Incomplete Contracts, Vertical Integration, and Supply Assurance

Patrick Bolton; Michael D. Whinston

Incomplete Contracts, Vertical Integration, and Supply Assurance

PATRICK BOLTON

London School of Economics and ECARE, Brussels

and

MICHAEL D. WHINSTON

Harvard University and NBER

This paper extends the analysis of transactions cost models of vertical integration to multilateral settings. Its main focus is on supply assurance concerns which arise when several downstream firms are competing for inputs in limited supply. Integration reduces supply assurance concerns for an integrating firm but it may increase them for others. Therefore, to explain the scope of any firm, one must consider the overall network of production and distribution relations. Three fundamental questions are addressed: (1) What are the effects of different integration structures?; (2) What are the determinants of the socially efficient integration structures?; (3) In what way do equilibrium integration structures differ from socially efficient structures?

1. INTRODUCTION

Recently, a great deal of attention has focused on theories of vertical integration in which the motives for integration arise from an inability to write comprehensive contracts (e.g., Williamson (1975, 1985), Klein, Crawford, and Alchian (1978), Grossman and Hart (1986), and Holmstrom and Tirole (1989). Notably, however, this literature has focused almost exclusively on *bilateral* trading relations, where a single supplier produces an input for use by a single buyer. Yet, this bilateral scenario is rarely descriptive of reality. In most settings, economies of scale and/or scope dictate that supply or purchasing relationships be *multilateral,* as when a manufacturer supplies inputs to a number of firms or a retailer handles numerous manufacturers' products.

When supply relationships are multilateral, however, new complications arise in determining the scope of the firm. For example, when a number of buyers rely on a single source of supply, shortages may make the input supplier only able to satisfy some subset of the buyers' needs, giving rise to the kind of supply assurance concerns that are often cited as a motive for vertical integration (e.g., Chandler (1964), Porter (1980)). However, while integration may reduce supply assurance concerns for an integrating firm, it may increase them for others. As a result, to explain the scope of any one firm, we must study the overall network of production and distribution relations.[1]

---

1. This point is, at best, recognized tangentially in the literature. Williamson (1985, Chapter 4), for instance, implicitly introduces aspects of a multilateral setting when he discusses a trade-off between transaction cost benefits and scale-economy losses of vertical integration. The scale economy losses he discusses arise because the integrated firm is presumed to be unable to sell externally, a problem that, as he notes (footnote 8, p. 92), must itself arise from some unmodeled transaction costs associated with the multilateral supply setting.

In this paper, we investigate these issues, extending the analysis of transactions cost/incomplete contracting models of vertical integration (most notably that of Grossman and Hart (1986)) to multilateral settings. Our focus is on settings where the sort of supply assurance concerns described above are present. The model we analyze is the simplest in which these problems can arise: a single upstream asset produces an input that is used by two downstream assets to produce their final products; with some probability, however, upstream capacity is insufficient to satisfy downstream needs. Further, to focus on the pure input assurance issue, we assume away any downstream competitive effects (the downstream goods may be different products or may be sold in different markets).

Our analysis focuses on three basic questions: First, what are the effects of different ownership structures? Second, what are the determinants of the socially optimal owner-ship structure? Third, how do private incentives differ from what is socially optimal? In addition, we highlight how moving from a bilateral setting to a multilateral one affects the answers to these questions.

The paper is organized as follows. We begin in Section 2 with a description of our model. Its basic assumptions are of three types: technological, contracting, and bargaining. Following Grossman and Hart (1986), we assume that ex ante contracts are necessarily incomplete. This incompleteness results in ex post bargaining over procurement of the input, a process we model in a manner similar to Rubinstein (1982). While the structure of ownership affects the division of surplus that arises in this bargaining, equilibrium allocations are always efficient here. Thus, as in Grossman and Hart (1986), the efficiency implications of ownership in our model come entirely through ownership's effect on ex ante investment.

In Section 3, we analyze the investment effects of the various possible ownership structures. Among our findings, we show that a move from non-integration (each asset being separately owned) to vertical integration (one downstream asset being owned jointly with the upstream asset) is typically accompanied by an increase in investment in the integrated downstream asset, a shift in supply patterns toward internal supply, and a decrease in the non-integrated downstream firm's payoff. Thus, as is often suggested in the informal literature on the subject, vertical integration in our model involves exter-nalities and a form of market foreclosure.

In Section 4, we investigate the welfare properties of these ownership structures. We first compare the investment levels arising in each structure to the levels that would be chosen by a planner. We show that while both downstream firms *under*-invest under non-integration, a move to vertical integration results in *over*-investment by the *integrated* firm. We then consider the determination of the socially optimal ownership structure and identify circumstances in which various ownership structures are best. We find that the multilateral supply setting has important implications for the determination of the optimal structure. For example, while integration would always be optimal in our model were there only a single downstream asset, in a multilateral setting non-integration is optimal in a broad range of cases. Indeed, we identify a class of cases in which non-integration achieves the first-best. Most interestingly, these cases arise precisely when *supply assurance concerns are greatest.*

This last finding is striking in light of the empirical literature on vertical integration, which often attributes firms' integration decisions to supply assurance concerns. In Section 5, we investigate the relation between the private and social incentives to integrate. In particular, we identify ownership structures that are stable in the sense that there is no mutually advantageous trade of assets by a coalition of individuals away from the structure. We find that, in our model, when the optimal structure is other than non-integration, it

is necessarily the unique stable structure. This need not be true when non-integration is optimal, however. Indeed, we show that in precisely the circumstances identified in Section 5 where non-integration is necessarily optimal, it is *never* stable. The fundamental cause of this problem lies in the negative externality that vertical integration imposes on the non-integrated downstream firm.

In Section 6 we discuss a number of extensions of our model, including the effects of alternative bargaining solutions, the introduction of managerial quasi-rents and upstream investment, the possibility of building of multiple upstream assets and chains of integration, and joint ownership.

Section 7 concludes.

In addition to the papers on incomplete contracts and vertical integration that we have already mentioned, our work here is related to several other analyses of vertical integration in the literature. One is the literature on supply assurance motives for vertical integration, most notably Carlton (1979) and Green (1986). In these papers, prices are assumed to be unable to respond to shifts in demand and the motive for integration arises from a desire to avoid input rationing. In addition, to keep all firms from integrating, these papers assume that there is some exogenous cost of integration (Green simply posits that integrated firms are less efficient, while Carlton assumes that an integrated firm cannot sell its inputs on the open market). In contrast, here prices respond fully to supply shortages through bargaining and both the costs of integration and the level of external sales by an integrated firm are endogenously derived.

The antitrust literature on foreclosure effects of vertical integration is also related to our work here. In addition to an extensive informal literature on the subject, three recent papers (Ordover, Salop, and Saloner (1990); Salinger (1988); Hart and Tirole (1990)) attempt to formally model this issue. In contrast to our model, downstream firms compete with one another in these papers and the resulting motives for integration differ from the input assurance concerns studied here. Nevertheless, our work shares the feature that integration can reduce the competitiveness of non-integrated firms.[2]

Lastly, in a recent paper, Hart and Moore (1990) consider a formal model of incomplete contracting and ownership in a multilateral setting that shares a number of features with our analysis. In their paper, however, investment is assumed to be *complementary* across individuals, in contrast to its basic substitutability in the supply assurance settings that we consider here. We also consider private incentives for integration, an issue not addressed in their work.

2. THE MODEL

There are three basic components to our model: technological assumptions, contracting assumptions, and bargaining assumptions. We present these in turn in subsections (i), (ii), and (iii).

(i) Technological Assumptions

We consider a simple vertical structure consisting of three productive assets: one upstream asset denoted by U and two downstream assets denoted by $D1$ and $D2$. The good produced by the upstream asset is an essential input for the production of the two

---

2. Bolton–Whinston (1991) contains an extensive discussion and comparison of Hart–Tirole's (1990) model 1 with our own. The expropriation of ex post surplus in Hart–Tirole's model 3 is also similar to some of the incentives observed in Section 6(iv)'s discussion of multiple upstream assets.

downstream goods. To keep things simple we suppose that each downstream plant requires one unit of U's input (in addition to other inputs) to produce one unit of output. The goods produced by these downstream assets do not compete with one another for consumer demand; for example, they may be different goods or could be the same good sold in different markets.

Each of the downstream assets has a single customer who desires at most one unit of its output. This customer's valuation for Di's product is randomly related to the level of investment in unobservable effort by asset Di's manager (we shall say more about managers below). We denote this valuation, net of Di's costs (other than the cost of U's input), by $v_i(I_i, s)$ for $i = 1, 2$ where $I_i \in [0, \bar{I}_i]$ is the effort level of Di's manager and $s \in S$ is a random state of nature; $v_i(I_i, s)$ is taken to be differentiable, non-decreasing, and concave in I_i. Note that in this formulation the investment of effort, I_i, could be either a cost-reduction activity or one designed to increase the quality of Di's product. The monetary equivalent cost of this effort for Di's manager is given by the function $c_i(I_i)$ where $c_i'(I_i) \geqq 0$ and $c_i''(I_i) > 0$. We let $\mu(s)$ denote the probability measure on the set of possible states of nature S. Finally, we assume that the final customer's net surplus, v_i, is observable and can be fully extracted by Di when it makes a sale.

To capture the idea of supply assurance concerns we assume that asset U has a random capacity level denoted by Q. With probability λ asset U can produce only one unit of output while with probability $(1 - \lambda)$ it can produce two. Thus, with probability λ a "shortage" will exist in which U is unable to supply the needs of the two downstream assets. For expositional simplicity we suppose that U's production costs per unit are zero and that $v_i(I_i, s) \geqq 0$ for all i, s, and $I_i \in [0, \bar{I}_i]$ (i.e., trade is always efficient); none of our results would change were these assumptions relaxed. Finally, we assume that there is no investment by the manager of U.[3]

We actually model final customers' net benefits from consumption as a flow of benefits occurring over an interval of time that we take to be of length one. At each instant that the final customer has Di's good during this interval he receives a flow net benefit of $v_i(I_i, s)$. Thus, if the consumer purchases the good at the beginning of this interval, the total net benefit is exactly $v_i(I_i, s)$, while if he does so after some fraction x of the interval has elapsed, his valuation is $(1 - x) \cdot v_i(I_i, s)$.[4] Below, the loss of these flow benefits will serve as the cost of delay in bargaining.

The basic technological timing is depicted in Figure 2.1. The contracting problem begins at time t_0. At time t_1, managers of the downstream assets choose their unobservable effort levels. At time t_2, the net values of the downstream firms and the upstream asset's capacity are realized (upstream production cannot take place before this time). These are assumed to be observable. Then, between times t_2 and $t_2 + 1$, production can occur and final customers can consume the products of the downstream firms.
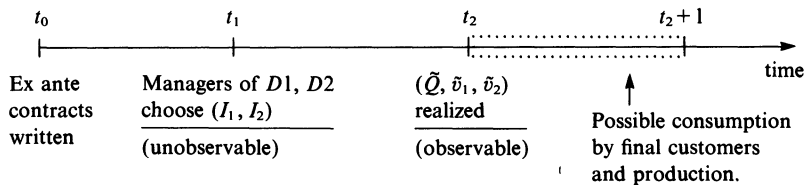

FIGURE 2.1

Managers are needed at two points in this sequence of events: at the investment stage and at the production stage. We assume that all individuals are equally capable managers. Thus, any owner of an asset can potentially manage his own asset. However, we assume that a single individual can manage at most one asset at a time. Thus, an owner of, say, two assets will have to hire a manager to run at least one of them. This assumption is an important one in the model for it prevents complete integration from always arising. Finally, we assume that managers develop no "quasi-rents" in the course of employment; that is, an individual who has managed an asset in the investment stage possesses no special expertise relative to other potential managers in the production stage. This strong assumption will serve to greatly limit (in fact completely) the ability of managers to earn returns from their unobservable effort; we discuss the effects of its relaxation in Section 6.

