between the two managers.

In fact it is possible to achieve the first-best as long as the $\mathbf{q}_i$ are ex ante contractible, even if the $\mathbf{a}_i$ are not. For if the date 0 contract specifies that $\mathbf{q}_i = \mathbf{q}_i^*$, party i has an incentive to choose $\mathbf{a}_i$ to maximize $B_i[\mathbf{a}_i, \phi_i(\mathbf{q}_1^*, \mathbf{q}_2^*)]$, that is, to set $\mathbf{a}_i = \mathbf{a}_i^*$.[12] If neither the $\mathbf{q}_i$ nor the $\mathbf{a}_i$ are ex ante contractible, however, the first-best cannot generally be achieved, as we shall now see.

Under our simplifying assumption that no date 1 variables are contractible as of date 0, the contract will consist simply of an allocation of ownership rights and a transfer payment between the managers. There are three interesting cases to consider. In the first case, the firms remain separately owned or nonintegrated; that is, manager 1 owns and controls firm 1's assets and manager 2 owns and controls firm 2's assets. In the second case, firm 1 owns firm 2; that is, manager 1 owns and controls the assets of both firms (we call this firm 1 control). In the third case, firm 2 owns firm 1 (we call this firm 2 control). There is a fourth case in which manager 1 owns firm 2's assets and manager 2 owns firm 1's assets. This case appears less interesting than the others since it seems likely in practice to give a much lower level of surplus than case 1. We therefore ignore it in what follows.[13]

A. *Case 1: Nonintegration*

In this case manager 1 has the right to choose $\mathbf{q}_1$ and manager 2 has the right to choose $\mathbf{q}_2$ at date 1. It is useful to start at date 1 and work backward. At date 1, $\mathbf{a}_1$ and $\mathbf{a}_2$ are predetermined, and the only question concerns the choices of $\mathbf{q}_1$ and $\mathbf{q}_2$. If no further negotiation takes place, $\mathbf{q}_1$ and $\mathbf{q}_2$ will be chosen simultaneously and noncooperatively by managers 1 and 2 to maximize $\phi_1(\mathbf{q}_1, \mathbf{q}_2)$ and $\phi_2(\mathbf{q}_1, \mathbf{q}_2)$, respectively (since B_i is increasing in ϕ_i). We make the following assumption.

ASSUMPTION 1. There exists a unique pair $(\hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2)$ satisfying: $\mathbf{q}_1 =$

[12] This depends on our simplifying assumption that $\mathbf{a}_i$ does not affect B_j. The results presented below can be extended to the case of externalities in the $\mathbf{a}$'s without difficulty.

[13] There is also a class of more complicated contracts that make asset ownership at date 1 a function of messages the managers of firms i and j send after they have observed each other's investment decision. An example of this is an option to own contract. Our results are not affected by the existence of such contracts, and so for simplicity we ignore them.

$\hat{\mathbf{q}}_1$ maximizes $\phi_1(\mathbf{q}_1, \hat{\mathbf{q}}_2)$ subject to $\mathbf{q}_1 \in Q_1$, and $\mathbf{q}_2 = \hat{\mathbf{q}}_2$ maximizes $\phi_2(\hat{\mathbf{q}}_1, \mathbf{q}_2)$ subject to $\mathbf{q}_2 \in Q_2$.

In other words the game in which each manager i maximizes ϕ_i has a unique Nash equilibrium. Note the role of the separability assumption on B_i; it ensures that $\hat{\mathbf{q}}_1$ and $\hat{\mathbf{q}}_2$ are independent of $\mathbf{a}_1$ and $\mathbf{a}_2$.

Of course, given $\mathbf{a}_1$ and $\mathbf{a}_2$, the noncooperative equilibrium $(\hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2)$ is unlikely to be ex post efficient in the sense of maximizing

$$B_1[\mathbf{a}_1, \phi_1(\mathbf{q}_1, \mathbf{q}_2)] + B_2[\mathbf{a}_2, \phi_2(\mathbf{q}_1, \mathbf{q}_2)]. \tag{3}$$

Therefore, the two parties can gain from writing a new contract at date 1 that specifies that $\mathbf{q}_1 = \mathbf{q}_1(\mathbf{a}_1, \mathbf{a}_2)$, $\mathbf{q}_2 = \mathbf{q}_2(\mathbf{a}_1, \mathbf{a}_2)$, where these are the maximizers of (3) (if there are several maximizers, choose any pair). We will use the notation $\hat{\mathbf{q}} \equiv (\hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2)$ and $\mathbf{q}(\mathbf{a}) \equiv [\mathbf{q}_1(\mathbf{a}), \mathbf{q}_2(\mathbf{a})]$, where $\mathbf{a} = (\mathbf{a}_1, \mathbf{a}_2)$. The new contract is feasible since $\mathbf{q}_1$ and $\mathbf{q}_2$ are ex post contractible. It will specify a transfer price p that serves to allocate the gains from renegotiation. Because we do not want to get into the details of contract renegotiation, we shall simply assume that the parties split the increase in total surplus 50:50; that is, the transfer price p satisfies

$$\begin{aligned} B_1\{\mathbf{a}_1, \phi_1[\mathbf{q}(\mathbf{a})]\} - p = {} & B_1[\mathbf{a}_1, \phi_1(\hat{\mathbf{q}})] + \tfrac{1}{2}(B_1\{\mathbf{a}_1, \phi_1[\mathbf{q}(\mathbf{a})]\} \\ & + B_2\{\mathbf{a}_2, \phi_2[\mathbf{q}(\mathbf{a})]\} - B_1[\mathbf{a}_1, \phi_1(\hat{\mathbf{q}})] \\ & - B_2[\mathbf{a}_2, \phi_2(\hat{\mathbf{q}})]) \equiv \xi_1(\mathbf{a}, \hat{\mathbf{q}}), \end{aligned} \tag{4}$$

$$\begin{aligned} p + B_2\{\mathbf{a}_2, \phi_2[\mathbf{q}(\mathbf{a})]\} = {} & B_2[\mathbf{a}_2, \phi_2(\hat{\mathbf{q}})] + \tfrac{1}{2}(B_1\{\mathbf{a}_1, \phi_1[\mathbf{q}(\mathbf{a})]\} \\ & + B_2\{\mathbf{a}_2, \phi_2[\mathbf{q}(\mathbf{a})]\} - B_1[\mathbf{a}_1, \phi_1(\hat{\mathbf{q}})] \\ & - B_2[\mathbf{a}_2, \phi_2(\hat{\mathbf{q}})]) \equiv \xi_2(\mathbf{a}, \hat{\mathbf{q}}). \end{aligned} \tag{5}$$

This is in fact the Nash bargaining solution. Note that most bargaining solutions will yield an ex post Pareto-optimal outcome given our assumptions that the parties have the same information and that bargaining (i.e., contract renegotiation) is costless (see, e.g., Rubinstein 1982). It should be clear from what follows that our results will generalize to many other divisions of the surplus.

