LE
7/26/07
11:20 AM

Hold-up, Asset Ownership, and Reference Points

by

Oliver Hart*

April 2007

*Harvard University. I am grateful to John Moore for discussions on some of the elements of this paper, to Mathias Dewatripont and Bob Gibbons for helpful comments, and to Georgy Egorov for excellent research assistance. Financial support from the U. S. National Science Foundation through the National Bureau of Economic Research is gratefully acknowledged.

Abstract

We study the trade-off between contractual flexibility and rigidity. Two parties can write a flexible contract that adjusts price to circumstances or a rigid contract that fixes price. A flexible contract leads to argument and shading. A rigid contract gives one party an incentive to hold up the other if value or cost falls outside the normal range. An optimal contract trades off shading and hold-up costs. Asset ownership can help. If one party owns a key asset, his outside option will be high when his value from trade is high. This reduces both shading and hold-up costs.

1. Introduction

This paper reexamines some of the themes of the incomplete contracts literature – in particular, the hold-up problem and asset ownership – through a new theoretical lens, the idea that contracts serve as reference points. We consider a buyer and seller who are involved in a (long-term) economic relationship where the buyer's value and seller's cost are initially uncertain. For the relationship to work out the parties need to cooperate in ways that cannot be specified in an initial contract. The buyer and seller face the following trade-off. On the one hand they can write a flexible contract that attempts to index the terms of trade – price – to the state of the world. However, to the extent that value and cost are not objective, such a contract will lead to argument, aggrievement and shading in the sense of Hart and Moore (2007); this in turn creates deadweight losses. On the other hand the parties can write a (relatively) rigid contract, e.g., a fixed price contract. A rigid contract has the advantage that there is less to argue about in "normal" times, but the disadvantage that, if value or cost falls outside the normal range, one party will have an incentive to engage in hold-up, i.e., to threaten to withhold cooperation unless the contract is renegotiated. We suppose that hold-up transforms a friendly relationship into a hostile one. The consequence is that the parties operate within the letter rather than the spirit of their (renegotiated) contract, causing deadweight losses that are at least as great as those from shading. However, even a hostile relationship is assumed to create more surplus than no trade, and so, if value or cost has moved sufficiently far outside the normal range, hold-up will occur.

The optimal contract trades off argument/shading costs and hold-up costs. We show that an appropriate allocation of asset ownership can mitigate these costs. To see how this works note that, if the buyer (resp., the seller) owns key assets, then this improves his outside opportunities, and so in states of the world where his value inside the relationship is high, his value outside the relationship is also likely to be high. But this reduces the seller's (resp., the buyer's) gains from hold-up. In other words, the range of parameters over which hold-up is avoided is expanded. One feature of our approach is that, in contrast to much of the literature, it focuses on ex post rather than ex ante inefficiencies. Indeed (noncontractible) ex ante investments play no role.

The ideas analyzed in this paper are not new. Klein (1996) and Baker, Gibbons and Murphy (2002) have argued that integration, asset ownership, and related decisions are made by parties to maximize the range of parameters over which contracts can be sustained and breach, i.e., hold-up, avoided (see also Halonen (2002) and Klein and Murphy (1997)). Indeed Klein argues that "hold-ups occur when market conditions change sufficiently to place the relationship outside the self-enforcing range."[1] Klein and Baker et al. focus on situations where the parties' contracts are implicit or relational. In contrast we study explicit contracts. However, our contracts, like relational contracts, have the feature that there is a limited range of parameters over which they can be enforced.

While Klein's analysis is mainly informal, Baker et al. (2002) provide a formal analysis of relational contracts and hold-up in a world of uncertainty. Baker et al. use a standard property rights model and focus on ex ante rather than ex post inefficiencies. They are concerned mainly with how relational contracts and asset ownership can help to mitigate the underinvestment problem (see also Halonen (2002)). In contrast our model stresses the role of ex ante contracts and asset ownership in reducing the ex post inefficiency losses of hold-up and shading.

[1] Klein (1996) cites the Fisher Body-General Motors and Alcoa-Essex cases as examples where exceptional events triggered hold-up. See also Klein (2007).

The paper is organized as follows. In Section 2 we lay out the model. Section 3 introduces asset ownership. Section 4 concludes.

2. The Model

We consider a buyer B and a seller S who are engaged in a long-term relationship. The parties meet at date 0 and can trade a widget at date 1. There is uncertainty at date 0, but this is resolved at date 1. There is symmetric information throughout and the parties face no wealth constraints. Each party has an outside option that he (or she) earns if trade does not occur. Let v, c denote B's value and S's cost if trade proceeds smoothly (i.e., the parties cooperate at date 1), and let r_b, r_s denote B and S's outside options.

We follow Hart-Moore (2007) in supposing that for the gains from trade to be fully realized each party must take a number of "cooperative" or "helpful" actions at date 1. Some of these actions are contractible at date 1, while others are noncontractible; but all are hard to describe in advance and so cannot be specified in a date 0 contract. To make matters as simple as possible we suppose that the cost of a helpful action is zero for the party taking it – thus each party is indifferent between being helpful and not -- but that such an action can yield considerable benefit to the other party. We suppose that all cooperative actions are chosen simultaneously and independently by B and S at date 1. See Figure 1 for a time-line.