The model also affords two other interpretations. First, we could envision each D as actually having *many* final customers but assume that U's product is an essential *fixed* input; v_i is then Di's *aggregate* net value. Alternatively, rather than viewing Di as engaging in a single durable good sale to its customer, we could model Di as making a flow of non-durable sales over the interval $[t_2, t_2+1]$. In this flow model, at each instant prior to getting U's good Di can service his customer's demand at some lower net value (that could equal zero). In this case, v_i is the incremental flow value of U's input to Di and the value of getting the input to downstream firm Di is again proportional to the amount of the consumption period remaining.[5] In this latter interpretation, Di's problem is not so much getting input, as it is getting the input that it really needs. The relevant notion of capacity then may be U's capacity to provide some specialized service or design to Di.

(ii) Contracting

We assume that at time t_0 some initial ownership distribution of the three assets exists. We restrict attention to ownership structures involving an allocation of simple *titles* which give undivided ownership rights over each asset to a single individual. Thus, there are five possible ownership structures at time t_0, consisting of four basic types of asset integration (we use curly brackets to indicate jointly owned assets; hence $(\{U\}, \{D1, D2\})$ is a "horizontal integration" structure where one individual owns asset U and another owns both D's):

Number of owners	Ownership structures
3 separate owners	Non-integration (NI): $(\{U\}, \{D1\}, \{D2\})$
2 separate owners	Vertical integration $\begin{cases} VI_1: (\{U, D1\}, \{D2\}) \\ VI_2: (\{D1\}, \{U, D2\}) \end{cases}$ (VI) Horizontal integration (HI): $(\{U\}, \{D1, D2\})$
single owner	Complete Integration (CI): $(\{U, D1, D2\})$

Ownership gives the owner of an asset complete control over the use of that asset and a full claim to its financial return except for any restrictions he chooses to impose

---

5. For this situation to exactly correspond to our model of a single durable goods sale, the input must not be usable by one downstream asset once it has been used by the other (e.g., it may be embodied in the initial asset's production process so that transfer is too costly).

on himself through the contracts that he signs with others; that is, in the terminology of Grossman and Hart (1986), the owner possesses the *residual rights of control* over his asset.

It is not difficult to see that if complete contracts could be written, the first-best could be achieved in our model regardless of ownership structure.[6] In contrast, here we assume that there are significant limitations on the set of contracts that can be employed. These limitations open the door for the allocation of ownership to matter. In fact, following Grossman and Hart's (1986) treatment in the bilateral context (see also Holmstrom and Tirole (1989) and Hart and Moore (1990)), we make the strong assumption that the only possible contracts are simple spot transactions for inputs or managerial services consisting of a simple lump-sum payment in exchange for the good or service. Thus, after the realization of values and capacity at time t_2, owners bargain over the procurement of the input, signing simple spot contracts for its exchange. With vertical or complete integration, however, the owner of the integrated firm can unilaterally decide to transfer some of U's input to its downstream asset(s).

The procurement of managerial services occurs on a competitive market. At time t_1 owners simultaneously hire and assign managers for the investment stage, possibly deciding to manage one of their assets themselves; we assume that these decisions are not observable to other owners and managers.[7,8] For convenience we take this competitive wage to be zero.[9] After time t_2, the owner of that asset again hires a manager for production if necessary (or runs the asset himself). We assume that all managerial actions in this stage are (ex post) contractible. With this assumption and the absence of managerial quasi-rents, standard competitive conditions will prevail; we can then simply "black box" this transaction, viewing production as instantaneous and the v_i's as including this competitive managerial wage.

Given these assumptions, the timing of our model can now be depicted as in Figure 2.2.
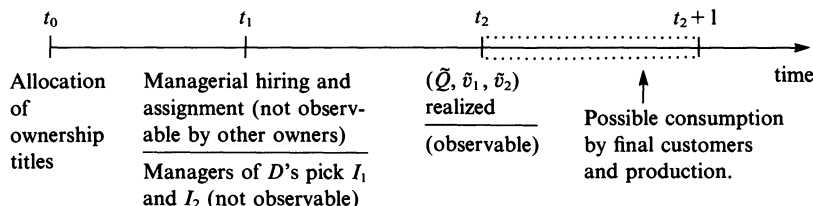

FIGURE 2.2

6. One contract that suffices has all owners plus any required managers (if there are fewer than three owners) agree to manage one of the three assets, to transfer the upstream unit's output in accordance with maximization of their aggregate payoff, and to give the individual managing each downstream asset the full payoff of his asset whenever he receives a unit of input. Payoffs to the three individuals are adjusted through ex ante lump-sum transfers.

7. The assumption of non-observability is not essential to any of our results. If these decisions were observable prior to the investment decisions, the hiring and assignment decisions could be used as a mechanism for strategically committing to a high or a low level of investment (depending on the choice), thereby complicating the analysis.

8. Note that we have implicitly ruled out the possibility that an owner of a downstream asset could hire another owner to be his manager. In our model, this type of managerial assignment will lead to an outcome different from that arising with a non-owner manager only if the individual hired owns U. We believe, however, that there are good (unmodeled) reasons for ruling out such managerial assignments. In circumstances where the input is not essential, for example, the manager of a downstream asset may be able to take grossly inefficient actions that serve only to raise the opportunity cost of not using U's input; while neither the owner of Di nor a non-owner manager will desire to take these actions, the owner of U will.

9. If the wage were positive ($w > 0$) each asset would have a fixed cost w that must be incurred at the investment stage if any later production is possible. In taking the wage to be zero, we are essentially assuming that this fixed cost does not result in the shut-down of any of the assets.

Our assumptions embody a number of important restrictions on the set of possible contracts. Like the existing literature on incomplete contracts and vertical integration in bilateral settings (e.g., Grossman and Hart (1986)), we take these contracting assumptions as a primitive rather than deriving them from more fundamental technological assumptions. Nevertheless, we feel it worthwhile to indicate some of our motivations in making these assumptions.

First, we do not allow any ex ante contracts specifying exchange of the input at or after time t_2. One reason why this restriction may be reasonable is that in many circumstances it may be difficult, if not impossible, to describe precisely the input characteristics required in the future even though these characteristics might be easily described ex post. If a contract for future delivery only vaguely describes the characteristics of the good to be delivered, the supplier may have strong incentives ex post to deliver an ineffective input. Then, despite the presence of the contract, the parties will essentially be left to bargain over the procurement of the input that is really needed. For instance, IBM may not be capable of describing what sort of microchip it wants Intel to deliver five years from now for a computer that it is currently developing, even though in due time it will be able to describe its needs precisely. In this case, IBM may not gain much from writing a long-term future delivery contract with Intel. In practice, of course, most situations lie between this extreme and that of complete contracting; we view our assumptions merely as a particularly stark way of capturing the important elements of contractual incompleteness present in many procurement settings.

A second restriction is that the owner of an asset cannot contract away any of his return stream on the margin. One example of a contract of this sort would be a contingent compensation package, say based on observed profits, for a manager hired during the investment stage. Another would be a profit-sharing agreement between owners of different assets. One reason why such contracts may be ineffective, however, is that the rights of control afforded by ownership (such as the right to decide the inputs that will be purchased) may inevitably provide an owner with the ability to undermine such schemes by altering observed returns; for example, he may purchase "inputs" that are really for his personal consumption or he may buy inputs from another firm that he owns at a very high price. Thus, marginal financial returns of an asset may inevitably be received primarily by its owner.[10] In the case of managerial compensation, for example, this point corresponds to the commonly cited idea (Williamson (1985)) that managers face "low-powered" incentives in contrast to the "high-powered" incentives faced by owners.[11]

We also do not allow contracts that specify so-called "announcement games" (as in Maskin (1977)) or that otherwise control the bargaining process that takes place between

---

10. Thus, the returns of the firms, though observable, may not be verifiable. See Hart (1988) for a formal derivation of this point in a simpler context than that here. Recent work by Holmstrom and Milgrom (1991) in the context of managerial compensation shows that it may be enough for certain important elements of returns to be non-contractable (either because they are simply unmeasurable or because of the manipulation problems discussed in the text) to undermine the desirability of contracting on the observable portion of returns. For example, even if current profits are observable, the continuation value of the asset may not be. If the owner makes compensation contingent on current performance, the manager may be led to inefficiently depreciate the asset in order to raise the current returns. Holmstrom and Tirole (1990) make a similar point and develop a bilateral model that closely parallels the assumptions we make here.

11. Another form of contingent contracting that might be used to (imperfectly) transfer financial returns on the margin are contracts that make payments conditional on whether trade occurs ex post. These may be difficult to write for the same reason that input exchange agreements may be problematic, namely, that it may be difficult to specify exactly what is meant by "trade".

the agents on the view that the enforcement costs of such contracts are too high to make their use viable.

Finally, in taking the allocation of titles to be the determinant of ownership rights, we do not allow for the rights of control over an asset to be split. A splitting of control might be done in two ways. First, forms of joint ownership might arise in which all control decisions must be agreed to in some pre-specified way (e.g., unanimous agreement) by the owners of an asset. We discuss such joint ownership arrangements in Section 6. Second, in some circumstances it may be possible to allocate control over various aspects of the firm to different individuals (e.g., individual A has control over market-decisions, while individual B has control over production decisions). As in the bilateral literature, we assume that such divisions are not practical; a similar style of analysis regarding the allocation of an asset's various control rights would be necessary if they were.

(iii) Ex post Bargaining

To model the ex post bargaining among owners over the input, we specify non-cooperative bargaining processes as in Rubinstein (1982) with the difference that, instead of modeling the costs of disagreement in terms of discounting, we model them in terms of foregone sales opportunities in the final output market. When an agreement to exchange U's input is delayed, the sale of the final good is also delayed and, as a result, the available net surplus is lower. Specifically, we suppose that there are T bargaining periods between t_2 and t_2+1 and let $\Delta = 1/T$. If U and Di reach agreement at time $t_2+\Delta$ instead of t_2, the value of the input at time $t_2+\Delta$ to Di is $(T-1)\cdot\Delta\cdot v_i(I_i, s)$ instead of $v_i(I_i, s)$ With this basic structure, we focus on the limit of the subgame perfect equilibria as $T \to \infty$.

Bargaining occurs in three different ownership structures in our multilateral setting: horizontal integration, vertical integration and non-integration. Under horizontal integration we have a situation of bilateral bargaining between the owner of the two downstream units and the owner of the upstream unit. We model this game like a standard alternating-offer Rubinstein (1982) bargaining game. The upstream owner and the downstream owner alternate in making offers from period to period, starting at time t_2. When one of them receives an offer at, say, time $t_2+\tau\Delta$, he can accept or reject the offer at that time. If any units are still available at time $t_2+\tau\Delta+\Delta$, that player can then make an offer. If there are q ($\leqq Q$) units still available at that time, the player making an offer can only make an offer for less than or equal to q units.