We assume that $\mathbf{a}_1$ and $\mathbf{a}_2$ are chosen noncooperatively by the agents at date 0, taking into account the renegotiation at date 1, that is, with regard to the overall payoffs ξ_1 and ξ_2. A Nash equilibrium in date 0 investments is a pair $(\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2) \in A_1 \times A_2$ such that

$$\xi_1(\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2, \hat{\mathbf{q}}) \geq \xi_1(\mathbf{a}_1, \tilde{\mathbf{a}}_2, \hat{\mathbf{q}}) \quad \text{for all } \mathbf{a}_1 \in A_1, \tag{6}$$

$$\xi_2(\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2, \hat{\mathbf{q}}) \geq \xi_2(\tilde{\mathbf{a}}_1, \mathbf{a}_2, \hat{\mathbf{q}}) \quad \text{for all } \mathbf{a}_2 \in A_2. \tag{7}$$

The total ex ante surplus from the relationship in this equilibrium is then

$$B_1\{\tilde{\mathbf{a}}_1, \phi_1[\mathbf{q}(\tilde{\mathbf{a}})]\} + B_2\{\tilde{\mathbf{a}}_2, \phi_2[\mathbf{q}(\tilde{\mathbf{a}})]\}. \tag{8}$$

A sufficient condition for the existence of a Nash equilibrium in date 0 investments is that A_i is convex and ξ_i is concave in $\mathbf{a}_i$ $(i = 1, 2)$.

We have seen how to compute total surplus in the case of nonintegration.[14] This will generally be less than the first-best level of surplus since the ex ante investments will be inefficient. To see this, note that the first-order conditions for a Nash equilibrium are

$$\frac{\partial \xi_i}{\partial \mathbf{a}_i} = \frac{1}{2}\frac{\partial B_i}{\partial \mathbf{a}_i}[\mathbf{a}_i, \phi_i(\hat{\mathbf{q}})] + \frac{1}{2}\frac{\partial B_i}{\partial \mathbf{a}_i}\{\mathbf{a}_i, \phi_i[\mathbf{q}(\mathbf{a})]\} = 0, \quad i = 1, 2, \tag{9}$$

where we are using the envelope theorem to eliminate remaining terms involving the ex post efficient $\mathbf{q}(\mathbf{a})$. This contrasts with the first-order conditions for the solution of (2),

$$\frac{\partial B_i}{\partial \mathbf{a}_i}\{\mathbf{a}_i, \phi_i[\mathbf{q}(\mathbf{a})]\} = 0, \quad i = 1, 2. \tag{10}$$

The inefficiency arises, then, because manager i puts 50 percent weight on the noncooperative outcome $\hat{\mathbf{q}}$, which is generally ex post inefficient, instead of all the weight on the cooperative outcome, which is ex post efficient; this is in spite of the fact that the noncooperative outcome never occurs! To the extent that the marginal and total benefits of $\mathbf{a}_i$ move in the same direction, the choice of the $\mathbf{a}_i$ can be substantially distorted. It is worth emphasizing that in this model all the inefficiency is due to the wrong choice of ex ante investment levels. The assumption of costless renegotiation ensures that there is no ex post inefficiency, and so if ex ante investments (more

[14] The reader may be concerned about our assumption that the manager can think clearly enough about $\mathbf{q}$ to solve (6) but that it is too costly to contract for $\mathbf{q}$ or design a mechanism to implement a particular $\mathbf{q}$. This assumption can be understood if we imagine that the noncontractible represents a special service that will be required of a firm at date 1 and that the type of service that is appropriate depends on the realization of a state of nature. Let there be N states of nature. The states are defined in such a way that state s requires the choice of activities from an M-dimensional space denoted by Q_s. The idea is that different activities are required for different states; i.e., while elements of Q_s, Q_t, $s \neq t$, are both M-dimensional Euclidean vectors, their coordinates refer to entirely distinct activities (different machines, e.g.). Further, in state s, the benefit function B is assumed to depend on the noncontractibles only through the chosen element $\mathbf{q}_s$ in Q_s, say $B = B(a, \mathbf{q}_s; s)$; if in state s some vector of activities in Q_t is chosen, $t \neq s$, no benefits are derived. Suppose in addition that we can normalize the spaces of activities so that $B(\mathbf{a}, \mathbf{q}_s, s) = B(\mathbf{a}, \mathbf{q})$, where $\mathbf{q}$ lies in a single space Q (where the coordinates of $\mathbf{q}$, of course, continue to refer to different activities in different states). Then, from an ex ante point of view, the manager, taking each s as equally likely, thinks of his objective as $B(\mathbf{a}, \mathbf{q})$, where $\mathbf{q}$ is a typical value assigned to the vector $\mathbf{q}_s$. Further, any element $\mathbf{q}$ in Q is contractible ex post (so that ownership has some value). However, to make $\mathbf{q}$ ex ante contractible, it would be necessary to specify different coordinates of $\mathbf{q}$ for each of the N states, and we assume that this is too costly.

precisely the noncontractible ones) are unimportant, the first-best can always be achieved.[15]

B. *Case 2: Firm 1 Control*

In this case firm 1 owns firm 2, and so manager 1 has the right to choose $\mathbf{q}_1$ and $\mathbf{q}_2$ at date 1. At date 1, manager 1 will now choose $(\mathbf{q}_1, \mathbf{q}_2)$ to maximize ϕ_1 if no further negotiation takes place. We make the following assumption.

ASSUMPTION 2. There is a unique pair $(\bar{\mathbf{q}}_1, \bar{\mathbf{q}}_2)$ such that $(\bar{\mathbf{q}}_1, \bar{\mathbf{q}}_2)$ solves: maximize $\phi_1(\mathbf{q}_1, \mathbf{q}_2)$ subject to $(\mathbf{q}_1, \mathbf{q}_2) \in Q_1 \times Q_2$.

The pair $(\bar{\mathbf{q}}_1, \bar{\mathbf{q}}_2)$ will generally not be ex post Pareto optimal, and so recontracting at date 1 will lead to the pair $\mathbf{q}_1(\mathbf{a}), \mathbf{q}_2(\mathbf{a})$, as in the case of nonintegration. *We will continue to assume that the parties split the gains from renegotiation 50:50.* That is, owning an additional firm increases a manager's bargaining power only by raising his status quo utility, that is, his utility in the event of no renegotiation (relative to given $\mathbf{a}_1$ and $\mathbf{a}_2$). Given the 50 percent sharing rule, manager i's final payoff is as in (4)–(5) with $(\bar{\mathbf{q}}_1, \bar{\mathbf{q}}_2)$ replacing $(\hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2)$. The date 0 Nash equilibrium in investments and the final level of surplus are also defined as in the case of nonintegration, again with $(\bar{\mathbf{q}}_1, \bar{\mathbf{q}}_2)$ replacing $(\hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2)$. Firm 1 control will generally lead to inefficient ex ante investments since $(\bar{\mathbf{q}}_1, \bar{\mathbf{q}}_2) \neq [\mathbf{q}_1(\mathbf{a}), \mathbf{q}_2(\mathbf{a})]$ (see [9]–[10]).

C. *Case 3: Firm 2 Control*

In this case, firm 2 owns firm 1, and so manager 2 has the right to choose $\mathbf{q}_1$ and $\mathbf{q}_2$ at date 1. Now, at date 1, manager 2 will choose $(\mathbf{q}_1, \mathbf{q}_2)$ to maximize ϕ_2 if no further negotiation takes place. We make the following assumption.