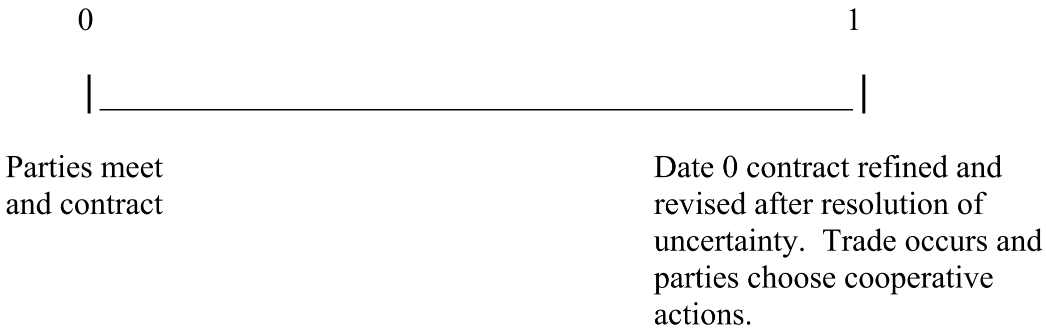

Figure 1

In an ideal world, in order to elicit cooperation and avoid hold-up, the parties would write a date 0 contract that indexes the trading price p to v, c, r_b, and r_s. However, although these variables are observable at date 1, they are not verifiable. We follow Hart and Moore (2007) in supposing that indexing on observable but nonverifiable information can lead to argument about which state has occurred, or, to put it another way, about what the appropriate price is; and this in turn causes aggrievement and the withholding of cooperation. In particular, as in Hart and Moore (2007, Section 3), we assume that at date 0 all the parties can do is to write a contract that specifies an interval of trading prices $[\underline{p}, \overline{p}]$. This contract serves as a reference point for date 1 entitlements in the sense that neither party feels entitled to an outcome outside those permitted by the contract. However, within the contract there can be disagreement about the appropriate outcome. Following Hart and Moore (2007) we suppose that each party feels entitled to the <u>best</u> feasible outcome for them. Of course, this means that typically at least one party, and usually both, will be aggrieved about the outcome that actually occurs. The only way to avoid or reduce aggrievement is to pick a tighter contract, i.e., a tighter interval of prices $[\underline{p}, \overline{p}]$.[2]

[2] We are assuming that a tight contract written at date 0 avoids argument and aggrievement, whereas a tight contract written at date 1 does not. Behind this assumption is the idea that there is a fundamental transformation between dates 0 and 1: at date 0 there is a (relatively) competitive market for buyers and sellers, whereas at date 1 the parties

We make the following assumptions:

(A1) If all helpful actions are taken at date 1, the value of the widget to B is v and the cost to S is c, where $v > c$. Hence net surplus = $v - c$ in this case.

(A2) If all the contractible helpful actions, but none of the noncontractible helpful actions, are taken at date 1, the value of the widget to B is $v - \frac{1}{2} \lambda (v - c)$ and the cost to S is $c + \frac{1}{2} \lambda (v - c)$, where $0 < \lambda < 1$. Hence net surplus = $(1 - \lambda)(v - c)$ in this case.

(A3) If none of the helpful actions (contractible or otherwise) is taken at date 1, B's value is very low and negative (approximately, $-\infty$) and S's cost is very high and positive (approximately, $+\infty$). In this case each party walks away from the contract (neither party has an incentive to enforce it) and no trade occurs.

Note that the import of (A2) is that withholding noncontractible helpful actions moves v and c in the direction of $\frac{1}{2}(v + c)$.

To understand the role of these assumptions, it is useful to distinguish between two cases. In the first the parties agree to one of the outcomes specified by the date 0 contract, i.e., a price $p \in [\underline{p}, \overline{p}]$. As part of this agreement the parties undertake the contractible helpful actions, but one or both parties may cut back on the noncontractible helpful actions to the extent that he is aggrieved about the price that is chosen.

---

have far fewer trading options (perhaps because they have made relationship-specific investments). We suppose that the date 0 market provides a relatively objective measure of what the parties bring to the relationship – there is little to argue about at this date – and that, once this information is embodied in a date 0 contract, the contract anchors and serves as a reference point for future entitlements. For further discussion, see Hart and Moore (2007).

In the second case, one party holds up the other party by threatening not to undertake any helpful actions, even the contractible ones, unless he receives a sidepayment. In effect he "coerces" the other party to renegotiate the contract. We assume that hold-up is viewed as an egregious act – it is a breach of the spirit of the date 0 contract – and causes the victim to withhold all cooperation, contractible and noncontractible, as a punishment. The result is a Nash equilibrium where neither party cooperates. This yields the no-trade outcome described in (A3). However, renegotiation is possible, as outlined below.

Let us elaborate on these two cases. Assume first that the parties stick to the date 0 contract (Case 1). Following Hart and Moore (2007), we suppose that each party feels entitled to the best feasible outcome permitted by the contract $[\underline{p}, \overline{p}]$ and that he is aggrieved to the extent that he gets less than this. In rough terms B feels entitled to pay $\underline{p}$ and S to receive $\overline{p}$ (we will be more precise below). We assume that, if B's payoff is \$a less than under the best feasible outcome, he punishes S by cutting back on noncontractible helpful actions to the point where S's payoff falls by $\$\theta a$, up to the maximum loss $\frac{1}{2} \lambda (v - c)$ that B can impose on S by withholding all noncontractible helpful actions (see (A2)). Similarly, if S's payoff is \$a less than under the best feasible outcome, she punishes B by cutting back on noncontractible helpful actions to the point where B's payoff is reduced by $\$\theta a$, up to the maximum loss $\frac{1}{2} \lambda (v - c)$ that S can impose on B. The parameter θ is exogenous and the same for B and S, and $0 < \theta \leq 1$.

Consider next the case where one party holds up the other (Case 2). As discussed above, this triggers the no-trade outcome described in (A3), yielding payoffs r_b, r_s for B, S, respectively. However, even if the parties hate each other they can still agree (contractually) to undertake some of the helpful actions at date 1. In other words renegotiation is possible. At the same time, since the relationship is poisoned, neither party will provide noncontractible cooperation in the future. In effect the parties agree to have a cold but correct relationship.[3] Renegotiation therefore yields surplus $(1 - \lambda)(v - c)$ by (A2). Thus if

(A4) $(1 - \lambda)(v - c) > r_b + r_s$,

the parties will renegotiate away from the no-trade outcome. We will assume that (A4) holds in what follows, but we will discuss below what happens if (A4) is relaxed.

Before analyzing the optimal contract, we put a little more structure on the random variables r_b, r_s. We suppose

(2.1) $r_b = \alpha_b + \beta_b v + \varphi + \gamma_b \varepsilon$,

(2.2) $r_s = \alpha_s - \beta_s c + \gamma_s \eta$.

Here α_b, β_b, γ_b, α_s, β_s and γ_s are constants, $1 > \beta_b > 0$, $1 > \beta_s > 0$, $\gamma_b > 0$, $\gamma_s > 0$ and φ, ε, η are independent random variables with mean zero. (2.1) – (2.2) captures the idea that B and S's outside options are positively correlated with v, c, respectively, but are also subject to exogenous noise (ε, η). The rationale for including the noise term φ in (2.1) will become clear in Section 3.