With ownership structure VI_i we again have a situation of bilateral bargaining, now between the owner of the integrated structure $\{U, Di\}$ and the owner of the non-integrated downstream firm Dj. When capacity $Q=2$, this bargaining problem is very much like that under horizontal integration: the owner of $\{U, Di\}$ will immediately use one unit of U's input in Di and will then bargain (with alternating offers) with the owner of Dj over the price of the second unit. When $Q=1$, however, matters are a bit different. In particular, when $v_j \geqq v_i$ (so that a sale to the owner of Dj is efficient), the nature of the bilateral bargaining problem is one where the owner of $\{U, Di\}$ has what the bargaining literature refers to as an "*outside option*" to simply use the unit in his own downstream asset (see, for example, Shaked and Sutton (1984), Rubinstein, and Wolinsky (1986), and Sutton (1986)).[12] More specifically, we suppose that the owner of $\{U, Di\}$ and the owner of $\{D_j\}$ alternate in making offers for the purchase or sale of one unit (recall that Dj has no use for more than one unit). In periods where the owner of $\{D_j\}$ makes offers, in addition

---

12. The standard terminology is a bit unfortunate: here *internal use* serves as an *outside option* for the integrated firm.

to responding by accepting or rejecting Dj's offer, the owner of $\{U, Di\}$ can unilaterally use U's input in asset Di.[13]

Under non-integration we have a three-player bargaining situation. We extend the alternating offer framework as follows: the upstream firm alternates in making offers with the two downstream firms. When it is the upstream firm's turn to make offers, it cannot offer to sell more units than it has available. Thus, if capacity $Q = 1$, it can offer either one unit to one of the downstream firms or can make no offers. When two units are available ($Q = 2$), the upstream firm can offer two units to one of the D's, one unit to each D, or no units to either D. When it is the upstream firm's turn to receive offers (if any units of input are still available), the two downstream firms simultaneously make offers for one unit. We restrict such offers to be no more than a downstream firm's value for the unit of input.[14] The upstream firm can accept or reject these offers subject only to the restriction that it not accept more offers than units it has available.

Given space limitations, we shall not derive the limiting equilibria of these bargaining models here. The derivations for HI and VI follow well-known lines; that for NI is available from the authors.

Under all of these ownership structures, the limiting bargaining outcome is efficient; that is, if there are two units of input, both D's receive a unit, while if there is only one unit it goes to the downstream asset with the highest valuation (this allocation is done unilaterally by the D's owner under structure HI). In addition, trade occurs, in the limit, immediately at time t_2. Thus, in our model (as in Grossman and Hart (1986)), any efficiency consequences of ownership must arise through effects on ex ante investment.

The payoffs that ar:-e from this bargaining under various ownership structures depend in part on whether, in the parlance of the bargaining literature, the outside options that owners may possess are "binding". If $v_i(I_i, s) \geq v_j(I_j, s)/2$ for all (I_1, I_2, s), these outside options are *always* binding. Through much of the paper (Sections 3–5), we focus on this case to illustrate the basic effects at work in the simplest possible setting. In this case, bargaining payoffs are as depicted in Table 2.1. In the Table, the payoffs under vertical integration are given for structure VI_1 only, $\vee$ and $\wedge$ are the max and min operators respectively, and a box depicts the payoff to the owner of jointly owned assets.

To understand the payoffs in Table 2.1, note that under HI the surplus from trade is split evenly between the bargainers. This is also true under vertical integration when there are two units available. When only one unit of input is available under structure VI_1 and $v_2 > v_1$, trade occurs and the option to use the input internally in $D1$ allows the owner of $\{U, D1\}$ to extract a payment of v_1 from $D2$'s owner.

Under non-integration, the payoff of the owner of $\{Dj\}$ for any given levels of v_i and v_j is *identical* to that arising under ownership structure VI_i (compare $D2$'s payoff under NI and VI_1 in Table 2.1). Intuitively, this arises because under non-integration, when $v_j > v_i$ and $Q = 1$, the possibility of selling to Di acts as U's "outside option" when bargaining with Dj. In particular, because Di anticipates that he will not receive the input, he is willing to pay all of v_i to purchase the good. This makes the value of U's

---

13. We follow the existing bargaining literature in this timing (see, for example, Sutton (1986)). The same results emerge if one assumes that acceptances/rejections involve a time delay of a period (so that rejections are made with a counteroffer) and allows the internal use option to be exercised at any instant by the integrated owner (when $Q = 1$, he will always exercise it immediately after hearing Dj's response if he does so at all).

14. This restriction rules out equilibria for the case where $Q = 1$ in which the low valuation firm bids more than its value and, by doing so, raises the price that is paid by the high value firm for the input (a similar problem arises in a simple Bertrand duopoly model with differing costs). In the absence of this simplifying assumption we would need to employ the stronger notion of trembling hand perfection (of the agent normal form) rather than just subgame perfection (Selten (1975)) to rule out such equilibria.

TABLE 2.1

Ownership Structure	Capacity	U	Bargaining Payoff to:	
			D1	D2
HI	$Q=2$	$\dfrac{v_1+v_2}{2}$	$\boxed{\dfrac{v_1+v_2}{2}}$	
	$Q=1$	$\dfrac{v_1 \vee v_2}{2}$	$\boxed{\dfrac{v_1 \vee v_2}{2}}$	
VI_1	$Q=2$	$\boxed{v_1+\dfrac{v_2}{2}}$		$\dfrac{v_2}{2}$
	$Q=1$	$\boxed{v_1}$		$(v_1 \vee v_2)-v_1$
NI	$Q=2$	$\dfrac{v_1+v_2}{2}$	$\dfrac{v_1}{2}$	$\dfrac{v_2}{2}$
	$Q=1$	$(v_1 \wedge v_2)$	$(v_1 \vee v_2)-v_2$	$(v_1 \vee v_2)-v_1$

outside option under NI equal to the value of internal use for the owner of $\{U, Di\}$ under ownership structure VI_i.

Note that this implies that for given realizations of values (v_1, v_2), Dj is indifferent between buying its inputs from an independent supplier and buying from one that is owned by the other downstream owner. Thus, in our model, vertical integration does not impose negative externalities on an unintegrated firm through bargaining effects.

If outside options do not always bind, the bargaining payoffs depicted in Table 2.1 must be amended because internal use of the input by the integrated firm under VI_i may be dominated by the option of continuing to bargain (as may a sale to the inefficient firm under NI). While we focus in Sections 3–5 on the case where outside options always bind, in Section 6 we discuss how our results are affected when this is not so (we discuss some alternative bargaining processes there as well).

3. INVESTMENT EFFECTS OF OWNERSHIP

We now study the consequences of ownership for ex ante investment in our model, restricting attention here to the case where outside options always bind. As we noted in the previous section, it is these investment effects that generate efficiency implications of ownership in our model.

Consider the four possible types of ownership structures:

Complete Integration. The case of complete integration is the simplest. The owner's ex post payoff from investments of (I_1, I_2) is the aggregate gross return of the assets. Letting $S_i(I_1, I_2) \equiv \{(I_1, I_2) \mid v_i(I_1, I_2) \geqq v_j(I_1, I_2)\}$ denote the set of states in which it is efficient for Di to get U's input when $Q=1$ and investments are I_1 and I_2, this is given by (the arguments of $S_i(\cdot, \cdot)$ are suppressed to ease notation):

$$\lambda\left\{\int_{S_1} v_1(I_1,s)\,d\mu(s)+\int_{S_2} v_2(I_2,s)\,d\mu(s)\right\}+(1-\lambda)\left\{\int_S [v_1(I_1,s)+v_2(I_2,s)]\,d\mu(s)\right\}.$$

$$(3.1)$$

Since any manager that is hired to run a downstream firm sets his effort equal to zero

(because he earns no ex post return on his investment of effort), the owner will clearly manage one of the D's, say Di, and pick I_i to maximize expression (3.1) less the effort cost $c_i(I_i)$, taking $I_j = 0$.

Horizontal Integration. Under HI the owner of $\{D1, D2\}$ will also manage one of his downstream assets. If he manages Di, then $I_j = 0$ and given the bargaining outcome described in Table 2.1, he will pick I_i to maximize, one-half times expression (3.1) less $c_i(I_i)$, taking $I_j = 0$.

Non-integration. Under NI, each owner of a downstream asset will manage his own asset (he can always pick investment to be zero and thereby replicate the outcome that would occur if he hired a manager). Consider the investment incentives of the owner of Di. Using the bargaining payoffs in Table 2.1, his expected ex post payoff given investments (I_1, I_2) is given by:

$$\lambda \int_{S_i} [v_i(I_i, s) - v_j(I_j, s)] d\mu(s) + (1 - \lambda) \int_S \frac{v_i(I_i, s)}{2} d\mu(s), \qquad (3.2)$$

a weighted average of his payoffs when $Q = 1$ and he gets the input (which happens for $s \in s_i(I_1, I_2)$) and when $Q = 2$. From (3.2), Di's optimal choice of I_i sets $c_i'(I_i)$ equal to:[15]

$$\lambda \int_{S_i} \left(\frac{\partial v_i(I_i, s)}{\partial I_i} \right) d\mu(s) + (1 - \lambda) \frac{1}{2} \int_S \left(\frac{\partial v_i(I_i, s)}{\partial I_i} \right) d\mu(s). \qquad (3.3)$$

To interpret these two terms, note that when capacity $Q = 1$ (the first term), Di receives the full marginal increase in value in states in which he receives the input ($s \in S_i$) and he receives nothing otherwise while, when $Q = 2$, Di always receives the good but loses half of any increase in value to the owner of U through bargaining.

A useful way to view these effects graphically comes from considering the special case where $S = [0, 1]$ and $\mu(s)$ is a uniform distribution on this interval. This allows us to graph values $v_i(I_i, s)$ for the various possible realizations of s and to measure expected values as the areas under these curves. For example, in Figure 3.1(a), $v_1(I_1, s)$ is drawn for two possible values of $D1$'s investment, I_1 and $I_1' > I_1$. The expected value of $v_1(I_1, s)$ is the shaded area (A) in the figure; cross-hatched area (B) is the change in its expected value when I_1 is increased to I_1'. Similarly, Figure 3.1(b) depicts the same two value curves for $v_1(\cdot, \cdot)$ along with a single value curve for $v_2(\cdot, \cdot)$; for convenience, the drawing assumes that values for the two firms are perfectly negatively correlated. Di's expected payoff when $Q = 1$ and investments are (I_1, I_2) is the shaded area (C) in Figure 3.1(b). Discrete approximations of the two marginal payoff terms in expression (3.3) can be seen in Figures 3.1(a) and (b). The first term (for $Q = 1$) equals the cross-hatched area (D) in Figure 3.1(b); the second (for $Q = 2$) corresponds to one-half of area (B) in Figure 3.1(a).

If we let $I_i^*(I_j \mid NI)$ denote Di's best-response under non-integration to investment level I_j by Dj, $I_i^*(I_j \mid NI)$ must be non-increasing in I_j (if it is multivalued then, if $I_j' > I_j$, all elements of $I_i^*(I_j' \mid NI)$ are weakly less than min $\{I_i^*(I_j \mid NI)\}$). To see this, note that I_j in expression (3.3) enters only through the determination of the set $S_i(I_1, I_2)$. Since set $S_i(I_1, I_2)$ (weakly) contracts when I_j increases, Di's marginal investment incentives must (weakly) fall when I_j increases. Put less formally, the reason for this effect is straightforward: Di's marginal return to investment is increasing in his "market share"— the set of states in which he receives the input when $Q = 1$—and increases in I_j reduce this share.