ASSUMPTION 3. There is a unique pair $(\underline{\mathbf{q}}_1, \underline{\mathbf{q}}_2)$ such that $(\underline{\mathbf{q}}_1, \underline{\mathbf{q}}_2)$ solves: maximize $\phi_2(\mathbf{q}_1, \mathbf{q}_2)$ subject to $(\mathbf{q}_1, \mathbf{q}_2) \in Q_1 \times Q_2$.

This case is the same as the previous one with $(\underline{\mathbf{q}}_1, \underline{\mathbf{q}}_2)$ replacing $(\bar{\mathbf{q}}_1, \bar{\mathbf{q}}_2)$ everywhere. Again ex ante investments will generally be inefficient.

We consider now which of the three cases above represents the optimal ownership structure. We saw in (9) that the inefficiency in the $\mathbf{a}$'s is due to the fact that manager i puts 50 percent weight on the noncooperative solution $(\tilde{\mathbf{q}}_1, \tilde{\mathbf{q}}_2)$—which equals $(\hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2)$ under non-

[15] The result that the conflict over the division of surplus at date 1 can lead to a distortion in investment at date 0 is similar to the finding of Grout (1984). In Grout's model, however, investment expenditure is observable, there are no noncontractibles, and the inefficiency in ex ante investment results from the assumed impossibility of writing binding contracts.

integration, $(\bar{\mathbf{q}}_1, \bar{\mathbf{q}}_2)$ under firm 1 control, and $(\underline{\mathbf{q}}_1, \underline{\mathbf{q}}_2)$ under firm 2 control—instead of 100 percent on the cooperative solution $[\mathbf{q}_1(\mathbf{a}), \mathbf{q}_2(\mathbf{a})]$. It is clear, therefore, that if one of the pairs $(\hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2)$, $(\bar{\mathbf{q}}_1, \bar{\mathbf{q}}_2)$, $(\underline{\mathbf{q}}_1, \underline{\mathbf{q}}_2)$ happens to be very close to $[\mathbf{q}_1(\mathbf{a}), \mathbf{q}_2(\mathbf{a})]$, there will be little inefficiency in the $\mathbf{a}$'s and the corresponding ownership structure will achieve approximately the first-best. Examples of this are provided in proposition 1.

PROPOSITION 1. (A) Suppose that ϕ_i depends primarily on $\mathbf{q}_i$ in the sense that $\phi_1(\mathbf{q}_1, \mathbf{q}_2) = \alpha_1(\mathbf{q}_1) + \epsilon_1\beta_1(\mathbf{q}_2)$, $\phi_2(\mathbf{q}_1, \mathbf{q}_2) = \alpha_2(\mathbf{q}_2) + \epsilon_2\beta_2(\mathbf{q}_1)$, where $\epsilon_1, \epsilon_2 > 0$ are small. Then nonintegration yields approximately the first-best, while firm 1 and firm 2 control generally do not. (B) Suppose that ϕ_2 hardly depends on $\mathbf{q}_1$ and $\mathbf{q}_2$ in the sense that $\phi_2(\mathbf{q}_1, \mathbf{q}_2) = \alpha_2 + \epsilon_2\delta_2(\mathbf{q}_1, \mathbf{q}_2)$, where $\epsilon_2 > 0$ is small. Then firm 1 control yields approximately the first-best, while nonintegration and firm 2 control generally do not. (C) Suppose that ϕ_1 hardly depends on $\mathbf{q}_1$ and $\mathbf{q}_2$ in the sense that $\phi_1(\mathbf{q}_1, \mathbf{q}_2) = \alpha_1 + \epsilon_1\delta_1(\mathbf{q}_1, \mathbf{q}_2)$, where $\epsilon_1 > 0$ is small. Then firm 2 control yields approximately the first-best, while nonintegration and firm 1 control generally do not.

To understand (and establish) part A, note that, under nonintegration, manager 1 chooses $\mathbf{q}_1 = \hat{\mathbf{q}}_1$ to maximize $\alpha_1(\mathbf{q}_1)$ and manager 2 chooses $\mathbf{q}_2 = \hat{\mathbf{q}}_2$ to maximize $\alpha_2(\mathbf{q}_2)$. If A holds, however, it is clear that in the limit $\epsilon_1 = \epsilon_2 = 0$ and $\hat{\mathbf{q}}_i$ is ex post efficient; that is, $[\mathbf{q}_1(\mathbf{a}), \mathbf{q}_2(\mathbf{a})] = (\hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2)$ for all $\mathbf{a}_1$ and $\mathbf{a}_2$. Hence (4)–(5) imply that in the limit $\mathbf{a}_i = \tilde{\mathbf{a}}_i$ maximizes $B_i[\mathbf{a}_i, \alpha_i(\hat{\mathbf{q}}_i)]$, and so $\mathbf{a}_1 = \tilde{\mathbf{a}}_1$ and $\mathbf{a}_2 = \tilde{\mathbf{a}}_2$ are ex ante efficient. Therefore, by continuity, for ϵ_1 and ϵ_2 small, nonintegration achieves approximately the first-best.

Firm 1 or firm 2 control, in contrast, may lead to great inefficiencies in case A. Under firm 1 control, in the absence of renegotiation, manager 1 chooses $\mathbf{q}_1 = \hat{\mathbf{q}}_1$ to maximize $\alpha_1(\mathbf{q}_1)$ (which is ex post efficient) and $\mathbf{q}_2 = \bar{\mathbf{q}}_2$ to maximize $\beta_1(\mathbf{q}_2)$ (which is ex post inefficient). This means that in the limit $\epsilon_1 = \epsilon_2 = 0$, $\mathbf{a}_1$ is chosen efficiently, but $\mathbf{a}_2$ is chosen to maximize

$$\tfrac{1}{2}B_2[\mathbf{a}_2, \alpha_2(\bar{\mathbf{q}}_2)] + \tfrac{1}{2}B_2[\mathbf{a}_2, \alpha_2(\hat{\mathbf{q}}_2)], \tag{11}$$

which may be very inefficient if $\bar{\mathbf{q}}_2$ is far from $\hat{\mathbf{q}}_2$. Similarly under firm 2 control, $\mathbf{a}_2$ is chosen efficiently while $\mathbf{a}_1$ is not.

Parts B and C follow similarly. Under B, firm 2 cares little about $\mathbf{q}_1$ and $\mathbf{q}_2$, and so if firm 1 has control over these, it will make an approximately ex post efficient choice. This will in turn lead to approximately ex ante efficient choices of $\mathbf{a}_1$ and $\mathbf{a}_2$. Under C, firm 2 control over $\mathbf{q}_1$ and $\mathbf{q}_2$ leads to approximately efficient ex post and ex ante outcomes.