Given (2.1) – (2.2), we can represent the state of the world by the 5-tuple $\omega = (v, c, \varphi, \varepsilon, \eta)$. Both parties observe ω at date 1. Recall that a contract is a price interval $[\underline{p}, \overline{p}]$.[4] To

[3] We suppose that the parties cannot negotiate around this coldness. One could imagine that the buyer, anticipating that the seller is about to hold him up, would make a sidepayment to the seller to deflect the hold-up and preserve the relationship. We take the view that, given that there is a perceived threat in the background, the relationship is poisoned nonetheless.

[4] If the parties do not write a contract at all at date 0, we interpret this to mean that they have put no restrictions on p, i.e., $\underline{p} = -\infty$, $\overline{p} = \infty$. For other interpretations of "no contract," see Hart and Moore (2007).

simplify matters we suppose that B and S bargain within this interval and settle on a price half way between B's preferred price and S's preferred price.

To see what happens in state ω, it is useful to define $p_L(\omega)$ to be the price p such that S is indifferent between receiving p under the contract and holding B up; and $p_H(\omega)$ to be the price p such that B is indifferent between paying p under the contract and holding S up. As we have remarked, hold-up triggers the no-trade outcome described in (A3) as a threat point; but the parties then renegotiate to the outcome described in (A2). We suppose that the parties split the gains from negotiation 50 : 50. It follows that

$$p_L(\omega) - c = r_s + \tfrac{1}{2}\left((1-\lambda)(v-c) - r_b - r_s\right),$$

$$v - p_H(\omega) = r_b + \tfrac{1}{2}\left((1-\lambda)(v-c) - r_b - r_s\right).$$

Hence, from (2.1) – (2.2),

$$p_L(\omega) = \tfrac{1}{2}\left[\alpha_s + \gamma_s\eta - \alpha_b - \varphi - \gamma_b\varepsilon + ((1-\lambda) - \beta_b)v + ((1+\lambda) - \beta_s)c\right], \tag{2.3}$$

$$p_H(\omega) = \tfrac{1}{2}\left[\alpha_s + \gamma_s\eta - \alpha_b - \varphi - \gamma_b\varepsilon + ((1+\lambda) - \beta_b)v + ((1-\lambda) - \beta_s)c\right]. \tag{2.4}$$

Note that $p_H(\omega) > p_L(\omega)$. We may conclude that any price between $p_L(\omega)$ and $p_H(\omega)$ will avoid hold-up, while a price below $p_L(\omega)$ will trigger hold-up by S and a price above $p_H(\omega)$ will trigger hold-up by B.

Define $H(\omega, \underline{p}, \bar{p}) \equiv [p_L(\omega), p_H(\omega)] \cap [\underline{p}, \bar{p}]$. Either $H(\omega, \underline{p}, \bar{p}) \neq \phi$, in which case hold-up can be avoided, or $H(\omega, \underline{p}, \bar{p}) = \phi$, in which case it cannot be avoided. Examples of each case are illustrated in Figure 2.

$p_L(\omega)$	$p_H(\omega)$	$\underline{p}$	$\bar{p}$

Figure 2(a)

$p_L(\omega)$	$\underline{p}$	$p_H(\omega)$	$\bar{p}$

Figure 2(b)

$\underline{p}$	$p_L(\omega)$	$\bar{p}$	$p_H(\omega)$

Figure 2(c)

Figure 2(d)

In Figure 2(a), contractual prices are too high and B will hold up S ($p > p_H(\omega)$). In Figure 2(d), contractual prices are too low and S will hold up B ($p < p_L(\omega)$). In Figure 2(b) the parties can avoid hold up by choosing $\underline{p} \le p \le p_H(\omega)$. In Figure 2(c) the parties can avoid hold up by choosing $p_L(\omega) \le p \le \bar{p}$.

Although hold-up is avoided in Figures 2(b) and 2(c), there is aggrievement. Our assumption is that each party feels entitled to the best outcome permitted by the contract. However, each party recognizes that he faces the feasibility constraint that the other party can trigger hold-up, i.e., B can't expect to pay less than $p_L(\omega)$ or S to receive more than $p_H(\omega)$. In other words B feels entitled to $p = \text{Max}(p_L(\omega), \underline{p})$ and S feels entitled to $p = \text{Min}(p_H(\omega), \bar{p})$. Note that the assumption that entitlements are constrained by what's feasible simplifies the analysis but is not crucial (a similar assumption is made in Hart and Moore (2007)).

Thus in Case 1 aggregate aggrievement is given by $\text{Min}(p_H(\omega), \bar{p}) - \text{Max}(p_L(\omega), \underline{p})$, and net surplus by

$$(2.5) \quad W_1(\omega, \underline{p}, \bar{p}) = v - c - \theta\,\{\text{Min}(p_H(\omega), \bar{p}) - \text{Max}(p_L(\omega), \underline{p})\}.$$

In contrast, in Case 2 where $H(\omega, \underline{p}, \overline{p}) = \phi$, hold-up occurs, followed by renegotiation, and net surplus is given by

$$(2.6) \quad W_2(\omega) = (1 - \lambda)(v - c).$$

It is easy to see that

$$(2.7) \quad W_1(\omega, \underline{p}, \overline{p}) \geq W_2(\omega).$$

This follows from the fact that $0 < \theta \leq 1$ and

$$(2.8) \qquad \begin{aligned} \text{Min}\,(p_H(\omega), \overline{p}) - \text{Max}\,(p_L(\omega), \underline{p}) & \\ &\leq p_H(\omega) - p_L(\omega) \\ &= \lambda(v - c), \end{aligned}$$

where we are using (2.3) – (2.4). In other words, however large the price range $[\underline{p}, \overline{p}]$ is, net surplus is higher if hold-up is avoided than if it occurs. Note that, given that B and S choose a price midway between $\text{Min}\,(p_H(\omega), \overline{p})$ and $\text{Max}\,(p_L(\omega), \underline{p})$, (2.8) implies that the equilibrium losses B and S impose on each other are no greater than ½ $\lambda(v - c)$, which is consistent with (A2).