---

15. To derive this expression, note that the induced change in set S_i has no first-order effect on profits (holding the level of I_j fixed).
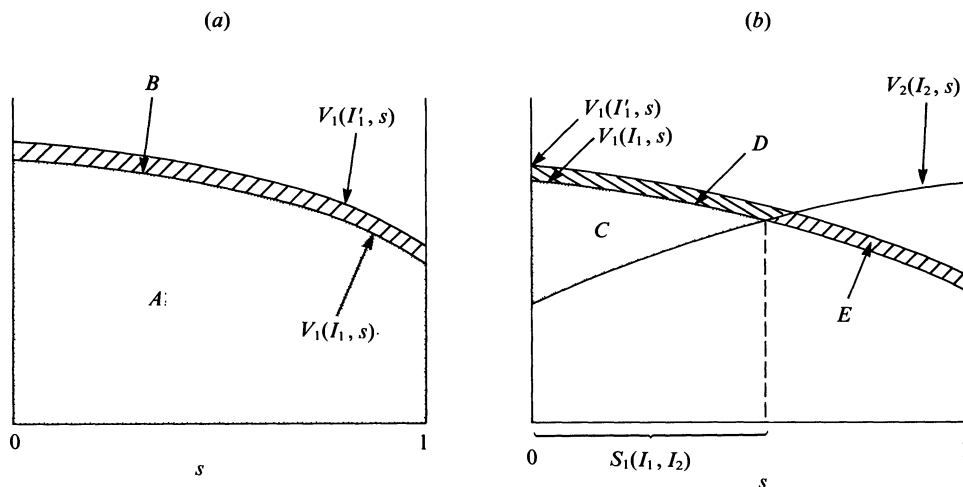
FIGURE 3.1

Vertical Integration. Under ownership structure VI_j, both owners again manage their downstream assets. From Table 2.1 and the discussion accompanying it, we know that the non-integrated firm Di's expected payoff under VI_j is exactly the same as under non-integration for any (I_1, I_2); thus, its investment incentives given I_j must be exactly those given above in (3.3). Now consider those of the owner of $\{U, D_j\}$. From Table 2.1, his expected ex post payoff is given by:

$$\lambda \int_S v_j(I_j, s) d\mu(s) + (1-\lambda) \int_S \left(v_j(I_j, s) + \frac{v_i(I_i, s)}{2} \right) d\mu(s). \qquad (3.4)$$

His optimal effort level therefore sets $c_i'(I_i)$ equal to:

$$\lambda \int_S \left(\frac{\partial v_j(I_j, s)}{\partial I_j} \right) d\mu(s) + (1-\lambda) \int_S \left(\frac{\partial v_j(I_j, s)}{\partial I_j} \right) d\mu(s). \qquad (3.5)$$

Both when $Q = 1$ and when $Q = 2$, the owner of $\{U, Dj\}$ receives the full marginal increase in his value in *every* state s. This is clear when $Q = 2$ since he uses a unit of input in asset D_j in all states s. When $Q = 1$, however, this result arises because the option to use the input internally gives the owner of $\{U, Dj\}$ a payoff of v_j *whether or not he actually ends up using the input himself.* These marginal effects can also be seen in Figure 3.1: the marginal effect when $Q = 2$ is now equal to area (B) in Figure 3.1(a), while that when $Q = 1$ is now equal to the sum of area (D) plus area (E) in Figure 3.1(b).

Note from (3.5) that Dj's optimal investment level $I_j^*(I_i | VI_j)$, which is unique under our concavity assumptions on $c_j(\cdot)$ and $v_j(\cdot, s)$, is independent of I_i. Thus, if Di's best-response correspondence is single-valued, there are unique equilibrium investment levels in structure $VI_j, [I_1^*(VI_j), I_2^*(VI_j)]$.

The comparison of the outcomes under non-integration and vertical integration, summarized in the following result, is of considerable interest:

Proposition 3.1. *If $[I_1^*(NI), I_2^*(NI)]$ and $[I_1^*(VI_j), I_2^*(VI_j)]$ are equilibrium levels of investment under NI and VI_j respectively, then,*

  (1) *$Dj's$ equilibrium investment level is (weakly) higher under structure VI_j than in this non-integration equilibrium while $Di's$ equilibrium investment level is (weakly) lower. That is, $I_j^*(VI_j) \geq I_j^*(NI)$ and $I_i^*(VI_j) \leq I_i^*(NI)$.*

(2) *Relative to supply patterns in this non-integration equilibrium, a move to vertical integration by Dj results in a (weak) shift toward internal supply. That is,* $S_j(I_1^*(NI), I_2^*(NI)) \subseteq S_j(I_1^*(VI_j), I_2^*(VI_j))$.

(3) *Vertical integration by Dj gives the owner of Di a (weakly) lower payoff than under non-integration.*

Moreover, if $\partial v_i(I_i, s)/\partial I_i > 0$ *and* $v_i(I_i, s)$ *is continuous in s for* $i = 1$ *and* 2 *and if* $\phi \neq S_j(I_1^*(NI), I_2^*(NI)) \subset S$, *then all of these inequalities are strict.*

Remark. Proposition 3.1 does not require a unique equilibrium under either *NI* or *VI$_j$*; it applies for any equilibria arising in these structures.

Proof. We argue only the case of weak inequalities here; the argument for strict inequalities follows similarly. Since $S_j(I_1, I_2) \subseteq S$ for all (I_1, I_2), (3.3) and (3.5) imply that $I_j^*(VI_j) \geq \max\{I_j^*(0 | NI)\}$. Because $I_j^*(I_i | NI)$ is non-increasing, $\max\{I_j^*(0 | NI)\} \geq \max\{I_j^*(I_i^*(NI) | NI)\} \geq I_j^*(NI)$, so we have $I_j^*(VI_j) \geq I_j^*(NI)$. Also, because $I_i^*(I_j | VI_j) = I_i^*(I_j | NI)$ is non-increasing in I_j, $I_i^*(VI_j) \leq I_i^*(NI)$. This establishes claim (1). Claim (2) follows from the fact that $S_j(I_1, I_2) \subseteq S_j(I_1', I_2')$ if $I_j' \geq I_j$ and $I_i' \leq I_i$. Finally, since Di's payoff is weakly decreasing in I_j holding I_i fixed, the fact that $I_j^*(VI_j) \geq I_j^*(NI)$ establishes claim (3). ‖

Claims (2) and (3) of Proposition 3.1 are of particular interest. A move to vertical integration causes a shift in supply patterns when $Q = 1$ toward the integrated downstream asset and imposes a negative externality on the downstream owner who remains unintegrated. Note, though, that these effects do not result from inefficient ex post trade, but rather, from the changes in investment that integration causes (for any *given* levels of I_1 and I_2, the payoff of a non-integrated downstream firm is unaffected by whether he buys from an independent upstream firm or one owned by the other downstream firm). A typical scenario is depicted in Figure 3.2 where the best-response functions under *NI* and *VI$_1$* are drawn (the dashed curves in the figure will be described in Section 4). There, vertical integration by *D*1 shifts the best-response function for I_1 rightward to the vertical line labeled $I_1^*(I_2 | VI_1)$. The effects of this shift are as described in Proposition 3.1.
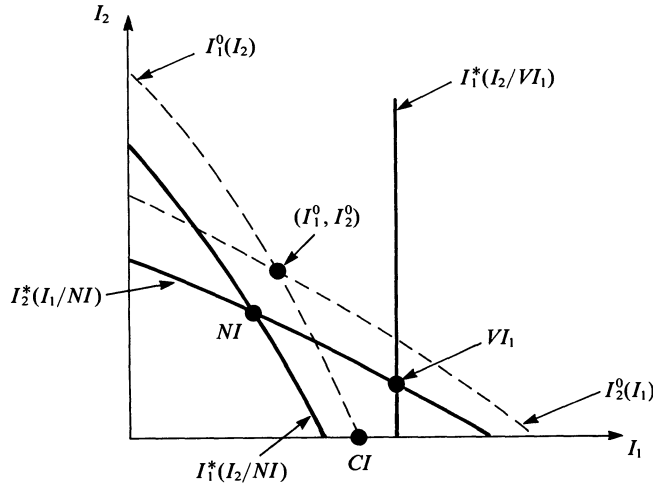

FIGURE 3.2

4. OPTIMAL OWNERSHIP STRUCTURES

We now investigate the welfare properties of these ownership structures. We begin in sub-section (i) by comparing the investment incentives under the various structures to socially optimal incentives. Then, in sub-section (ii), we investigate the determination of the optimal ownership structure.

(i) Comparisons to the First-best

In a first-best world where investment and trade could be directly controlled, a social planner would pick investment levels (I_1, I_2) to maximize expression (3.1) less investment costs $[c_1(I_1) + c_2(I_2)]$.[16] The marginal change in social welfare, gross of investment costs, for a differential change in I_i is then given by,

$$\lambda \left\{ \int_{S_i} \frac{\partial v_i(I_i, s)}{\partial I_i} d\mu(s) \right\} + (1-\lambda) \left\{ \int_S \frac{\partial v_i(I_i, s)}{\partial I_i} d\mu(s) \right\}. \tag{4.1}$$

When $Q = 2$, the social planner considers the full marginal increase in v_i for all states $s \in S$, while for the case where $Q = 1$, the planner considers the full marginal increase in v_i for those states in which Di will get the input $(s \in S_i)$. In terms of Figure 3.1, the marginal effect when $Q = 2$ is given by area (B) in Figure 3.1(a), while that when $Q = 2$ is given by area (D) in Figure 3.1(b).

Note that if we let $I_i^o(I_j)$ denote the planner's optimal choice of I_i for a given level of I_j, this correspondence must be non-increasing since the set S_i is (weakly) decreasing in I_j. We shall assume in what follows that aggregate welfare is strictly concave in I_1 and I_2. In that case, $I_i^o(I_j)$ is a differentiable function, $I_1^o(I_2)$ and $I_2^o(I_1)$ cross only once, and they have the relationship depicted by the dashed lines in Figure 3.2.

Comparison of expression (4.1) with expressions (3.3) and (3.5) allows us to establish the following result:

Proposition 4.1. *For a given level of I_j,*

(1) *Di under-invests relative to what is socially optimal under non-integration; that is,* $\max \{I_i^*(I_j | NI)\} \leq I_i^o(I_j)$.

(2) *Di under-invests relative to what is socially optimal under vertical integration by Dj; that is,* $I_i^*(I_j | VI_j) \leq I_i^o(I_j)$.

(3) *Di over-invests relative to what is socially optimal under vertical integration by Di; that is,* $I_i^*(I_j | VI_i) \geq I_i^o(I_j)$.

The under-investment result in Proposition 4.1 for non-integration arises because Di loses half of any value increase to U through bilateral bargaining. This effect parallels that which would arise in our model if we had only a *single* downstream firm which engaged in bargaining with U. The most interesting aspect of Proposition 4.1, however, is the over-investment by the integrated owner arising under vertical integration. Its cause is fundamentally linked to our multilateral setting and relates to the role of increases in v_i in raising Di's bargaining power. Under structure VI_i, the owner of $\{U, Di\}$ receives Di's value v_i not only when he uses some of U's input in Di ($Q = 1$ or $Q = 2$ and $s \in S_i$) but *also* when $Q = 1$ and he sells a unit to Dj ($s \in S_j$). The former cases create a socially optimal marginal return to investment for the owner of $\{U, Di\}$, while the latter- in which increases in v_i raise $\{U, Di\}$'s payoff by increasing the price that Dj must pay to get the input- create the excess investment incentive.

---

16. Since ex post trade is always efficient, it is sufficient for the planner to merely control I_1 and I_2.

Letting (I_1°, I_2°) denote the social optimum, we also have the following result regarding the equilibrium level of investment under structure VI_i:

Proposition 4.2. *Under Di vertical integration, the level of Di's investment is more than, and the level of Dj's investment is less than, their socially optimal levels. That is, $I_i^*(VI_i) \geq I_i^\circ$ and $I_j^*(VI_i) \leq I_j^\circ$.*

Proof. Since $I_j^*(I_i | VI_i) = I_j^*(I_i | NI) \leq I_j^\circ(I_i)$ any equilibrium under VI_i must be in the set $\{(I_1, I_2) | I_j \leq I_j^\circ(I_i)\}$. Also, since $I_i^*(I_j | VI_i) \geq I_i^\circ(I_j)$, any equilibrium under VI_i must be in $\{(I_1, I_2) | I_i \geq I_i^\circ(I_j)\}$. The concavity of $W(I_1, I_2)$ then yields the result.   ∥

Note that it is not necessarily true that $I_i^*(NI) \leq I_i^\circ$: with decreasing non-integration best-responses, we may get the reverse inequality holding for one of the downstream firms (Figure 3.2 makes this point clear).