Proposition 1 says that if the noncontractibles $\mathbf{q}_l$ ($l = 1$ or 2) have a small effect on firm j's benefit B_j, it is efficient for firm i to control them. The reason is that, if firm j controls them, j will use these rights

in such a way that i's ex ante expenditure is distorted, while if i owns them, there will be only a negligible distortion in j's investment (since j does not care about them). Note that j's ownership of $\mathbf{q}_l$ will lead to a serious distortion in i's expenditure only if $\partial^2 B_i[\mathbf{a}_i, \phi_i(\mathbf{q})]/\partial \mathbf{q}_l \partial \mathbf{a}_i$ is large, that is, if the marginal product of $\mathbf{a}_i$ is sensitive to $\mathbf{q}_l$. If $B_i[\mathbf{a}_i, \phi_i(\mathbf{q})] = f_i(\mathbf{a}_i) + \phi_i(\mathbf{q}_l)$, say, there is no distortion at all. To put it another way, proposition 1 tells us only that a particular ownership structure is optimal. It does not quantify the costs of being at a suboptimal structure. However, by choosing $(\partial/\partial \mathbf{q}_l)(\partial B_i/\partial \mathbf{a}_i)$ appropriately, we may easily construct examples in which this loss is extremely large.

REMARK 1. An interesting application of proposition 1 is to the special case in which one manager, manager 1, say, can run both firms by himself without any loss in efficiency (as in Aron [1984] and Mann and Wissink [1984]); that is, the firms may be engaged in complementary activities, and manager 1 may have some spare "capacity." This case can be captured by supposing that B_2 is approximately zero. We see from proposition 1 that under these conditions firm 1 control will dominate nonintegration or firm 2 control.

Proposition 1 deals with the special case in which the noncontractibles are important to one party but not to another. In general, both parties will care about the noncontractibles, and, as a result, each ownership structure will lead to a distortion in ex ante investments. The crucial question then is, Which ownership structure leads to the least significant distortion? Progress can be made in the analysis of this if we make some further assumptions. Recall that $B_i[\mathbf{a}_i, \phi_i(\mathbf{q}_1, \mathbf{q}_2)]$ is increasing in ϕ_i. We now make the following assumptions.

ASSUMPTION 4. Investment decisions are scalars and A_1, A_2 are intervals of the real line.

ASSUMPTION 5.

$$\frac{\partial^2 B_i}{\partial \phi_i \partial \mathbf{a}_i}[\mathbf{a}_i, \phi_i(\mathbf{q}_1, \mathbf{q}_2)] > 0.$$

That is, marginal benefit is high when average benefit is high.

ASSUMPTION 6.

$$\frac{\partial^2 B_i}{\partial \mathbf{a}_i^2}[\mathbf{a}_i, \phi_i(\mathbf{q}_1, \mathbf{q}_2)] < 0.$$

ASSUMPTION 7. The maximizers $\mathbf{q}_1(\mathbf{a})$ and $\mathbf{q}_2(\mathbf{a})$ of (3) are independent of $(\mathbf{a}_1, \mathbf{a}_2)$ in the relevant range; we write them as $\mathbf{q}_1^*$ and $\mathbf{q}_2^*$.

The fourth of these is a strong assumption. It says that the ex post efficient choice of the noncontractibles is independent of ex ante actions. The assumption is not reasonable if B_1 and B_2 are differentiable functions of the $\mathbf{q}$'s. However, it may hold if the $\mathbf{q}$'s take on only discrete values. In any case the argument that follows can be

generalized to the case in which assumption 7 is violated, at the cost of additional complexity.

The first-order conditions for the choice of ex ante investment by the managers are given by (9), where we replace $\phi_i(\hat{\mathbf{q}})$ by $\tilde{\phi}_i = \phi_i(\tilde{\mathbf{q}}_1, \tilde{\mathbf{q}}_2)$, the prerenegotiation outcome. Given assumption 7, (9) therefore becomes

$$\frac{1}{2}\frac{\partial B_i}{\partial \mathbf{a}_i}(\tilde{\mathbf{a}}_i, \tilde{\phi}_i) + \frac{1}{2}\frac{\partial B_i}{\partial \mathbf{a}_i}(\tilde{\mathbf{a}}_i, \phi_i^*) = 0, \tag{12}$$

where $\phi_i^* = \phi_i(\mathbf{q}_1^*, \mathbf{q}_2^*)$. On the other hand, the first-best investment decisions are characterized by

$$\frac{\partial B_i}{\partial \mathbf{a}_i}(\mathbf{a}_i^*, \phi_i^*) = 0. \tag{13}$$

Using assumption 5, we see that the left-hand side of (12) is positive (respectively negative) at $\mathbf{a}_i = \mathbf{a}_i^*$ if $\tilde{\phi}_i > \phi_i^*$ ($< \phi_i^*$). Hence, by assumption 6,

$$\tilde{\mathbf{a}}_i \lesseqgtr \mathbf{a}_i^* \quad \text{as} \quad \tilde{\phi}_i \lesseqgtr \phi_i^*. \tag{14}$$

Proposition 1 dealt with the case in which one of the ownership structures gave rise to a $(\tilde{\mathbf{q}}_1, \tilde{\mathbf{q}}_2)$ very close to $(\mathbf{q}_1^*, \mathbf{q}_2^*)$. Our concern now, however, is with cases in which $(\hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2)$, $(\bar{\mathbf{q}}_1, \bar{\mathbf{q}}_2)$, and $(\underline{\mathbf{q}}_1, \underline{\mathbf{q}}_2)$ are all quite "far" from $(\mathbf{q}_1^*, \mathbf{q}_2^*)$. We illustrate the situation in figure 1. The curve represents the efficient $\phi_2 - \phi_1$ combinations. We have drawn it to be continuous, but it could equally well be a set of discrete points. If firm 1 or firm 2 has control, the noncooperative outcome $(\tilde{\phi}_1, \tilde{\phi}_2)$ will lie on the efficiency frontier since one party controls $\mathbf{q}_1$ and $\mathbf{q}_2$. Under nonintegration, in contrast, the noncooperative outcome $(\hat{\phi}_1, \hat{\phi}_2)$ may well be highly inefficient because of the uncoordinated choice of $(\mathbf{q}_1, \mathbf{q}_2)$.

We can use figure 1 to determine the nature of the investment distortions corresponding to the different ownership structures. Since $\bar{\phi}_1 > \phi_1^*$ and $\bar{\phi}_2 < \phi_2^*$, (14) implies that $\bar{\mathbf{a}}_1 > \mathbf{a}_1^*$ and $\bar{\mathbf{a}}_2 < \mathbf{a}_2^*$; that is, under firm 1 control, firm 1 overinvests relative to the first-best and firm 2 underinvests. On the other hand, since $\underline{\phi}_1 < \phi_1^*$ and $\underline{\phi}_2 > \phi_2^*$, (14) implies that, under firm 2 control, firm 2 overinvests relative to the first-best and firm 1 underinvests. Nonintegration is more complicated since the nature of the distortion depends on the relationship of $(\hat{\phi}_1, \hat{\phi}_2)$ to (ϕ_1^*, ϕ_2^*). However, if the outcome $(\hat{\phi}_1, \hat{\phi}_2)$ is highly inefficient—which seems plausible in a number of cases—it will quite likely lie to the southwest of (ϕ_1^*, ϕ_2^*), that is, $\hat{\phi}_1 < \phi_1^*$ and $\hat{\phi}_2 < \phi_2^*$. Hence in this case $\hat{\mathbf{a}}_1 < \mathbf{a}_1^*$ and $\hat{\mathbf{a}}_2 < \mathbf{a}_2^*$; that is, nonintegration leads to underinvestment by both firms.