Since date 0 lump-sum transfers can be used to reallocate surplus, an optimal contract maximizes expected net surplus. Thus an optimal contract solves:

$$\text{(2.9)} \qquad \underset{\underline{p},\ \overline{p}}{\text{Max}} \left\{ \int_{H(\omega,\underline{p},\overline{p}) \neq \phi} W_1\ \omega, \underline{p}, \overline{p})\ dF(\omega) + \int_{H(\omega,\underline{p},\overline{p}) = \phi} W_2(\omega)\ dF(\omega) \right\},$$

where F is the distribution function of ω. We assume that F has bounded support. It is then easy to show that an optimal contract exists since the objective function is upper semicontinuous in [$\underline{p}$, $\overline{p}$]. The trade-off is the following. As $\underline{p}$ falls or $\overline{p}$ rises, the set H becomes larger and so it is more likely that hold-up can be avoided. This is good given that $W_1 \geq W_2$. However, shading represented by θ $\{\text{Min}\ (p_H\ (w), \overline{p}) - \text{Max}\ (p_L\ (w), \underline{p})\}$ rises, which means that W_1 falls, which is bad.

It is worth considering in which states hold-up will occur. We have seen in Figure 2 that B will hold up S when p_H (ω) is low and S will hold up B when p_L (ω) is high. From (2.3) – (2.4), the former occurs when v is low or c is low, and the latter occurs when v is high or c is high, as long as $\beta_b < 1 - \lambda$, $\beta_s < 1 - \lambda$. In words, if r_b varies positively but not too strongly with v, and r_s varies positively but not too strongly with c, B is vulnerable to hold up when v is high or c is high. Conversely for S.

It is useful to compute the optimal contract in some simple cases. Suppose that there is no uncertainty: $\omega = \omega_1$. Then the first-best can be achieved by selecting a single trading price $\hat{p}$ in the range [p_L (ω_1), p_H (ω_1)]. In other words set $\underline{p} = \hat{p} = \overline{p}$. There is then nothing to argue about, i.e., no aggrievement, and neither B nor S has an incentive to engage in hold-up.

It turns out that the first-best can also be achieved if there are just two states: $\omega = \omega_1$ or ω_2. To see why, note that there are two possibilities. Either [p_L (ω_1), p_H (ω_1)] $\cap$ [p_L (ω_2), p_H

$(\omega_2)]$ is non-empty or it is empty. In the first case, choose any price $\hat{p}$ in the intersection and set $\underline{p} = \hat{p} = \overline{p}$. In the second case, suppose without loss of generality that

$$p_L(\omega_1) < p_H(\omega_1) < p_L(\omega_2) < p_H(\omega_2).$$

Then set $\underline{p} = p_H(\omega_1)$, $\overline{p} = p_L(\omega_2)$. In state ω_1, $p = p_H(\omega_1)$ and in state ω_2 $p = p_L(\omega_2)$. Hold-up is avoided in both states and there is no aggrievement by S in ω_1 (since in this state $p_H(\omega_1)$ is the highest price that avoids hold-up by B); and no aggrievement by B in ω_2 (since in this state $p_L(\omega_2)$ is the lowest price that avoids hold-up by S).

We now present a three-state example in which the first-best cannot be achieved.

<u>Example 1</u>

Suppose that there are three states, ω_1, ω_2, ω_3, with probabilities π_1, π_2, π_3, respectively. Only v varies across the states: $v(\omega_1) = 20$, $v(\omega_2) = 60$, $v(\omega_3) = 80$. Assume $c = 10$, $r_b = r_s = 0$, $\lambda = \frac{1}{2}$. Then

$$p_L(\omega) = \tfrac{1}{4} v + \tfrac{3}{4} c,$$

$$p_H(\omega) = \tfrac{3}{4} v + \tfrac{1}{4} c.$$

The $[p_L(\omega), p_H(\omega)]$ intervals are illustrated in Figure 3.

12.5 17.5 22.5 27.5 47.5 62.5

|____________|___________|____________|_________________|____________|

$p_L(\omega_1)$ $p_H(\omega_1)$ $p_L(\omega_2)$ $p_L(\omega_3)$ $p_H(\omega_2)$ $p_H(\omega_3)$

Figure 3

We see that the $[p_L(\omega_2), p_H(\omega_2)]$, $[p_L(\omega_3), p_H(\omega_3)]$ intervals overlap, but neither overlaps with $[p_L(\omega_1), p_H(\omega_1)]$. There are three candidates for an optimal contract: one can avoid hold-up in all states with some aggrievement; one can avoid hold-up in ω_1 and ω_2 without aggrievement; or one can avoid hold-up in ω_2 and ω_3 without aggrievement.

Contract 1: Avoiding hold-up in all states

To avoid hold-up in all states, we need $\underline{p} \leq p_H(\omega_1)$, $\overline{p} \geq p_L(\omega_3)$. To minimize shading costs, we want the highest $\underline{p}$ and lowest $\overline{p}$. Hence set $\underline{p} = 17.5$, $\overline{p} = 27.5$. In ω_1, $p = 17.5$ and there is no aggrievement. In ω_2 there is aggrievement of 5 since B would like $p = 22.5$ and S would like $p = 27.5$. In ω_3, $p = 27.5$ and there is no aggrievement. Net surplus is given by

$$W = 10\,\pi_1 + (50 - 5\theta)\,\pi_2 + 70\,\pi_3.$$

Contract 2: Avoiding hold-up in ω_1, ω_2

In this case it is best to set $\underline{p} = p_H(\omega_1) = 17.5$, $\overline{p} = p_L(\omega_2) = 22.5$. This avoids aggrievement in states ω_1, ω_2 (see the above discussion of the two-state example). However, hold-up occurs in ω_3. Net surplus is given by

$$W = 10\,\pi_1 + 50\,\pi_2 + 35\,\pi_3.$$

Contract 3: Avoiding hold-up in ω_2, ω_3

In this case it is best to set $\underline{p} = \overline{p} = \hat{p}$, where $27.5 \le \hat{p} \le 47.5$. This avoids aggrievement in ω_2, ω_3. However, hold-up occurs in ω_1. Net surplus is given by

$$W = 5\,\pi_1 + 50\,\pi_2 + 70\,\pi_3.$$

Which of the above contracts is best depends on the parameters π_1, π_2, π_3, and θ. Contract 1 is optimal if θ or π_2 is small. Contract 2 is optimal if π_3 is small. Contract 3 is optimal if π_1 is small.