(ii) The optimal Ownership Structure

We now investigate the determination of the socially optimal ownership structure. We first note that horizontal integration can never be the uniquely optimal ownership structure.

Proposition 4.3. *Horizontal integration is (weakly) dominated by complete integration.*

Proof. Suppose that under HI the owner of $\{D1, D2\}$ manages asset Di. As we noted earlier, this must result in $I_j = 0$. If the owner under CI manages Di, however, then we also get $I_j = 0$, but since I_i is then set equal to $I_i^\circ(I_j)$, CI must (weakly) dominate HI from a social perspective.   ∥

Thus, we can reduce our focus to structures NI, VI_i, and CI. Given the results of sub-section (i), the relationship of the investment levels under these structures and those in the social optimum can be seen in Figure 3.2.[17] In general, it is difficult to say much about which structure is optimal as this involves making global comparisons among the socially sub-optimal (i.e., non-first-best) outcomes of the various structures. However, in certain extreme cases we can reach definitive conclusions. The following result identifies conditions under which each of these three structures is necessarily optimal.[18]

Proposition 4.4. *Non-integration, vertical integration, and complete integration can each be (uniquely) optimal. In particular,*

(1) *If one downstream firm has no investment decision ($\bar{I}_i = 0$) or has a marginal effect on v_i of zero in every state, then complete integration is optimal. More generally, this is true whenever the social optimum involves one downstream firm doing no investment. In these cases, complete integration achieves the first-best outcome.*

(2) *If the probability of a shortage is zero ($\lambda = 0$) then vertical integration, either by D1 or by D2, is optimal.*

(3) *If the probability of a shortage is one ($\lambda = 1$), then non-integration is optimal. In this case, non-integration implements the first-best outcome.*

The arguments for these results follow easily from our earlier discussions. The reason for (1) is clear: if one D does no investment in the social optimum, then complete integration can implement this outcome. More striking, perhaps, is the fact that vertical integration by the D with an investment decision is *strictly inefficient* in this case since it

17. In Figure 3.2, we assume that there is a unique equilibrium under each structure. If there are multiple equilibria, they all have the relationships depicted there.
18. Although these structures need not be uniquely optimal in such circumstances, it is not difficult to see that they will be in a wide subset of these cases.

leads to over-investment by this integrated firm. In case (2), a move from NI to VI_i causes no change in Dj's investment or payoff, but increases the joint payoff of assets U and Di, so VI_i dominates NI (this is similar to the bilateral case). It also weakly dominates CI since it results in the same level of I_i and, possibly, a positive level of I_j. Case (3) is the most interesting from our perspective. When $Q = 1$, the investment incentives under non-integration are exactly those of the social planner since the under-investment effect under non-integration arises solely because of state $Q = 2$.

Two points should be emphasized about Proposition 4.4's identification of circumstances in which non-integration is the optimal structure. First, it is easy to see that if there were only *one D* in our model, then vertical integration (which can also be viewed as complete integration in that case) would always be the socially optimal ownership structure. When we move to a multilateral setting, however, *transactions cost considerations alone* may make non-integration the optimal structure. Second, non-integration emerges as optimal here precisely when *supply assurance concerns are greatest*, that is, when $\lambda = 1$. This result is striking given the evidence suggesting that supply assurance concerns cause firms to vertically integrate. Thus, if this evidence is correct, our results suggest that there may be a divergence between the social and private incentives to integrate. We now turn to a formal investigation of the private incentives for integration in our model.

5. PRIVATE INCENTIVES: STABLE OWNERSHIP STRUCTURES

Up to this point we have taken the existing ownership structure as a given and have not discussed its determination. In this section, we consider this issue. In particular, we imagine that during the interval $[t_0, t_1)$ ownership arrangements can be changed.[19]Then, starting at time t_1 the game proceeds as before, with the ownership structure being that which has emerged during the period $[t_0, t_1)$. We assume throughout that the equilibrium investment levels arising under a given ownership structure are unaffected by the particular trades that have led to it.[20]Thus, we associate with each ownership structure a unique pair of investments. When we rank ownership structures according to social welfare, we do so for these investment levels.

A central question of our analysis concerns whether we should expect to see the socially optimal ownership structure arise in equilibrium. In the bilateral setting commonly examined in the literature the answer is immediate; since the two parties can arrange a mutually profitable exchange to the socially optimal structure from any suboptimal one, and cannot find any mutually advantageous trade away from it, only the socially optimal structure can arise in an equilibrium.

Matters are more complicated, however, in a multilateral setting since trades can potentially take place between subsets of the owners. This would not be a problem if individuals could write a contract at time t_0 that prevented them from engaging in any future trades of assets; then, as in the bilateral case, there could always be a mutually advantageous trade to the socially optimal structure with an agreement that forbade any future asset exchanges. In practice, however, it may be difficult to make use of such restrictions. Since an owner may need to trade his asset for a number of exogenous reasons, disallowing all future trades of assets may be grossly inefficient. Yet, it may be difficult to contractually specify the precise set that are permitted if some trades are to be allowed. In this section, we explore the determination of ownership structure when

---

19. Since the only efficiency consequences of ownership are for ex ante investment, there is never any incentive to alter the ownership structure after investments have been chosen.

20. Of course, this must be true when there is a unique equilibrium given the ownership structure.

individuals can agree only to simple exchanges of titles and ask whether this process can be expected to lead to a socially optimal ownership structure. We do so by attempting to identify which ownership structures, if any, are stable in the sense that at no time during $[t_0, t_1)$ can any parties arrange a mutually desirable exchange that breaks that structure.

An important part of the question of whether a particular trade is mutually desirable, of course, concerns the set of future trades that may occur following the one in question. In the analysis below we employ a *weak* notion of stability by imposing relatively stringent conditions for a successful deviating trade to occur. Despite the weakness of the notion of stability we consider, however, we get surprisingly far in determining the possible outcomes of this process. Specifically, we use the following concept which we call *quasi-stability*;

Definition 5.1. An ownership structure is *quasi-stable* if and only if there does not exist a trade that is *guaranteed* to strictly raise the payoffs of all parties to that trade. By *guaranteed* we mean that this payoff increase occurs for each individual involved in the trade regardless of whether any further trades occur that do not involve that individual.

Some notation may help make Definition 5.1 more precise. Consider an initial ownership structure X and some individual i. Let $\Gamma(i, X)$ denote the set of all *finite sequences of trades* starting from structure X that do not involve individual i. An element of such a sequence is a description of the asset and monetary exchanges that take place in that trade. In addition, let $\xi(i, X)$ denote the set of *terminal ownership structures* that result from trade sequences in set $\Gamma(i, X)$ starting from structure X. According to Definition 5.1, and ownership structure $\tilde{X}$ is not quasi-stable if there is a trade among some set of individuals I resulting in some new ownership structure X such that each individual $i \in I$ has a strictly higher payoff *for any* structure in $\xi(i, X)$ (including any monetary transfers from the trade) than he has in structure $\tilde{X}$. We shall say that a trade satisfying this criterion "*q-breaks*" ownership structure $\tilde{X}$.

Clearly, no structure that is not quasi-stable can be regarded as a reasonable outcome. The converse, however, cannot be claimed. Nevertheless, it turns out that in our model we can get quite far using only this concept.

Before proceeding we first briefly set some notation: let $\pi(\{\cdot\cdot\} \mid X)$ denote the payoff, net of investment costs, to the owner of assets $\{\cdot\cdot\}$ under ownership structure X and let $\Pi(X)$ denote the aggregate payoff net of investment costs to the owners under structure X.

We first argue that if some structure other than non-integration is the socially optimal structure, then it is the only possible outcome.

Proposition 5.1. *If a one or two-owner structure is socially optimal, then it is quasi-stable. Furthermore, if it is uniquely optimal, then no other ownership structure can be quasi-stable.*

Proof. Suppose, first, that CI is optimal. No trade to another structure can break this structure: if new owners are to be guaranteed non-negative profits in the event that no further trades occur, the initial owner cannot be better off since the aggregate payoff of the assets falls (weakly). Similarly, if CI is uniquely optimal, then no other structure can be quasi-stable: starting from any other structure a (possibly multilateral) trade to CI can be constructed that is guaranteed to raise all participants' payoffs since the new owner's guaranteed payoff is his payoff under CI and this is strictly larger than the initial aggregate payoff.

Now suppose a two-owner structure is optimal. Recall that this cannot be *HI*. Thus, assume without loss of generality that it is VI_i. We first argue that VI_i is quasi-stable. To see this, note that no trade that involves both owners can q-break this structure since the aggregate payoff falls (weakly) if no further trades occur. On the other hand, the only trade involving one owner that could possibly q-break this structure involves the owner of $\{U, Di\}$ and results in *NI*. But, since $\Pi(NI) \leqq \Pi(VI_i)$ and (by Proposition 3.1) $\pi(\{Dj\} \mid NI) \geqq \pi(\{D_j\} \mid VI_i)$, it must be that $[\pi(\{U\} \mid NI) + \pi(\{Di\} \mid NI)] \leqq \pi(\{U, Di\} \mid VI_i)$. Thus, such a trade cannot q-break VI_i since the joint payoff to assets U and Di falls (weakly) if no further trade occurs.

Finally, no other structure can be quasi-stable if VI_i is uniquely optimal: from any other structure, a trade to VI_i can be constructed that is guaranteed to raise all participants' payoffs. In particular, the resulting owners' guaranteed payoffs are exactly their payoffs under VI_i: the new owner of $\{U, Di\}$ is clearly unaffected by any trade involving asset $\{Dj\}$ and, since $\pi(\{Dj\} \mid NI) \geqq \pi(\{Dj\} \mid VI_i)$, the new owner of asset Dj can only benefit from a trade involving $\{U, Di\}$'s owner. ‖

Proposition 5.1 indicates that in our model the only possible difficulty for obtaining the socially optimal outcome arises when non-integration is the socially preferred outcome. Our next proposition shows that in such cases it may be impossible to sustain the socially efficient ownership structure. Indeed, the proposition shows that in precisely the circumstances that we identified in Proposition 4.4 in which non-integration achieves the first-best outcome, it is *not* quasi-stable. The important implication of this result when combined with Proposition 5.1 is that, in our model, there is *excess integration from a social standpoint*.

Proposition 5.2. *If non-integration is socially optimal, it may not be quasi-stable. In particular, if $\lambda = 1$, outside options always bind, and non-integration is the unique socially optimal structure, then non-integration is not quasi-stable.*

Proof. Under the hypotheses described, a sale of asset U to asset Di's owner can be constructed that is guaranteed to strictly increase both owners' payoffs. To see this, note first that, for any given level of I_i, the joint payoff to assets U and Di is independent of I_j under both *NI* and VI_i (since outside options are binding it equals exactly the expected value of v_i less I_i). Since the new owner of $\{U, Di\}$ will choose I_i to maximize the joint payoff of these two assets (and there is a unique maximizer), this implies that $\pi(\{U, Di\} \mid VI_i) \geqq [\pi(\{U\} \mid NI) + \pi(\{Di\} \mid NI)]$ with strict inequality if $I_i^*(VI_i) \neq I_i^*(NI)$. But $\Pi(NI) > \Pi(VI_i)$ implies that investments differ under the two ownership structures and, since $I_j^*(I_i \mid NI) = I_j^*(I_i \mid VI_i)$ for all I_i, this means that $I_i^*(VI_i) \neq I_i^*(NI)$. Since the new owner of $\{U, Di\}$ is guaranteed a payoff of at least $\pi(\{U, Di\} \mid VI_i)$, a trade can be constructed that q-breaks *NI*. ‖

The key elements leading to Proposition 5.2 are the negative externality that a move to VI_i imposes on Dj combined with the fact that the joint payoff to assets U and Di is unaffected by any induced changes in I_j when $Q = 1$ and outside options bind.