It may be useful to put these results in words. Under firm i control,

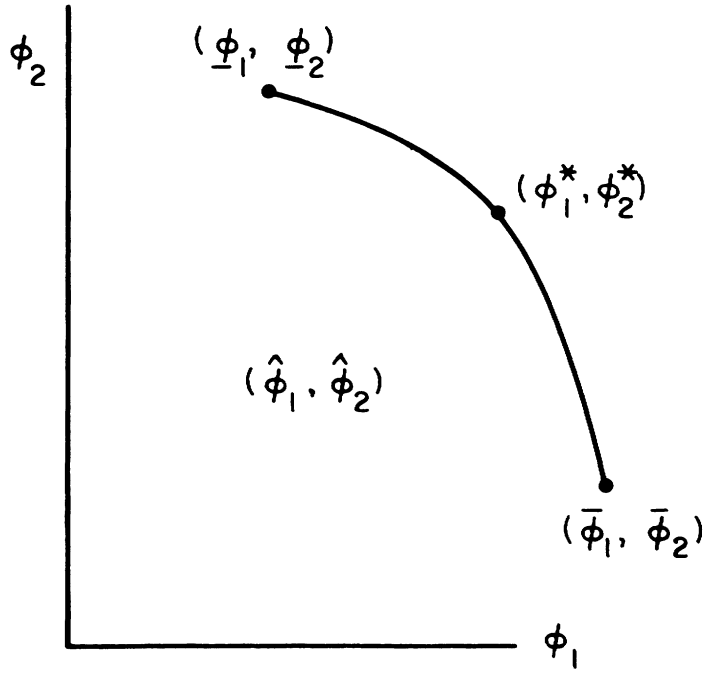

FIG. 1

firm i has a great deal of power ex post and hence will receive a high benefit in any date 1 renegotiation. Under assumption 5, however, high total benefits go together with high marginal benefits of investment, and so the consequence will be that firm i will overinvest. Firm j, on the other hand, with a low total and marginal benefit, will underinvest. Nonintegration, in contrast, gives both firms some power and will lead to moderate investment levels by each (note that, if $\hat{\phi}_2 > \bar{\phi}_2$ and $\hat{\phi}_1 > \underline{\phi}_1$, then $\hat{\mathbf{a}}_2 > \bar{\mathbf{a}}_2$ and $\hat{\mathbf{a}}_1 > \underline{\mathbf{a}}_1$; i.e., firm i's investment under nonintegration is greater than under firm j control).

The trade-offs should now be fairly clear. Firm 1 control will be desirable when firm 1's ex ante investment is much more important than firm 2's (so that firm 2's underinvestment under firm 1 control is relatively unimportant) and when overinvestment by firm 1 under firm 1 control is a less severe problem than underinvestment by firm 1 as in, for example, the nonintegrated solution. Firm 2 control will be desirable when firm 2's investment decision is much more important than firm 1's and when overinvestment by firm 2 is a less severe problem than underinvestment. Finally, nonintegration is desirable if $\mathbf{a}_1$ and $\mathbf{a}_2$ are both "important" in some sense, so that it is preferable to have both of them at a medium level than to have one very high and the other very low as under integration.

It is worth emphasizing that, because of the overinvestment problem, we cannot be sure that firm 1 control will be desirable even if firm 2's investment is irrelevant; that is, $\partial B_2/\partial \mathbf{a}_2 \equiv 0$. The overinvestment effect is a consequence of our assumption that the benefits B_j accrue to manager j and are inalienable from him. In a slight variant of our model, however, B_2, say, is perfectly alienable in the sense that there is a way for manager 1 to capture B_2 as long as he controls firm 2's assets. Under these conditions, the overinvestment effect disappears.

In particular, let $B_2 = f_2[\mathbf{a}_2, \phi_2(\mathbf{q}_1, \mathbf{q}_2)] - C_2(\mathbf{a}_2)$, where f_2 is a date 1 variable benefit and C_2 is a sunk investment cost. Imagine that, if firm 1 owns firm 2's assets, manager 1 has the option at date 1 of firing manager 2 and replacing him by another equally skilled manager at date 1 (training is unimportant), and this new manager can be offered a contract that pays him $-f_2$ (for simplicity, we suppose that the opportunity costs of both the old and new managers are zero). This means that if firm 1's manager has control, his benefit becomes $B_1 + f_2$; that is, the benefit f_2 is transferred. The fact that manager 2 will not receive f_2 will, of course, have a very adverse effect on his date 0 incentives. However, in the case in which manager 2's date 0 investments are unimportant, it is clear that firm 1 control will achieve the first-best since firm 1 will face the social objective function. We see then that the alienability of B_j and the irrelevance of $\mathbf{a}_j$ are jointly sufficient conditions for firm i control to be optimal.[16]

REMARK 2. One simplifying assumption we have made is that, when firm i owns firm j, it can control *all* the residual rights, $\mathbf{q}_j$. In reality, a subvector $\mathbf{q}_j^j$ of $\mathbf{q}_j$ may always remain under the control of manager j, say because manager j is the only person with the ability to control this particular aspect of the firm's operation. Our analysis can easily be generalized to this case. The main difference is that, even under integration, the prerenegotiation choice of $(\mathbf{q}_i, \mathbf{q}_j)$ will involve a lack of coordination by firms 1 and 2. Note that ownership rights are likely to be less important the more components of $\mathbf{q}_j$ remain under manager j's control. For example, suppose firm j is a law firm with a single lawyer and firm i is firm j's single client. Then if the client buys the law firm he may no more be able to get the lawyer to provide a special service than if the lawyer were in private practice. That is, the value of controlling firm j's assets may be very small in this case.

[16] In practice, the replacement of a manager may well be publicly observable and, hence, a contractible. So that we can stick with our framework in which there are no contractibles, we suppose that a replacement involves a move from one job to another in the company, which may not be verifiable (the manager may be "kicked upstairs" to a job with no perquisites, e.g.). That is, a replacement is part of the noncontractible $\mathbf{q}_2$.

IV. An Application

The main result of the last section can be stated as follows. If total and marginal benefits of investment move together, firm i ownership of firms i and j will lead to overinvestment by firm i and underinvestment by firm j. On the other hand, nonintegration will lead to moderate investment levels by each firm. The optimal ownership structure will be chosen to minimize the overall loss in surplus due to investment distortions. We now apply this result to the insurance industry.

Any real industry is, of course, far more complex than our model. One important difference is that in practice some variables will be contractible at date 0. We will therefore interpret our model with considerable latitude in what follows.

In the insurance industry some firms have a sales force that sells primarily its own company's products.[17] These companies are called direct writers, and their sales force may include employees (with virtually no ownership rights to office equipment) or agents who are independent contractors (who may own their office equipment and the building housing their agency office). Aside from the ownership of some office equipment, there are no major differences between employees and nonemployees; typically, both are on commissions, and the differences in commissions between the two types just reflect in an obvious way the differences in who bears office expenditures. However, in all cases direct writers are distinguished by the fact that the insurance company and not the agent owns the list of policyholders. Ownership of the list of policyholders entitles the insurance company to sell insurance to the policyholder if the agent terminates the relationship with the insurance company. Insurance company ownership of the list also means that the agent has no right to renew the insurance policy with a different company; he cannot leave the company and take his clients with him.