3. Asset Ownership

In this section we explore the idea that asset ownership can improve the parties' contractual relationship. We take a simple view of asset ownership. Asset ownership matters because it determines which assets each party can walk away with if trade does not occur.[5] This in turn affects parties' outside options and their incentives to engage in hold-up.

Let A_b be the set of assets B owns and A_s the set of assets S owns. We suppose

$$(3.1) \quad A_b \cap A_s = \phi,\ A_b \cup A_s \subseteq A,$$

[5] As in Hart and Moore (1990). See also Grossman and Hart (1986).

where A is the set of all assets (assumed finite). The first part of (3.1) says that B and S can't walk away with the same asset. The inclusion in the second part reflects the possibility that if an asset is jointly owned neither party can walk away with it.

We now suppose that the coefficients α_b, β_b, γ_b, α_s, β_s, γ_s depend on asset ownership. In particular, $\alpha_b = \alpha_b\,(A_b)$, $\beta_b = \beta_b\,(A_b)$, $\gamma_b = \gamma_b\,(A_b)$, $\alpha_s = \alpha_s\,(A_s)$, $\beta_s = \beta_s\,(A_s)$, $\gamma_s = \gamma_s\,(A_s)$. We also make assumptions similar to those in the property rights literature (see Grossman and Hart (1986) and Hart and Moore (1990)) about how these coefficients vary with asset ownership. In particular, we assume that owning more assets increases both <u>total</u> and <u>marginal</u> payoffs of r_b, r_s, with respect to v and c. That is,

(3.2) α_b and β_b are nondecreasing in A_b ,

(3.3) α_s and β_s are nondecreasing in A_s.

We also make an assumption that we came across in Section 2:

(3.4) $\beta_b\,(A) < 1 - \lambda$, $\beta_s\,(A) < 1 - \lambda$.

(3.4) ensures that $p_L\,(\omega)$ is increasing in v and $p_H\,(\omega)$ is increasing in c. Finally, we suppose that (A4) holds for all ownership structures.

A contract is now a 4-tuple (A_b, A_s, $\underline{p}$, $\overline{p}$), specifying an asset ownership allocation (A_b, A_s) and a price interval [$\underline{p}$, $\overline{p}$], where A_b, A_s satisfy (3.1). An optimal contract solves:

$$\text{(3.5)} \qquad \max_{(A_b, A_s, \underline{p}, \overline{p})} \quad \int\limits_{H(\omega, A_b, A_s, \underline{p}, \overline{p}) \neq \phi} W_1\,(\omega, A_b, A_s, \underline{p}, \overline{p})\, dF(\omega) + \int\limits_{H(\omega, A_b, A_s, \underline{p}, \overline{p}) = \phi} W_2\,(\omega)\, dF(\omega)\ ,$$

where W_1 and H are now indexed by the asset ownership allocation since (A_b, A_s) affect p_L and p_H.

We begin our analysis of asset ownership with a simple observation. If β_b, γ_b are independent of A_b and β_s, γ_s are independent of A_s, asset ownership doesn't matter. To see this note that under these conditions a change in asset ownership changes p_L and p_H by the same constant (via α_b, α_s). If $\underline{p}$, $\overline{p}$ are also changed by this constant, everything is as before. Thus for asset ownership to matter the marginal terms β_b, γ_b, β_s, γ_s must vary with A_b, A_s. The most interesting parameters are the β's and we begin by focusing on these.

Suppose assets are transferred at date 0 from S to B. Then, given (3.2) – (3.3), β_b rises and β_s falls. As is clear from (2.3) – (2.4), this makes p_L (ω) and p_H (ω) less sensitive to v than before since

$$\frac{\partial p_L}{\partial v}\ (\omega) = ½\ ((1 - \lambda) - \beta_b\ ),$$

$$\frac{\partial p_H}{\partial v}\ (\omega) = ½\ ((1 + \lambda) - \beta_b\ ),$$

and these both decrease. On the other hand, p_L (ω) and p_H (ω) become more sensitive to c since

$$\frac{\partial p_L}{\partial c}(\omega) = \tfrac{1}{2}\left((1+\lambda) - \beta_s\right),$$

$$\frac{\partial p_H}{\partial c}(\omega) = \tfrac{1}{2}\left((1-\lambda) - \beta_s\right),$$

and these both increase.

Intuitively, a reduction in sensitivity is good because it makes it more likely that the interval $[p_L(\omega), p_H(\omega)]$ will intersect with the fixed interval $[\underline{p}, \overline{p}]$. In other words hold-up is less likely, and shading costs are also likely to fall. Conversely with an increase in sensitivity. The suggested conclusion is that, if the variation in v is large relative to the variation in c, assigning assets to B is a good idea; whereas, if the variation in c is large relative to the variation in v, assigning assets to S is a good idea.

We formalize this in the next proposition. For simplicity the proposition focuses on the case where either only v varies or only c varies.

<u>Proposition 1</u>. (1) Suppose that $\varphi = \varepsilon = \eta \equiv 0$, and $c \equiv c_0$ where c_0 is a constant. Then it is optimal for B to own all the assets, i.e., $A_b = A$, $A_s = \phi$.

(2) Suppose that $\varphi = \varepsilon = \eta \equiv 0$ and $v \equiv v_0$, where v_0 is a constant. Then it is optimal for S to own all the assets, i.e., $A_s = A$, $A_b = \phi$.

<u>Proof</u>. We prove (1). Let $(A_b, A_s, [\underline{p}, \overline{p}])$ be an optimal contract. The proof proceeds in two steps. We first replace $[\underline{p}, \overline{p}]$ by another price interval, and show that shading costs fall. We

then allocate all the assets to B and make another change in the price interval, and show that shading costs fall again. Finally, the hold-up region becomes (weakly) smaller.

Index the state by v. Let $\underline{v}$ be the smallest value of v, and $\overline{v}$ the largest of v, in the support of F such that no hold-up occurs under contract (A_b, A_s, $[\underline{p}, \overline{p}]$). Then

$$(3.6) \qquad [p_L(v), p_H(v)] \cap [\underline{p}, \overline{p}] \neq \phi$$

for $v = \underline{v}$, $v = \overline{v}$. Since $p_L(v)$, $p_H(v)$ are increasing in v, (3.6) must also hold for $\underline{v} \leq v \leq \overline{v}$, i.e., hold-up does not occur for intermediate v's. Note that (3.6) implies that $p_H(\underline{v}) \geq \underline{p}$, $p_L(\overline{v}) \leq \overline{p}$.