To complete the analysis of this section, two further questions must be answered. First, when non-integration is socially optimal but is not quasi-stable, what ownership structure should we expect to see? The next proposition shows that it must be the (socially) next best ownership structure.

Proposition 5.3. *When non-integration is socially optimal but is not quasi-stable, the ownership structure with the next highest social payoff is quasi-stable and is the unique quasi-stable structure if it is uniquely second-best.*

Proof. See appendix. ‖

Second, when non-integration is socially optimal and quasi-stable (as it can be when $\lambda < 1$ if non-integrated firm Dj's investment level falls sufficiently after a merger of U and Di to make such a merger unprofitable), is it the unique quasi-stable outcome? Not necessarily. Even if the joint payoff of U and Dj are higher under NI than under VI_j, it may be the case that there is no trade from VI_j to NI that is *guaranteed* to raise all participants' payoffs. In particular, the new owner of Dj may fear a future merger by U and Di to VI_i, while the new owner of asset U may fear a merger by the D's to HI, and so VI_j could be quasi-stable. Similar concerns may lead CI to be quasi-stable. All we can say, following arguments similar to Proposition 5.3's proof, is that only the second-best structure can also be quasi-stable in this case. To go further, one must strengthen the notion of stability. One possibility is pursued in Bolton and Whinston (1990), where a strengthened stability notion is defined that selects non-integration as the unique outcome whenever it is quasi-stable.

6. EXTENSIONS OF THE ANALYSIS

In this section, we explore how the analysis of Sections 3–5 is modified when the model is extended to incorporate non-binding outside options (as well as other bargaining solutions), managerial quasi-rents, upstream investment, multiple upstream assets, and joint ownership.

(i) Outside Options Not Binding and Other Bargaining Solutions

As noted earlier, for expositional purposes we have chosen to focus in Sections 3–5 on the case where outside options are always binding. What happens when they are not?

The effect of this change is to alter the ex post (gross) payoffs for the cases of vertical integration and non-integration when capacity $Q = 1$; in all other cases, either bargaining does not occur (i.e., under structure CI) or outside options are irrelevant (under HI and when $Q = 2$). The entries for these two cases in Table 2.1 become (below we let $\tilde{v} \equiv (v_1 \vee v_2)$):

Ownership Structure	Capacity	U	Bargaining Payoff to: D1	D2
VI_1	$Q = 1$		$v_1 \vee \dfrac{v_2}{2}$	$\tilde{v} - \left(v_1 \vee \dfrac{v_2}{2}\right)$
NI	$Q = 1$	$\sum_i \left(v_i \vee \dfrac{v_i}{2}\right) - \tilde{v}$	$\tilde{v} - \left(v_2 \vee \dfrac{v_1}{2}\right)$	$\tilde{v} - \left(v_1 \vee \dfrac{v_2}{2}\right)$

Consider, for example, the case of VI_1 when $Q = 1$ and suppose that $v_2 > v_2/2 > v_1$. In this case, a sale to $D2$ is efficient, but the integrated owner's threat to use the input internally lacks credibility because he can get at least $v_2/2$ (his payoff if he bargains without any outside option) by continuing to bargain. Hence, the owners of $\{U, D1\}$ and $D2$ must split the surplus in this case, each receiving $v_2/2$, their payoff in the absence of the option of internal use. A similar point applies to the credibility of U's threat to sell to $D1$ when $v_2 > v_2/2 > v_1$ under NI. Note that as in Section 2, for given values of v_1 and v_2, the payoff to a non-integrated firm is the same whether buying from a non-integrated or an integrated upstream firm.

The expected gross and marginal payoffs in this case can be represented in figures similar to those used earlier. Under structure VI_1, for example, the payoff when capacity $Q = 1$ to the owners of $\{U, D1\}$ and $D2$ are depicted in Figure 6.1 by areas (A) and (B) respectively. The marginal investment incentives can be seen, as before, by shifting the curves in the figure.
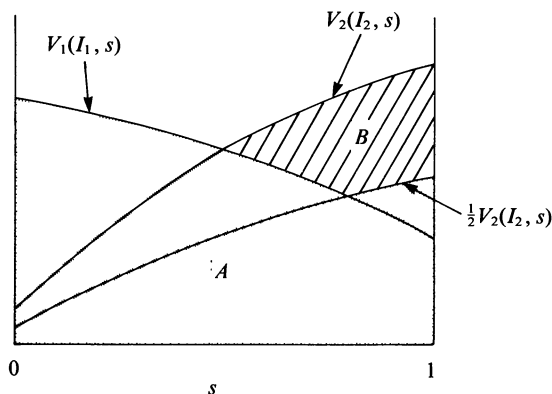

FIGURE 6.1

There are two primary changes arising from introducing non-binding outside options. First, a non-integrated D's best-response function $I_i^*(I_j \mid NI)$ is no longer necessarily decreasing in I_j. Nonetheless, our general results about equilibrium investments under NI and VI continue to hold if we impose some mild additional conditions.[21] Second, if outside options do not always bind, non-integration no longer achieves the first-best when $\lambda = 1$. Nevertheless, all of the other results in Sections 4 and 5 continue to hold (when NI is socially optimal, it still may or may not be quasi-stable).

The first-best efficiency of non-integration when $\lambda = 1$ may also not hold in bargaining processes in which the "outside option" of internal use is treated differently. As an example, suppose that the owner of U rents the input in every period instead of selling it to one of the D's and that the upstream and downstream owners alternate in making take-it-or-leave-it offers for the rental of the input in every rental period. In this case, when $Q = 1$ and $v_i > v_j$, the owners of U and Di split the gain in surplus over U's outside option, so that the owner of Di receives a payoff of $(v_i - v_j)/2$. If surplus is split in this way, then as when outside options do not always bind, non-integration will no longer achieve the first-best when $\lambda = 1$, but our basic conclusions remain unaffected.

(ii) Managerial Quasi-Rents for Downstream Managers

Above we made the strong assumption that a manager who runs an asset during the investment stage possesses no advantage over any other potential manager in the production stage. Thus, we have excluded such possibilities as managerial learning-by-doing

---

21. In particular, if each Di's non-integration payoff function is strictly concave in its investment level I_i and if the non-integration best-response functions satisfy a standard stability condition, then part (3) and the increase in I_j in part (1) and (2) of Proposition 3.1 continue to hold: vertical integration leads to an increase in the integrated downstream firm's investment and to a reduction in the non-integrated firm's expected payoff. Part (2) and Part (1)'s result regarding I_i also hold if the best-response of the non-integrated firm is non-increasing in the investment of the other downstream firm.

and investment in specific human capital. Since managers may capture some of the ex post surplus when such quasi-rents do exist, and hence may have positive ex ante investment incentives, it is natural to wonder how our results are affected by relaxing this assumption.

The existence of quasi-rents can only affect the outcomes of CI and HI in our model, since in all other structures both downstream assets are run by owners. Thus, here we focus only on these two structures. To consider a simple case, we assume that at the production stage there is a cost, given by $\theta > 0$, of replacing an incumbent manager.[22]

Suppose, without loss of generality, that asset $D1$ is run by a manager at the investment stage while asset $D2$ is run by the owner of the D's. We consider the following bargaining game during the production stage: first, the owner of the D's acquires the inputs (under CI this is automatic, while under HI he must bargain with U's owner) and then he bargains with $D1$'s manager over production if any is required (that is, if either $Q = 2$ or $v_1 > v_2$). This latter bargaining is again modeled as an alternating offer bargaining game where disagreement brings about a delay in production. In this game, the owner can also invoke some outside options: when $Q = 2$, the outside option is to replace the incumbent manager and earn $v_1 - \theta$; when $Q = 1$ and $s \in S_1(I_1, I_2)$, the best outside option for the owner may be either replacement of the manager or internal use of the input in the owner-managed asset, which yields $\max\{v_1 - \theta, v_2\}$. The outcome of this bargaining (under either CI or HI) therefore has immediate agreement and the manager receiving $\min\{\theta, v_1/2\}$ when $Q = 2$, and $\min\{\theta, v_1 - v_2, v_1/2\}$ when $Q = 1$ and $s \in S_1(I_1, I_2)$.[23]

How are managerial investment incentives under CI and HI affected by the presence of these rents? Consider CI first. When $Q = 2$, in all states s such that $v_1(I_1; s)/2 > \theta$ the manager earns exactly θ and faces a zero marginal return on investment; in all other states he earns $v_1/2$ and faces a gross marginal return on investment of $(1/2)(\partial v_1/\partial I_1)$. The interesting point to note is that as θ increases the manager eventually faces investment incentives that exactly equal those in structure VI_2.

Figures 6.2(a) and 6.2(b) depict the manager's expected payoff when $Q = 1$ for two given levels of θ, $\theta_b > \theta_a$. When θ is sufficiently large, the manager receives the full marginal increment in v_1 if $v_1/2 > v_1 - v_2$ and half the marginal increment if $v_1/2 \leqq v_1 - v_2$, so that again the manager's expected marginal return is identical to that arising under VI_2. In summary, as θ increases, the manager's incentives under CI come to approach, and ultimately equal, the incentives arising under VI_2. The same effect holds for the owner of $\{U, D1, D2\}$: his incentives approach and ultimately equal those of the owner of $\{U, D2\}$ under VI_2. Thus, as the level of managerial quasi-rents, θ, grows large, CI comes to resemble VI_2.

Under HI, the manager of D_2 has identical incentives to those arising under CI, since the bargaining solution is in all cases the same. However, because the upstream and downstream owners bargain over the joint surplus (net of the manager's earnings) the downstream owner's investment incentives differ from those of the owner of the completely integrated structure; in fact, the downstream owner's incentives equal exactly half of those arising under complete integration, for any level of θ. Thus, the excess incentives for his investment that arise as θ gets large under CI are somewhat ameliorated under HI. This could lead HI to be socially preferred to CI.[24]

22. More complicated cases might have the level of quasi-rents affected by managerial actions.

23. Note that under HI the input price agreed to is irrelevant for the bargaining solution between the owner and manager since the price for the input is a sunk cost at the time of bargaining.

24. Under certain conditions, HI may even implement the first-bet outcome. Specifically, the downstream owner's expected marginal returns may equal their first-best level if the surplus appropriation by the upstream owner exactly offsets the excess incentives arising under CI.
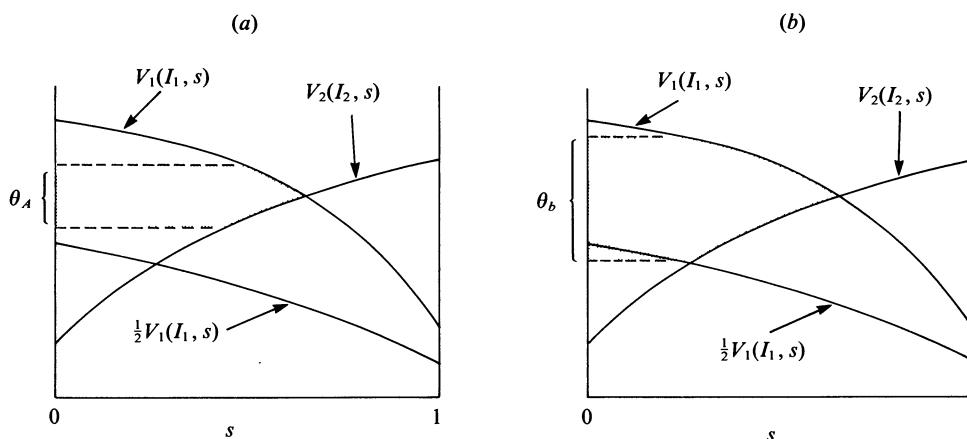
FIGURE 6.2

(iii) Upstream Investment

In the model analyzed above we assumed that only downstream managers engage in *ex ante* value enhancing investment activities. This simplifying assumption both provided a clear bilateral benchmark (with only one downstream firm vertical integration is the uniquely efficient ownership structure) and also helped make the analysis of the investment effects of ownership more tractable. With the addition of upstream investment, the analysis becomes significantly more complicated. Here, we briefly discuss some of the effects of introducing upstream investment for the special case where *only* upstream investment is present. Our particular focus in these remarks is on the comparison between *VI* and *NI*.[25]

Suppose that the upstream manager can raise the final net value of D_i's customer $v_i(I_i, s)$ by investing in effort I_i, and that his effort cost function is $c_u(I_1, I_2)$ with $\partial c_u(I_1, I_2)/\partial I_i > 0$ $(i = 1, 2)$ and $\partial^2 c_u(I_1; I_2)/\partial I_1 \partial I_2 > 0$ (i.e., I_1 and I_2 are substitutes).