Insurance companies that are not direct writers sell insurance through independent agents and brokers (whom we will lump together as independent agents in distinction to the "captive" agents discussed above). The independent agents are distinguished by the fact that they, rather than the insurance company, own the list. An independent agent can sell any insurance company's product to his client. If the agent terminates his relationship with a particular insurance company, that company has no right to solicit the business from

[17] Our statements about the structure of the insurance industry are based on Strickler (1981), Webb et al. (1984), and conversations with professionals in the insurance industry. We are very grateful to Naava Grossman for her help in finding general information and data sources and for providing general information herself. We would also like to thank Peter Thistle.

the agent's list. Even without termination of the relationship, if the agent thinks that a client would be happier with the insurance of another company, the agent can encourage the client to change companies.

An insurance company has a number of expenditures that, given characteristics of the (contractible) commission structure to be explained below, can create ex post surplus between the insurance company and its agents or brokers. These expenditures include training of agents, client list-building expenditures (such as advertising), product development, and policyholder services. An insurance agent can have similar expenditures. To the extent that the efforts of the parties in generating these expenditures are not verifiable, they cannot be reimbursed directly without the creation of moral hazards. Instead the contract between the parties will specify payments as a function of observables, for example, commissions to the agent for policies produced for the insurance company.

We will use our framework to analyze the determinants of who owns the list of policyholders. (We assume that the agent does not want to own the whole insurance company.) Note that, since there is only one asset here (namely the client list), the choice is, in the language of our model, between firm 1 control and firm 2 control. Nonintegration has no meaning. To proceed, we must provide a model of the insurance industry. Space limitations permit only the simplest model. We assume that the agent devotes effort that is not verifiable to acquiring and keeping clients. The greater this effort, the more likely it is that a typical client will renew his insurance in the future, that is, that he will be persistent. Examples of such effort are the care with which the agent tailors the initial policy to the client's needs and the efficiency with which he deals with a claim once the policy is in force. Note that it is important for what follows that this effort yield dividends in the future, not just at the time when it is incurred; for example, a claim dealt with speedily today is likely to encourage the client to renew next year and the year after. To simplify the exposition, we assume that the agent can either "work" and produce only persistent clients or "not work" and produce only temporary clients, and that, if effort were verifiable, the insurance company would be prepared to compensate the agent for the extra effort of delivering persistent clients. An immediate implication of these assumptions is that, if the agent is paid a commission for the initial acquisition of the client and no later commission as a function of the persistence of the client, then the agent will deliver only temporary clients, and this is inefficient relative to the first-best. (Note that similar incentive problems will arise if some clients are naturally more persistent than others and the agent must devote extra effort to

finding the more persistent clients; the analysis below applies also to this case.)

In order to induce the agent to produce persistent clients, the commission structure must be back-loaded to reward the agent's initial effort costs. Specifically, the agent must get an initial commission somewhat lower than the acquisition cost of a client but get renewal commissions that are in excess of the agent's servicing costs associated with obtaining the renewal; that is, the renewal premium must have some component of a reward for the effort of delivering persistent clients.

The back-loading of commissions, in and of itself, has no particular implication for who owns the list, unless there are noncontractibles. We will be concerned with two kinds of noncontractibles that could interfere with the commission structure above: (1) noncontractibles that can hurt the agent if the company owns the list and (2) noncontractibles that can hurt the company if the agent owns the list.

Important examples of type 1 have to do with the fact that the insurance company can make the product it is selling less competitive (e.g., by raising its price or lowering the quality of its services relative to other insurance companies) and hence make the client more likely to want to switch insurance companies. For example, an insurance company can decide that it does not want to insure automobiles in a particular region, so it raises its prices or lowers the quality of its services in that region. Or the insurance company can change the type and quality of its advertising, which affects the likelihood that a client will renew his policy. It is very difficult for an insurance company to write a contract with agents that specifies all the relevant ways in which, and contingencies under which, the company will support the competitive position of its particular products; that is, these actions really are noncontractible. Such noncontractibles can seriously distort the agent's effort decision if the firm owns the list and the commission structure is back-loaded. In particular, once the commission structure is back-loaded, the agent will lose the renewal premium and thus be unable to recover his cost of delivering persistent clients when the company takes actions that lead the client to want to switch insurance companies. On the other hand, if the agent owns the list, then the back-loading of the commission structure does not distort the agent's action because the agent can switch the client to another company when the first company is a bad match for the client. In the notation of our model, this is a case in which the **q** of the insurance company is very important for the agent's ex ante effort.

There are also type 2 noncontractibles, that is, noncontractibles that can hurt the company if the agent owns the list. First, if the company develops an unanticipated new insurance product, then the agent's

clients cannot be solicited without the agent's permission when the agent owns the list. Second, when the agent owns the list, he can encourage his clients to switch to other companies if this seems advantageous (to him or to them). In some states of the world, such a switch may be efficient, but in other states it will merely increase the agent's profits at the expense of those of the company. The ability of the agent to switch customers in this way will distort the company's ex ante investments.[18]

So the trade-off between the different ownership structures is as in Section III. As in that section, we suppose that marginal benefits are small when average benefits are small. It follows that, if the company owns the list, the agent will have an insufficient incentive to deliver persistent clients; that is, he will underinvest in this activity. The company, on the other hand, will have at least the socially correct incentive to invest in list building and similar activities; that is, it will if anything overinvest in these activities. In contrast, if the agent owns the list, the company will underinvest in list building, but the agent will work hard to deliver persistent clients.

Further understanding about list ownership can be gained by considering what would happen if the reason for the back-loading of commissions disappeared. Recall that the back-loading was necessary because (*a*) the agent devoted nonverifiable effort to the servicing of clients and (*b*) the persistence of the client was sensitive to this effort. Much can be explained by noting that some kinds of insurance policies are more likely to be renewed than others, and this can make reason *b* much less of a factor. An example is "whole life" insurance. A life insurance policy will involve a longer-term contract than automobile insurance or fire and casualty insurance because a short-term policy gives very little protection to a person against the event that he will be sick but not die during the term of the life insurance policy and then be uninsurable thereafter. As a result, a life insurance customer has less of a tendency to switch insurance companies than does an automobile insurance customer. Moreover, to the extent that life insurance renewals do not occur, it is not because the agent has given the customer bad service on his claims! When renewals are relatively

[18] In each of the examples of noncontractibles we have indicated actions each party could take that would put the other party at a disadvantage. In some of these cases, e.g., the insurance company changing the support it provides to a given product, the noncontractible action does not involve direct manipulation of the item of which we are trying to explain the ownership, namely the client list. We have taken as given that the insurance agent does not want to own the whole insurance company. Hence the relevant variable that will allow the agent to increase his control over the renewal premiums to be generated by a particular client is the ownership of the list rather than direct control over the insurance company's marketing and product support program.

insensitive to the agent's actions, the commission structure need not be as back-loaded, and hence the argument for the agent to own the list is weakened. Further, even with some back-loading, to the extent that one company's **q**'s do not affect the desire of a client to switch given that his insurance is a long-term contract, the agent has less need to own the list (see proposition 1).

Our analysis therefore predicts that, in products in which the renewal is not guaranteed and is sensitive to the agent's actions, the agent will be more likely to own the list, whereas in products in which the renewal is more certain and is less sensitive to the agent's actions, the company will be more likely to own the list. We now argue that these predictions are consistent with facts characterizing the insurance industry.