Now define a new price interval $[\underline{p}', \overline{p}']$, where

$$\underline{p}' = p_H(\underline{v}) \text{ and } \overline{p}' = \text{Max}\,(p_L(\overline{v}), p_H(\underline{v})).$$

Clearly $\underline{p}' \geq \underline{p}$. Also either $\overline{p}' \leq \overline{p}$ or $\overline{p}' = \underline{p}'$. In the first case the new price interval is a subset of the previous price interval. In the second case it is a singleton. In both cases

$$[p_L(v), p_H(v)] \cap [\underline{p}', \overline{p}'] \neq \phi$$

for $\underline{v} \leq v \leq \overline{v}$. Hence the new price interval avoids hold-up for $\underline{v} \leq v \leq \overline{v}$ just like the old one. In addition aggrievement and shading costs are lower under the new price interval given that either the new price interval is a subset of the previous price interval or it is a singleton (in which case shading costs are zero).

Now assign all the assets to B, i.e., set $A_b = A$, $A_s = \phi$. Call this the new ownership structure. Define a new price interval $[\underline{p}'', \overline{p}'']$, given by

$$(3.7) \qquad \underline{p}'' = p_H^N(\underline{v}), \ \overline{p}'' = \text{Max}\,(p_L^N(\overline{v}), p_H^N(\underline{v})),$$

where p_L^N, p_H^N represent the values of p_L, p_H under the new ownership structure. The price interval $[\underline{p}'', \overline{p}'']$ avoids hold-up under the new ownership structure when $\underline{v} \leq v \leq \overline{v}$.

We show next that shading costs are lower for each $\underline{v} \leq v \leq \overline{v}$ under the new ownership structure and price interval $[\underline{p}'', \overline{p}'']$ than under the old ownership structure and $[\underline{p}', \overline{p}']$ (which in turn are lower than those under the old ownership structure and $[\underline{p}, \overline{p}]$). That is, we demonstrate that

$$(3.8) \qquad \text{Min}\,(p_H^N(v), \ \overline{p}'') - \text{Max}\,(p_L^N(v), \underline{p}'')$$
$$\leq \text{Min}\,(p_H(v), \ \overline{p}') - \text{Max}\,(p_L(v), \underline{p}').$$

There are several cases to consider. Note first that if $\underline{p}', \overline{p}' = p_H(\underline{v}) \geq p_L(\overline{v})$, i.e., the right-hand side (RHS) of (3.8) is zero, then

$$
\begin{aligned}
p_H^N(\underline{v}) - p_L^N(\bar{v}) &= -\left(p_H^N(\bar{v}) - p_H^N(\underline{v})\right) \\
&\qquad + p_H^N(\bar{v}) - p_L^N(\bar{v}) \\
&\geq -\left(p_H(\bar{v}) - p_H(\underline{v})\right) \\
&\qquad + p_H(\bar{v}) - p_L(\bar{v}) \\
&= p_H(\underline{v}) - p_L(\bar{v}) \\
&\geq 0,
\end{aligned}
$$

where we are using the fact that $p_H^N(\bar{v}) - p_H^N(\underline{v}) \leq p_H(\bar{v}) - p_H(\underline{v})$ since $\frac{\partial p_H}{\partial v}$ falls the more assets B owns (by (2.4) and (3.2)), and $p_H^N(\bar{v}) - p_L^N(\bar{v}) = p_H(\bar{v}) - p_L(\bar{v})$, i.e., $p_H - p_L$ is independent of the ownership structure (see (2.3) – (2.4)). Hence $p_H^N(\underline{v}) \geq p_L^N(\bar{v})$. It follows from (3.7) that $\underline{p}'' = \bar{p}'' = p_H^N(\underline{v})$ and so the left-hand side (LHS) of (3.8) is zero. Therefore (3.8) holds.

Consider next the case where $\underline{p}' = p_H(\underline{v}) < \bar{p}' = p_L(\bar{v})$. If $\underline{p}'' = \bar{p}'' = p_H^N(\underline{v}) \geq p_L^N(\bar{v})$, (3.8) again holds. So suppose $\underline{p}'' = p_H^N(\underline{v}) < p_L^N(\bar{v}) = \bar{p}''$. We must show that

$$
\begin{aligned}
(3.9)\qquad & \text{Min}\left(p_H^N(v),\ p_L^N(\bar{v})\right) - \text{Max}\left(p_L^N(v), p_H^N(\underline{v})\right) \\
& \leq \text{Min}\left(p_H(v), p_L(\bar{v})\right) - \text{Max}\left(p_L(v), p_H(\underline{v})\right).
\end{aligned}
$$

We can rewrite (3.9) as

$$
(3.10)\quad \text{Min}\left\{\begin{matrix} p_H^N(v)-p_H^N(\underline{v}), p_L^N(\bar{v})-p_H^N(\underline{v}), \\ p_H^N(v)-p_L^N(v), p_L^N(\bar{v})-p_L^N(v)\end{matrix}\right\} \leq \text{Min}\left\{\begin{matrix} p_H(v)-p_H(\underline{v}), p_L(\bar{v})-p_H(\underline{v}), \\ p_H(v)-p_L(v), p_L(\bar{v})-p_L(v)\end{matrix}\right\}.
$$

To establish (3.10) one shows that each component in the min formula on the LHS of (3.10) is no greater than the corresponding component on the RHS of (3.10). This follows from the fact that $\frac{\partial p_H}{\partial v}, \frac{\partial p_L}{\partial v}$ are nonincreasing in the assets that B owns, and that $p_H - p_L$ is independent of ownership structure for a given v. Hence (3.9) holds and so does (3.8).

In summary, the new ownership structure (in which B owns all the assets) and price range $[\underline{p}'', \overline{p}'']$ yields (weakly) lower shading costs than the original ownership structure and price range $[\underline{p}, \overline{p}]$. Also the hold-up region is no larger (hold-up does not occur for $\underline{v} \leq v \leq \overline{v}$). This shows that allocating all the assets to B is optimal. Q.E.D.