Consider, first, investment levels under *NI*. Unlike the downstream investment case, *NI* does *not* generally implement the first-best outcome when $\lambda = 1$ and outside options bind. To see why, note that the upstream firm gets *none* of the increment in value of the "winner" (the downstream asset with the higher value v_i), but it gets *all* of the increment in value of the loser. This may lead to either over-investment or to under-investment in I_i.[26] When $\lambda \neq 1$ or outside options do not always bind, these distortions are compounded with those arising from surplus extraction by the downstream firms, an effect that pushes the outcome toward under-investment.

In contrast, under VI_i the owner of $\{U, Di\}$ completely specializes his investment when $\lambda = 1$ and outside options always bind, setting $I_j = 0$ and over-investing in I_i. This effect arises because the owner of $\{U, D1\}$ receives v_i both when he uses U's input

---

25. The outcomes of horizontal integration and complete integration are easy to discern. The reader may be concerned that in this setting complete integration is always optimal. However, if we introduce downstream investment that affects some fixed cost realized at the investment stage rather than ex post values, then the analysis in the text would remain unchanged, but complete integration would in general not be optimal.

26. In a few special cases, however, the first best is implemented when $\lambda = 1$ and outside options bind. For example, whenever investment affects the "lower envelope" of the value functions, $v_i(I_i, s)$, in the same way as it affects the "upper envelope", the first-best can be implemented under *NI*. One case where that property holds is when the model is symmetric and $\partial v_i(I_i; s)/\partial I_i$ is independent of s for $i = 1, 2,$. Another arises when the cost function has a Leontieff form (so that the upstream manager always chooses $I_i = I_j = I$) and $\partial v_1(I; s)/\partial I - \partial v_2(I; s)/\partial I$. This latter case arises, for example, if upstream investment reduces upstream marginal production costs.

internally and when he sells it to Dj; therefore in every state he appropriates all of the increments in v_i and none of those in v_j. In addition, the reduction in investment by the external customer is also reinforced by the increase in the marginal cost of I_j caused by the increase in I_i. Hence, I_j falls even when $\lambda = 0$.

The effects of vertical integration on upstream investment are therefore not very different from those on downstream investment. In both cases integration leads to a shift in investment toward the internal customer and a decrease in the non-integrated firm's payoff relative to non-integration.[27]

(iv) Multiple Upstream Assets

The empirical literature on vertical integration decisions often notes the occurrence of "chains of integration", in which the integration of one set of firms leads other firms to themselves integrate (see, for example, Chandler (1964) and Scherer (1980)). Along with these observations often goes a suggestion that the resulting level of integration is excessive. To consider such issues, we need to extend our model to allow for multiple upstream assets. Here we do so by assuming that although prior to time t_0 there is a single upstream asset, additional upstream assets can be built during the time interval $t \in [t_0, t_1)$ at a cost of K. The rest of the model is as before, except for one difference. To keep to cases where our previous bargaining analysis applies, we assume that investment is *supplier-specific* (otherwise we would get into issues of bargaining with two buyers and two sellers). Specifically, each Di can tailor his investment toward only *one* supplier and if Di does no investment toward a particular supplier, then its value for that supplier's input is *zero* for *all states s*. In what follows we show that this extension of our model is indeed characterized by socially excessive chains of integration and we also discuss how the possibility of additional upstream assets being built can affect which *single U* asset structures may arise.

We begin with a discussion of social optimality before turning to an analysis of private incentives. From our previous analysis we know that among ownership structures with a single upstream asset, HI is dominated by CI. In addition, here CI is in turn dominated by either VI_1 or VI_2 (since under CI, one of the downstream investment levels must be zero), so that NI, VI_1, and VI_2 are the only possible socially optimal ownership structures involving a single upstream asset.[28] Among ownership structures with more than one upstream asset, on the other hand, the *parallel integration structure* $PI \equiv (\{U1, D1\}, \{U2, D2\})$ is uniquely best socially, achieving the first-best payoff conditional on having more than one U (supply shortages disappear once we have more than one upstream asset and—as in the simple bilateral model—vertical integration then achieves efficient investment levels).[29]

The determination of the socially optimal structure mixes considerations of scale economies and transactions costs. An implication of our earlier results is that the presence

---

27. Note that the negative externality on Dj resulting from a move from NI *to* VI_i depends in part on our assumption about the sign of the cross-partial derivative of the cost function $c_u(\cdot)$. If we reversed our assumption, so that upstream investments are complementary, then the over-investment in I_i resulting from VI_i could raise I_j as well (if $\lambda < 1$), creating a positive externality for D_j.

28. Strictly speaking, *strict* domination in each case requires that there be some change in investment (e.g., that under VI_i, Dj invests a positive amount).

29. One complication here is that when there are several upstream assets we need to verify that equilibrium behaviour under structure $(\{U1, D1\}, \{U2, D2\})$ has each D investing towards its U (potentially Di may invest towards Uj because Dj is investing toward Ui, and vice versa). In Bolton and Whinston (1990) (Proposition 7.2) we show that in any equilibrium of structure $(\{U1, D1\}, \{U2, D2\})$, each D will necessarily invest toward his own U.

of incomplete contracting makes it more likely that the socially optimal structure here will be one with two upstream assets (since the best structure with two upstream assets always achieves the first-best outcome conditional on having more than one upstream asset).

The analysis of private incentives in the context of this extended model is considerably more complicated than in the single U case. In Bolton and Whinston (1990), we analyze private incentives using a similar approach to that in the previous section (i.e., an extension of the quasi-stability concept). Given space constraints, here we shall instead illustrate some of the basic points that emerge from that analysis in a simple, although somewhat *ad hoc*, extensive-form game of asset trades and upstream asset building between times t_0 and t_1.

In this game, at time t_0 there is initially a single upstream asset $U1$ and the three assets $U1$, $D1$, and $D2$ are separately owned. The game then consists of three stages: in stage 1, the owner of $U1$ makes a merger offer to any of the two D's (he chooses which one) and that downstream firm—say Di—can then accept or reject $U1$'s offer; in stage 2, regardless of the outcome of this bargaining, the other downstream firm Dj can build its own upstream asset at a cost of K; finally, if Dj builds in stage 2 and no merger has previously occurred, then in stage 3 the owner of $U1$ can make a take-it-or-leave-it merger offer to Di.

In what follows, the following notation is useful: let $I_i^{**} = \mathrm{argmax}_{I_i}\,[Ev_i(I_i, s) - c_i(I_i)]$ and $\hat{I}_i = \mathrm{argmax}_{I_i}\,[\frac{1}{2}Ev_i(I_i, s)c_i(I_i)]$. I_i^{**} is the efficient investment level in asset Di with more than one upstream asset; $\hat{I}_i$ is a non-integrated Di's privately optimal investment when he is certain to get a unit of input but must split his surplus with his supplier's owner. Also let $v_i^{**} \equiv Ev_i(I_i^{**}, s)$, $c_i^{**} \equiv c_i(I_i^{**})$, $\hat{v}_i \equiv Ev_i(\hat{I}_i, s)$, and $\hat{c}_i \equiv c_i(\hat{I}_i)$.

We note first that this model has a chains of integration property. This is true in two senses. The first is that if a downstream firm, say Dj, builds an upstream asset when no previous merger has occurred, then $U1$ and the other downstream firm Di will consummate a merger in response (since $U1$ will be Di's sole supplier and Di will be $U1$'s sole buyer we are back to a bilateral setting).[30] The second is that a merger (in stage 1) of assets $U1$ and Di increases Dj's incentive to build. This is so because the merger has a negative externality on Dj, lowering his payoff from not building, while his payoff from building is the same, $(v_j^{**} - c_j^{**}) - K$, whether $U1$ and Di have already merged or not. Thus, in this model, integration by one downstream firm tends to cause the other to integrate as well.

Second, from a social standpoint, there is an excessive tendency for an additional upstream asset to be built:

Proposition 6.1. *Whenever structure $PI \equiv \{(U1, D1), (U2, D2)\}$ is socially optimal, it is the unique subgame-perfect Nash equilibrium of the extensive form merger/building game.*

Proof. See appendix. ‖

The cause of this excessive building (an example of which is given below) is the negative externality that a D has on the other two assets when it builds.

Lastly, the presence of the opportunity to build additional upstream assets can affect *which* single upstream asset structures can emerge (when one still does). To see this,

---

30. As in footnote 30, matters are actually somewhat more complicated than this because there may be another continuation equilibrium where Dj invests toward $U1$ and Di invests toward $U2$. However, it is not difficult to show that $U1$ and Di will want to merge in this case as well.

consider the special case where $\lambda = 0$ (so that capacity is always $Q = 2$). Define $\varepsilon_i \equiv (v_i^{**} - c_i^{**}) - (\hat{v}_i - \hat{c}_i) \geqq 0$ to be the social gain from moving from NI to VI_i here and assume that $\varepsilon_1 > \varepsilon_2 > 0$ so that $\Pi(VI_1) > \Pi(VI_2) > \pi(NI)$. In this case, if there were no opportunity to build (i.e., if the last two stages of the game were deleted), then VI_1 would be the unique subgame-perfect equilibrium outcome.

With the possibility of building, however, VI_1 may not be the single upstream asset structure that we observe. To see this, denote the private gain to Dj from building as $\theta_j \equiv [(v_j^* - c_j^*) - (\frac{1}{2}\hat{v}_j - \hat{c}_j)]$ (note that it is the same whether $U1$ and Di have previously merged or not in this example) and consider a situation where $\theta_2 > K > \theta_1$. In this case, $D2$ will build if it is not intregrated with $U1$, while $D1$ will not, and $\Pi(VI_2) > \Pi(PI)$ (this follows because building by $D1$ from VI_2 has a negative externality). In the equilibrium of this model, $U1$ will merge with $D2$ and the result will be structure VI_2 ($U1$ earns exactly θ_1 by making an optimal offer to $D1$, and earns at least θ_2 by making $D2$ an optimal offer).

The idea here is that $D2$ may be a "large" (high value) firm which, while it does not gain much in investment efficiency from merger, loses a lot in bargaining with a separately owned upstream firm (indeed, we could have $\varepsilon_2 = 0$); in contrast, the "small" (low value) firm $D1$ does not lose enough in bargaining to justify building its own upstream unit. (This interpretation comes from imagining an extension of this model in which downstream firms sell multiple units of their final good). Thus, the single U structures we observe may reflect not only the investment efficiency effects analyzed earlier, but also downstream assets' threats of building additional upstream assets. Interestingly, these considerations introduce "firm size" as a factor in the determination of equilibrium ownership structures.[31]

(v) Joint Ownership

It is natural to wonder about the effects of allowing joint ownership of assets among several individuals. In analyzing such cases, however, one must confront a number of fundamental issues. The first concerns the question of how control is allocated within such organizations. Many possibilities can arise; decisions may be subject to unanimous consent by all owners, majority voting based on individuals' share of ownership, a random allocation of control to one of the owners, or even an allocation of control determined through more complicated announcement games (e.g. Maskin (1977)). Because these various governance structures may significantly vary in their transactions costs, it is difficult to determine the relevant set to analyze. A second concern arises in trying to model simultaneous bargaining processes internal and external to the firm. For example, with a unanimity structure all owners of a firm need to be involved in making decisions involving bargaining with outsiders. These matters can be particularly complicated when a single individual is a part owner of several different firms. Fully analyzing these issues is beyond the scope of this paper. Nevertheless, a couple of simple points seem likely to emerge from any such analysis.