One important fact is that about 65 percent of the premiums in property-casualty insurance are generated by agents who own the client list, while in life insurance about 12 percent of the premiums are generated by agents who own the list.[19] Most property-casualty insurance is sold for a shorter term than most life insurance. Table 1 gives a more detailed breakdown for life insurance. It can be seen that term insurance is sold far more often by agents who retain list ownership than is whole life insurance. Note that term life insurance is for a period of a few years and then must be renewed. Hence the renewal is more important for term than for whole life insurance.

Another important fact is that there is great variation regarding list ownership among products in the property-casualty product area. For example, independent agents have a 47 percent share of the market for private passenger automobile liability insurance, while they have a 96 percent share of surety insurance (see Webb et al. 1984, 1:85–88). Marvel (1982) has shown that there is a positive correlation across property-casualty products between the importance of independent agents (as measured by their market share) and the size of an agent's client acquisition costs (as measured by advertising and other acquisition expenses). We think that this is some support in favor of our conclusion that the agent will own the list when the agent's marginal incentives are relatively important in generating the renewal.[20] Table 1 is suggestive of a similar point for life insurance.

[19] The property-casualty number comes from Webb et al. (1984, 1:85); the life insurance number is from Life Insurance Marketing and Research Association (LIMRA) (1977, p. 9) and is the fraction of premiums written by insurance brokers (as opposed to captive agents) in 1977 for the United States. The LIMRA study also estimates that brokers tend to specialize somewhat in term policies rather than whole life policies.

[20] Marvel (1982) offers an alternative explanation for the correlation. He argues that there are situations in which it is more efficient for the company to advertise than for the agent. In these situations, the insurance company helps bring the client to the agent. According to Marvel, an agent who did not have an exclusive dealings contract

TABLE 1

LIFE INSURANCE

Product	Importance of Agent List Ownership*
Substandard insurance	55.9
Term insurance	46.2
Group/pension	43.1
Whole life insurance	19.4

SOURCE.—Czepiec (1984), table 1.

NOTE.—An agent who uses a brokerage insurance company as the insurance provider for his client is an agent who is the owner of his client list. The percentage figure refers to agents who claim to "frequently" place their clients with a brokerage insurance company. A brokerage insurance company is an insurance company that uses independent "agents" who are called brokers because they do not have a legal agency relationship with the company but instead represent the client.

* Measured as percentage of agents who use brokerage companies to sell the indicated product.

The selling of substandard insurance and group/pension insurance involves substantial effort on the part of the agent to find an insurer that is a good match for the client. The willingness of the client to

with the insurance company could then switch the customer to another insurance company that does not advertise and thus can pay higher commissions. This argument faces the following difficulty. First, if the company advertises the *specific* benefits of *its* product, why should the customer allow the agent to switch him to another insurance company? Marvel seems to be assuming that the agent uses a "bait and switch" sort of tactic against his customers. Second, if the insurance company convinces the customer about the *general* benefits of insurance, then how does an exclusive dealings contract protect the insurance company? The customer will just go to a cheaper company that advertises somewhat less (which he can find in the Yellow Pages rather than through television). Another piece of evidence that Marvel presents in favor of his argument is that exclusive dealings companies tend to spend more on advertising than do companies without exclusive dealings contracts. This correlation is consistent with our explanation as well. If, for any reason, a company is assured of more policies per customer it acquires, then it may spend more on acquisition costs. Further, Marvel does not explain the fact that life insurance tends to be sold through captive agents far more frequently than property-casualty insurance. He also does not distinguish ownership of the renewal from exclusive dealings. A company can own the renewal without having an exclusive dealings contract. Industry sources are emphatic in pointing out that "the most important characteristic of the independent agency system in comparison with the exclusive agency system is the independent agent's 'ownership of renewals or expirations' " (Strickler 1981, p. 294). *The renewal plays no role in Marvel's argument.* Nevertheless, Marvel's argument can be modified to supplement ours as follows. First, an exclusive dealings contract is one method of enforcing list ownership rights. Second, if, unlike Marvel, we assume that the company is advertising the specific high quality of its agent force (e.g., "your State Farm agent is always available"), then an exclusive dealings contract would be a method of recovering the expenditures from its agents. Note that an insurance company that is involved in selling for a longer time than any one agent or customer has reputational incentives to choose agents of high quality so that its advertising is to some extent truthful. We would then argue that companies will have exclusive dealings contracts when they are better able to convince customers of the agent's quality than is the agent or they are able to select agents of high quality.

maintain his insurance coverage with the agent depends on the quality of the match. Hence the ownership of the list by the agent provides him with more protection from the noncontractible acts of the insurance company than he would receive with company list ownership and the back-loading of the commissions.

V. Conclusions

When two parties enter into a relationship in which assets will be used to generate income, the parties can, in principle, contractually specify exactly who will have control over each dimension of each asset in each particular future contingency. We have argued that there is often a low-cost alternative to contracts that allocate all specific rights of control. In particular, when it is too costly for one party to specify a long list of the particular rights it desires over another party's assets, then it may be optimal for the first party to purchase all rights except those specifically mentioned in the contract. Ownership is the purchase of these residual rights of control. Vertical integration is the purchase of the assets of a supplier (or of a purchaser) for the purpose of acquiring the residual rights of control.

The literature on transactions costs has emphasized that incomplete contracts can cause a nonintegrated relationship to yield outcomes that are inferior to those that would be achieved with complete contracts. It is implicitly assumed that integration yields the outcome that would arise under complete contracts. We argue that the relevant comparison is not between the nonintegrated outcome and the complete contract outcome but instead between a contract that allocates residual rights to one party and a contract that allocates them to another. We have emphasized the symmetry of control—namely, that when residual rights are purchased by one party they are lost by a second party—and this inevitably creates distortions. That is, integration shifts the incentives for opportunistic and distortionary behavior, but it does not remove these incentives.

Our model emphasizes the distortions, due to contractual incompleteness, that can prevent a party from getting the ex post return required to compensate for his ex ante investment. To the extent that the marginal and average values of investment move together, the allocation of ownership rights, by changing the average investment return, will affect the level of investment. We have seen that, if firm i owns firm j, firm i will use its residual rights of control to obtain a large share of the ex post surplus, and this will cause firm i to overinvest and firm j to underinvest. Under nonintegration, on the other hand, the ex post surplus will be divided more evenly, and so each firm will invest to a moderate extent. Integration is therefore optimal

when one firm's investment decision is particularly important relative to the other firm's, whereas nonintegration is desirable when both investment decisions are "somewhat" important.

It should be noted, however, that contractual incompleteness can lead to other distortions. For instance, even if all ex ante investments can be verified and hence are reimbursable, residual rights may matter if the ex post distribution of the surplus is important for other reasons, for example, because of the risk aversion of the parties. An example is where manager 1 has an investment project but does not wish to finance it entirely himself since he would then bear all the risk. One possibility is to raise the funds externally from the market, which is risk neutral, say. The outside investors, who we suppose are led by manager 2, should then receive as their return a sizable fraction of the project's benefits. If manager 1 retains control of the project, however, he may be able to divert these benefits ex post from the investors to himself through his choice of noncontractibles, and knowing this, the investors may withhold some of their funds. In order to encourage outside investment, therefore, manager 1 may have to hand over some control to manager 2, for example, by giving him ownership rights over some of the assets.[21]

It should also be noted that if there is some barrier to ex post renegotiation, caused, for example, by the presence of transaction costs or asymmetric information, control of residual rights will be important in affecting the *size* of the ex post surplus as well as its *distribution* (even in the absence of ex ante investments). An analysis of the costs and benefits of ownership in this case may be found in Grossman and Hart (1984). A related idea is discussed by Farrell (1985).