Proposition 1 can be summarized as follows. If v varies a lot relative to c, then B should own all the assets because this makes p_L and p_H vary less with v. As a result, for a given price range, it is less likely that v will be so high that $p_L(v)$ will exceed $\overline{p}$, causing S to hold up B; or that v will be so low that $p_H(v)$ will fall below $\underline{p}$, causing B to hold up S. Conversely, if c varies a lot relative to v, then S should own all the assets.

Proposition 1 is reminiscent of the result in the property rights literature that says that one party should own all the assets if his investment is important. The proposition is restrictive in that it deals only with the case where one party's payoff varies. Unfortunately, the general case is hard to analyze. To make progress we simplify matters considerably. We suppose that with high probability v, c take on "normal" values $v = v_0$, $c = c_0$, while with small probability v, c can each take on an "exceptional" value. Since exceptional values are unusual, we ignore the possibility that v and c can take on an exceptional value at the same time. We also suppose that there is a small amount of exogenous noise through the random variable φ, but we set $\varepsilon = \eta = 0$.

Define an asset to be idiosyncratic to B if owning it increases the sensitivity of r_b to v and not owning it has no effect on the sensitivity of r_s to c. Define an asset to be idiosyncratic to S similarly.

Definition (i) Asset a is idiosyncratic to B if $\beta_b(A_b \cup \{a\}) > \beta_b(A_b)$ for all $A_b \subseteq A$, $A_b \cap \{a\} = \phi$, and $\beta_s(A_s \cup \{a\}) = \beta_s(A_s)$ for all $A_s \subseteq A$.

(ii) Asset a is idiosyncratic to S if $\beta_s(A_s \cup \{a\}) > \beta_s(A_s)$ for all $A_s \subseteq A$, $A_s \cap \{a\} = \phi$, and $\beta_b(A_b \cup \{a\}) = \beta_b(A_b)$ for all $A_b \subseteq A$.

The next proposition says that an asset that is idiosyncratic to a party should be owned by that party.

Proposition 2. Assume that $\varepsilon = \eta = 0$. Suppose that with probability $0 < \pi < 1$ event 1 occurs: $v = v_0$, $c = c_0$ and φ is uniformly distributed on [-k, k]; with probability $(1 - \pi)\,\alpha_v$ event 2 occurs: $c = c_0$, v has support $[v_L, v_H]$, where $v_L \leq v_0$ and $v_H \geq \dfrac{-\beta_b(A)v_0 + (1+\lambda)v_0 - 2\lambda c_0}{1-\lambda-\beta_b(A)}$, and φ is uniformly distributed on [-k, k]; with probability $(1 - \pi)\,\alpha_c$ event 3 occurs: $v = v_0$, c has support $[c_L, c_H]$, where $c_H \geq c_0$ and $c_L \leq \dfrac{-\beta_s(A)c_0 - 2\lambda v_0 + (1+\lambda)c_0}{1-\lambda-\beta_s(A)}$, and φ is uniformly distributed on [-k, k]. Here $\alpha_v > 0$, $\alpha_c > 0$, $\alpha_v + \alpha_c = 1$, $k > 0$, and φ is independent of v and c in events 2 and 3 respectively. Then for small enough k the following is true: if π is close to 1 it is uniquely optimal for B to own asset a if a is idiosyncratic to B and for S to own asset a if a is idiosyncratic to S.

Proof. Suppose a is idiosyncratic to B. We show that B should own a. The proof is by contradiction. If the proposition is false, then, however small k is, we can construct a sequence of optimal contracts $(A_{br}, A_{sr}, \underline{p}_r, \overline{p}_r)$ such that $a \in A_{sr}$ for all r, i.e., S owns asset a, and $\pi_r \rightarrow 1$ as $r \rightarrow \infty$. Without loss of generality (wlog) suppose that $A_{br} \rightarrow A_b(k)$, $A_{sr} \rightarrow A_s(k)$, $\underline{p}_r \rightarrow \underline{p}(k)$, $\overline{p}_r \rightarrow \overline{p}(k)$. Then $A_b(k)$, $A_s(k)$, $\underline{p}(k)$, $\overline{p}(k)$ must be optimal for the case where event 1 occurs with probability 1. For small k the first-best can be achieved (exactly) in event 1 since there is almost no uncertainty. A necessary condition for this is that there is a single trading price p(k) in the limit, i.e., $\underline{p}(k) = p(k) = \overline{p}(k)$ (so that shading costs are zero), and

$$(3.11) \qquad p_L(\omega, A_b(k), A_s(k)) \leq \underline{p}(k) = p(k) = \overline{p}(k) \leq p_H(\omega, A_b(k), A_s(k))$$

for all $-k \leq \varphi \leq k$, where $\omega = (v_0, c_0, \varphi)$ and we now suppress $\varepsilon = 0$, $\eta = 0$. Here p_L, p_H are as in (2.3) – 2.4) and are indexed by the limiting ownership structure.

Consider a new sequence of contracts $(A_{br}', A_{sr}', \underline{p}_r', \overline{p}_r')$, where the only difference between A_{br}', A_{sr}' and A_{br}, A_{sr} is that asset a is transferred to B, and

$$(3.12) \qquad \begin{aligned} \underline{p}_r' - \underline{p}_r = \overline{p}_r' - \overline{p}_r &= \tfrac{1}{2}\,[\alpha_s(A_{sr} \setminus \{a\}) - \alpha_s(A_{sr}) \\ &\qquad + \alpha_b(A_{br}) - \alpha_b(A_{br} \cup \{a\}) \\ &\qquad + \beta_b(A_{br})\, v_0 - \beta_b(A_{br} \cup \{a\})\, v_0] \\ &= \Delta_r(k). \end{aligned}$$

In other words we adjust $\underline{p}_r$, $\overline{p}_r$ by an amount equal to the change Δp_L, Δp_H in p_L, p_H that occurs as a result of the shift in ownership structure, where Δp_L, Δp_H are evaluated at $v = v_0$. Note that, given the assumption that a is idiosyncratic to B, Δp_L, Δp_H depend on v, but not on c (or φ).

What happens to expected net surplus as a result of this change? Expected net surplus is a weighted average of surplus in the three events 1, 2, 3. Given that $\underline{p}_r$, $\overline{p}_r$, p_{Lr}, (ω) p_{Hr} (ω) all shift by Δ_r when $v = v_0$, nothing changes in events 1 and 3 for all r. That is, for each state ω, hold-up occurs if and only if it did before, and the level of shading costs, if hold-up doesn't occur, remains constant. Thus net surplus is unchanged in events 1 and 3.