First, joint ownership of the upstream asset by the two downstream owners (either two owners possessing three assets jointly or each of two owners jointly owning U and each individually owning a D) together with a unanimity rule is unlikely to be the optimal ownership structure. Since bilateral bargaining between the joint owners gives each half of the total ex post surplus, the marginal returns to each owner from investment in the

---

31. When $\theta_2 > \theta_1 > K$, the equilibrium yields structure PI. Note that when θ_1 is close to K we still have $\Pi(VI_2) > \Pi(PI)$, and so we have an example here of socially excessive building of upstream assets.

downstream firm he manages will now always be half of the social marginal return. As a consequence, the downstream manager's best-response functions under this type of joint ownership will lie below those arising under non-integration. Thus, non-integration is likely to dominate this form of joint ownership.

Second, in situations where there is over-investment, joint ownership—if feasible— may help mitigate excess incentives. For instance, under VI_i, two person joint ownership of $\{U, Di\}$ with a unanimity rule will give the owner-manager of that firm only half of the marginal incentive arising with undivided ownership; thus, this structure could be socially preferable to either VI_i or NI. Similar effects can be achieved by allocating control stochastically among owners.[32] Note, though, that as long as ownership (e.g., probabilities of control) can be traded, private incentives may lead to an outcome other than the social optimum.

9. CONCLUSION

The preceding analysis reveals the importance of introducing the multilateral aspects of real-world production and distribution relations into the positive and normative analysis of the "scope of the firm". In our model, the presence of a multilateral supply setting has important implications for the effects and welfare properties of various ownership structures. Transactions costs savings are often a two-edged sword, with the alleviation of supply assurance concerns for merging parties often exacerbating supply assurance concerns for other downstream firms and leading to a form of market foreclosure. As a result of these externalities, the equilibrium and socially optimal ownership structures may differ markedly from those that would be derived based solely on bilateral considerations and as well as from each other.

APPENDIX

Proof of Proposition 5.3. (i) *If CI is a second-best structure.* Clearly the only trade that can q-break CI is a trade that results in NI. Without loss of generality we can restrict attention to trades that result in three new individuals owning the assets (we can always think of the original owner as a separate buyer and seller in other cases). Now consider a trade that q-breaks NI (by hypothesis, such a trade exists) and let Z be the structure that results from this trade. This trade must involve exactly two of the three new owners and result in their assets becoming jointly owned (no trade involving all three owners can q-break NI since the aggregate payoff falls). We denote the assets owned by these owners by A and B and the assets owned by the excluded owner by E. Now, in the initial trade away from CI, the new owners of assets A and B can each pay at most their profit level under NI if they are to be guaranteed non-negative payoffs (since there might be no further trades). The new owner of asset E, on the other hand, can pay at most $\pi(\{E\}|Z)$ if he is to be guaranteed a non-negative payoff. Since $\pi(\{A, B\}|Z) > [\pi(\{A\}|NI) + \pi(\{B\}|NI)]$, the initial owner must receive strictly less than $\Pi(Z)$ which, by hypothesis, is less than $\Pi(CI)$. Thus, this trade cannot break CI. Finally, if CI is uniquely second-best, then from any structure other than NI, a trade to structure CI can be constructed that is guaranteed to strictly increase the payoffs of all participants (the new owner of $\{U, D1, D2\}$ is guaranteed a payoff of at least $\Pi(CI)$).

(ii) *If VI_i is a second-best structure* (recall that HI cannot be). The only trade that can possibly q-break this structure is a trade that results in NI. An argument parallel to that above shows that no such trade involving both owners can q-break VI_i. The only alternative is a trade involving only the owner of $\{U, Di\}$. Again we can restrict attention to a sale of assets U and Di to two new owners without loss of generality. A necessary

condition for this trade to q-break structure VI_i is that it raise the joint payoff to these assets absent any further trades. If this is so, however, then the trade that q-breaks NI must result in either structure VI_j or structure HI and involve exactly two owners (in particular, a trade to structure VI_i cannot break NI).

Suppose, first, that a trade to VI_j q-breaks NI. Then it must be that $\pi(\{U, Dj\}| VI_j) > [\pi(\{U\}| NI) + \pi(\{Dj\}| NI)]$, and since $\Pi(NI) \geqq \Pi(VI_j)$, this implies that $\pi(\{Di\}| VI_j) < \pi(\{Di\}| NI)$. Thus the maximal amount that the initial owner of $\{U, Di\}$ can receive if the new owners are guaranteed non-negative payoffs is $[\pi(\{Di\}| VI_j) + \pi(\{U\}| NI)]$. But,

$$[\pi(\{Di\}| VI_j) + \pi(\{U\}| NI)] < \pi(\{Di\}| VI_j) + [\pi(\{U, Dj\}| VI_j) - \pi(\{Dj\}| NI)]$$

$$= \Pi(VI_j) - \pi(\{Dj\}| NI)$$

$$\leqq \Pi(VI_i) - \pi(\{Dj\}| VI_i)$$

$$= \pi(\{U, Di\}| VI_i),$$

where the weak inequality follows from the fact that, by hypothesis, $\Pi(VI_i) \geqq \Pi(VI_j)$ and that since VI_i results in backward integration, we know that $\pi(\{Dj\}| VI_i) \leqq \pi(\{Dj\}| NI)$. Thus, this trade cannot q-break VI_i,—a contradiction. A similar contradiction can be established for the case where a trade to HI q-breaks NI. Finally, if VI_i is uniquely second-best then starting from any structure other than NI a trade to VI_i can be constructed that is guaranteed to increase the payoffs of all participants (since the two new owners' guaranteed payoffs are exactly their payoffs under VI_i less any payments made). ‖

Proof of Proposition 6.1.   To establish the result, we argue that no matter who $U1$ decides to make an offer to in stage 1, the end result must be structure PI. Suppose that $U1$ decides to make an offer to Di. Note first that if Di accepts $U1$'s offer, then Dj will follow this by building its own asset. This is so because if building after this merger raises the social payoff (as we have assumed) and if it has a negative externality on $\{U1, Di\}$'s owner (as it does because this owner is weakly better off when there is a possibility of selling to Dj), it must raise Dj's payoff. In addition, because $U1$ and Di will react to Dj's building by merging if they have not merged already, the only way not to get structure PI is if $U1$ and Di do not initially merge and then Dj chooses not to build.

Now for this to be an equilibrium, if $U1$ and Di do not wish to merge, it must be that U and Di's joint payoff under Ni is higher than that under PI, which would result from their decision to merge, so that:

$$(v_i^{**} - c_i^{**}) \leqq \pi(U1| NI) + \pi(Di| NI).$$

But, combining this inequality with the assumption that $\Pi(PI) \equiv (v_1^{**} - c_1^{**}) + (v_2^{**} - c_2^{**}) - K > \Pi(NI)$ implies that

$$(v_j^{**} - c_j^{**}) - K \leqq \pi(Dj| NI).$$

But then Dj would prefer to build after $U1$ and Di fail to merge in stage 1 and so we must end up with structure PI. ‖

REFERENCES

BINMORE, K. G. (1985), "Bargaining and Coalitions", in A. E. Roth (ed.), *Game Theoretic Models of Bargaining* (Cambridge, England: Cambridge University Press).
BINMORE, K. G., RUBINSTEIN, A., and WOLINSKY, A. (1986). "The Nash Bargaining Solution in Economic Modeling", *Rand Journal of Economics*, 17, 176–88.
BINMORE, K. G., SHAKED, A. and SUTTON, J. (1989), "An Outside Option Experiment", *The Quarterly Journal of Economics*, 104, 753–70.
BOLTON, P. and WHINSTON, M. D. (1990), "Incomplete Contracts, Vertical Integration, and Supply Assurance" (Harvard Institute of Economic Research, Discussion Paper No. 1493 and Ecole Polytechnique, Laboratoire D'Econometrie Discussion Paper No. 332).
BOLTON, P. AND WHINSTON, M. D. (1991), "The 'Foreclosure' Effects of Vertical Mergers", *Journal of Institutional and Theoretical Economics*, 147, 207–26.
CARLTON, D. (1979), "Vertical Integration in Competitive Markets Under Uncertainty", *Journal of Industrial Economics*, 27, 189–209.

CHANDLER, A. D., Jr. (1964) *Strategy and Structure: Chapters in the History of American Industrial Experience* (Cambridge, MA: MIT Press).

CHANDLER, A. D., Jr. (1977) *The Visible Hand: The Managerial Revolution in American Business* (Cambridge, MA: Harvard University Press).

COASE, R. (1937), "The Nature of the Firm", *Econometrica*, **4**, 386–405.

GREEN, J. (1986), "Vertical Integration and Assurance of Markets", in J. E. Stiglitz and F. Mathewson (eds.), *New Developments in the Theory of Market Structure* (Cambridge: MIT Press).

GROSSMAN, S. J. and HART, O. D. (1986). "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration", *Journal of Political Economy*, **94**, 691–719.

HART, O. D. (1987), "Incomplete Contracts and the Theory of the Firm", *Journal of Law, Economics, and Organization*, **4**, 119–39.

HART, O. D. and MOORE, J. (1988), "Property Rights and the Theory of the Firm", *Journal of Political Economy*, **98**, 1119–58.

HART, O. D. and TIROLE, J. (1990), "Vertical Integration and Market Foreclosure", *Brookings Papers on Economic Activity, Microeconomics*, 205–76.

HOLMSTROM, B. and MILGROM, P. (1989), "Multi-task Principal-Agent Problems: Incentive Contracts, Asset Ownership, and Job Design", *Journal of Law, Economics, and Organization*, **7**, 24–52.

HOLMSTROM, B. and TIROLE, J. (1989), "The Theory of the Firm", in R. Schmalensee and R. D. Willig (eds.), *Handbook of Industrial Organization* (Amsterdam: North-Holland).

KLEIN, B., CRAWFORD, R. and ALCHIAN, A. (1978), "Vertical Integration, Appropriable Rents and the Competitive Contracting Process", *Journal of Law and Economics*, **21**, 297–326.

MASKIN, E. (1977), "Nash Equilibrium and Welfare Optimality" (mimeo).

ORDOVER, J., SALOP, S., and SALONER, G. (1988), "Equilibrium Vertical Foreclosure", *American Economic Review*, **80**, 127–42.

PORTER, M. E. (1980) *Competitive Strategy: Techniques for Analyzing Industries and Competitors* (New York: Free Press).

RUBINSTEIN, A. (1982), "Perfect Equilibrium in a Bargaining Model", *Econometrica*, **50**, 97–109.

SALINGER, M. (1988), "Vertical Mergers and Market Foreclosure", *Quarterly Journal of Economics*, **103**, 345–56.

SHAKED, A. and SUTTON, J. (1984), "Involuntary Unemployment as a Perfect Equilibrium in a Bargaining Model", *Econometrica*, **52**, 1351–1364.

SELTEN, R. (1975), "A Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games", *International Journal of Game Theory*, **4**, 25–55.

SUTTON, J. (1986), "Non-Cooperative Bargaining Theory: An Introduction", *Review of Economic Studies*, **53**, 709–724.

WILLIAMSON, O. E. (1985) *The Economic Institutions of Capitalism* (New York: Free Press).