Though we have emphasized residual rights of control over assets in order to explain who owns which assets, we can also use our theory to explain residual rights over actions. In particular, an employer-employee relationship differs from a contractor-contractee relationship in the allocation of residual rights of control over actions. An employer-employee relationship is typically characterized by the fact that many details of the job to be carried out are left to the employer's discretion; that is, the employer has many of the residual rights of control. In a contractor-contractee relationship, the job is specified in much greater detail, and the contractee typically has many of the residual rights of control over nonspecified actions. It may be useful in future work to apply our model to an analysis of the relative advantages of contractor-contractee and employer-employee relationships.

[21] We would like to thank John Minahan for a helpful discussion about this example.

It is worthwhile to consider which of the assumptions of the "Coase theorem" we drop in order to reach the conclusion that the distribution of ownership rights has efficiency consequences. The model of Sections II and III permits ex post bargaining of the type suggested in Coase (1960), but the ex ante efficiency of the relationship between the two parties will depend on how residual rights of control are allocated. The impossibility of ex ante bargaining over all aspects of the product to be delivered, that is, the incompleteness of the contract, is the source of our conclusion that the distribution of property rights has efficiency consequences.

References

Aron, Debra J. "Ability, Moral Hazard, and Firm Diversification, Part I." Mimeographed. Chicago: Univ. Chicago, Dept. Econ., May 1984.

Arrow, Kenneth J. "Vertical Integration and Communication." *Bell J. Econ.* 6 (Spring 1975): 173–83.

Chandler, Alfred D., Jr. *The Visible Hand.* Cambridge, Mass.: Harvard Univ. Press, 1977.

Coase, Ronald H. "The Nature of the Firm." *Economica* n.s. 4 (November 1937): 386–405. Reprinted in *Readings in Price Theory,* edited by George J. Stigler and Kenneth E. Boulding. Homewood, Ill.: Irwin (for American Econ. Assoc.), 1952.

———. "The Problem of Social Cost." *J. Law and Econ.* 3 (October 1960): 1–44.

Czepiec, Helena. "Measuring the Extent of Life Insurance Brokerage and Its Impact on Marketing Strategies." *Chartered Life Underwriters J.* (January 1984): 52–55.

Evans, David S., and Grossman, Sanford J. "Integration." In *Breaking Up Bell: Essays on Industrial Organization and Regulation,* edited by David S. Evans. New York: North-Holland, 1983.

Farrell, Joseph. "Allocating and Abrogating Rights: How Should Conflicts Be Resolved under Incomplete Information?" Mimeographed. Waltham, Mass.: GTE Labs, 1985.

Goldberg, Victor P., and Erickson, John E. "Long-Term Contracts for Petroleum Coke." Working Paper no. 206. Davis: Univ. California, Dept. Econ., September 1982.

Grossman, Sanford J., and Hart, Oliver D. "Vertical Integration and the Distribution of Property Rights." Mimeographed. Chicago: Univ. Chicago, Dept. Econ., 1984.

Grout, Paul A. "Investment and Wages in the Absence of Binding Contracts: A Nash Bargaining Approach." *Econometrica* 52 (March 1984): 449–60.

Hart, Oliver D., and Moore, J. "Incomplete Contracts and Renegotiation." Working Paper no. 367. Cambridge: Massachusetts Inst. Tech., 1985.

Holmes, Oliver Wendell. *The Common Law.* 1881. Reprint. Boston: Little, Brown, 1946.

Joskow, Paul. "Vertical Integration and Long-Term Contracts: The Case of Coal Burning Electric Generating Plants." *J. Law, Econ. and Organization* 1 (Fall 1985): 33–80.

Keren, Michael, and Levhari, David. "The Internal Organization of the Firm and the Shape of Average Costs." *Bell J. Econ.* 14 (Autumn 1983): 474–86.

Klein, Benjamin. "Transaction Cost Determinants of 'Unfair' Contractual Arrangements." *A.E.R. Papers and Proc.* 70 (May 1980): 356–62.

Klein, Benjamin; Crawford, Robert G.; and Alchian, Armen A. "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process." *J. Law and Econ.* 21 (October 1978): 297–326.

Kreps, David M. "Corporate Culture and Economic Theory." Mimeographed. Stanford, Calif.: Stanford Univ., 1984.

Life Insurance Marketing and Research Association. "Brokerage Ordinary Life Insurance Production by Ordinary Agencies and Branch Offices, United States and Canada in 1977." Research Report 1979-3. Farmington, Conn.: Life Insurance Marketing and Res. Assoc., 1977.

Mann, Duncan P., and Wissink, Jennifer P. "Inside vs. Outside Production: A Contracting Approach to Vertical Integration." Manuscript. Philadelphia: Univ. Pennsylvania, Center Study Organizational Innovation, 1984.

Marvel, Howard P. "Exclusive Dealing." *J. Law and Econ.* 25 (April 1982): 1–25.

Rosen, Sherwin. "Authority, Control, and the Distribution of Earnings." *Bell J. Econ.* 13 (Autumn 1982): 311–23.

Rubinstein, Ariel. "Perfect Equilibrium in a Bargaining Model." *Econometrica* 50 (January 1982): 97–109.

Strickler, Nancy E., ed. *Marketing Life and Health Insurance.* Atlanta: Life Office Management Assoc., 1981.

Teece, David J. "Economies of Scope and the Scope of the Enterprise." *J. Econ. Behavior and Organization* 1 (September 1980): 223–47.

Telser, Lester G. "A Theory of Self-enforcing Agreements." *J. Bus.* 53 (January 1980): 27–44.

Waldman, Michael. "Worker Allocation, Hierarchies and the Wage Distribution." *Rev. Econ. Studies* 51 (January 1984): 95–109.

Webb, Bernard L.; Laumie, J. J.; Rokes, W. P.; and Baglimi, N. A. *Insurance Company Operations.* 3d ed. Malvern, Pa.: American Inst. Property and Liability Underwriters, 1984.

Williamson, Oliver E. "Hierarchical Control and Optimum Firm Size." *J.P.E.* 75 (April 1967): 123–38.

———. "The Vertical Integration of Production: Market Failure Considerations." *A.E.R. Papers and Proc.* 61 (May 1971): 112–23.

———. "Transaction-Cost Economics: The Governance of Contractual Relations." *J. Law and Econ.* 22 (October 1979): 233–61.

———. "Credible Commitments: Using Hostages to Support Exchange." *A.E.R.* 73 (September 1983): 519–40.

Williamson, Oliver E.; Wachter, Michael L.; and Harris, Jeffrey E. "Understanding the Employment Relation: The Analysis of Idiosyncratic Exchange." *Bell J. Econ.* 6 (Spring 1975): 250–78.