Since the new contract cannot deliver higher expected net surplus than the original contract, given that the original contract is optimal, it follows that net surplus must be (weakly) lower in event 2. Let $r \to \infty$. Wlog $(A_{br}', A_{sr}', \underline{p}_r', \overline{p}_r') \to (A_b'(k), A_s'(k), \underline{p}'(k), \overline{p}'(k))$ and $\Delta_r(k) \to \Delta(k)$. Given (3.11) – (3.12), we must have

$$(3.13) \quad p_L(\omega, A_b'(k), A_s'(k)) \leq \underline{p}'(k) = p'(k) = \overline{p}'(k) \leq p_H(\omega, A_b'(k), A_s'(k))$$

for all $-k \leq \varphi \leq k$, where $\omega = (v_0, c_0, \varphi)$. By the above arguments the contract $(\underline{p}'(k), \overline{p}'(k), A_b'(k), A_s'(k))$ delivers surplus no higher in event 2 than the contract $(\underline{p}(k), \overline{p}(k), A_b(k), A_s(k))$.

We show that this conclusion is false. Since the primed and unprimed contracts both have a single trading price ($p'(k)$, $p(k)$, respectively), shading costs are zero in both contracts. We demonstrate that there is less hold-up in the primed contract. Since we are in event 2 index the state by (v, φ). Then, from (3.12),

$$(3.14)\quad p'(k) - p(k) = p_L((v_0, \varphi), A_b'(k), A_s'(k)) - p_L((v_0, \varphi), A_b(k), A_s(k))$$
$$= p_H((v_0, \varphi), A_b'(k), A_s'(k)) - p_H((v_0, \varphi), A_b(k), A_s(k))$$
$$= \Delta(k)$$

for all φ. Now hold-up occurs in the primed contract in state (v, φ) if and only if either $p'(k) < p_L((v, \varphi), A_b'(k), A_s'(k))$ or $p'(k) > p_H((v, \varphi), A_b'(k), A_s'(k))$. Consider the first. Given (3.13) and the fact that p_L is increasing in v, $p'(k) < p_L((v, \varphi), A_b'(k), A_s'(k))$ only if $v > v_0$. But, if $v > v_0$,

$$(3.15)\quad p_L((v, \varphi), A_b'(k), A_s'(k)) - p_L((v_0, \varphi), A_b'(k), A_s'(k))$$
$$= \tfrac{1}{2}[(1 - \lambda) - \beta_b(A_b \cup \{a\})](v - v_0)$$
$$< \tfrac{1}{2}[(1 - \lambda) - \beta_b(A_b)](v - v_0)$$
$$= p_L((v, \varphi), A_b(k), A_s(k))$$
$$- p_L((v_0, \varphi), A_b(k), A_s(k)),$$

since a is idiosyncratic to B. From (3.14) – (3.15), we may conclude that $p'(k) < p_L((v, \varphi), A_b'(k), A_s'(k)) \Rightarrow p(k) < p_L((v, \varphi), A_b(k), A_s(k))$, i.e., hold-up occurs in the unprimed contract if it occurs in the primed contract. A similar argument shows that $p'(k) > p_H((v, \varphi), A_b'(k), A_s'(k)) \Rightarrow p(k) > p_H((v, \varphi), A_b(k), A_s(k))$. Putting the two arguments together, we may conclude that hold-up costs are weakly lower in the primed contract than the unprimed one. In fact, they are strictly lower: this follows from the assumption about the support of v in Proposition 2, which ensures that $p_L((v, \varphi), A_b(k), A_s(k)) > p_H((v, \varphi), A_b(k), A_s(k))$ for large v and φ close to zero (i.e., hold-up does occur sometimes), but not for v close to v_0 (i.e., hold-up does not always occur). Contradiction. Q.E.D.

The intuition behind Proposition 2 is simple. If an asset is idiosyncratic to B, say, then transferring it to B makes B's payoff relative to his outside option less sensitive to exogenous events, without making S's payoff relative to her outside option more sensitive. This makes it easier to avoid hold-up.

A simple application of Proposition 2 is to the case of strictly complementary assets. Suppose assets a_1 and a_2 are strictly complementary. Then a_2 by itself is of no use to S, while a_1 and a_2 together may be very useful to B. Assume B owns a_1. Then we can define a new economy in which a_1 is <u>inalienable</u>, i.e., B always owns a_1, and the effective set of (alienable) assets is $A \setminus \{a_1\}$. For this economy, a_2 is idiosyncratic to B in the sense of Definition (i). Hence, according to Proposition 2, it is better for B to own $a_1$and a_2. The same argument shows that if S owns a_1, it is better for S to own both. The conclusion is that strictly complementary assets should be owned together (by B or S – without further information we cannot say which). A similar argument shows that joint ownership is suboptimal under the conditions of Proposition 2.

Of course, these results are very reminiscent of those obtained in the property rights literature (see particularly Hart and Moore (1990)). However, the driving force is different: uncertainty rather than ex ante investments.

So far we have emphasized the idea that ownership of an asset is good for one party because it reduces the variability of their payoff relative to their outside option. However, there is also a class of cases where ownership can increase variability and it may be better to take assets away from people. The next proposition describes a situation where it is better to take assets away from both parties, i.e., joint ownership is optimal. In this proposition v and c are constant while φ and ε or η vary.

Proposition 3. Assume γ_b, γ_s are strictly increasing in A_b, A_s respectively. Suppose that with probability $0 < \pi < 1$ event 1 occurs: $v = v_0$, $c = c_0$, $\varepsilon = 0$, $\eta = 0$ and φ is uniformly distributed on $[-k, k]$; with probability $(1-\pi)\,\alpha_\varepsilon$ event 2 occurs: $v = v_0$, $c = c_0$, $\eta = 0$, ε has support $[\varepsilon_L, \varepsilon_H]$, where $\varepsilon_L \leq \frac{2\lambda(c_0 - v_0)}{\gamma_b(\phi)}$ and $\varepsilon_H > 0$, and φ is uniformly distributed on $[-k, k]$; with probability $(1-\pi)\,\alpha_\eta$ event 3 occurs: $v = v_0$, $c = c_0$, $\varepsilon = 0$, η has support $[\eta_L, \eta_H]$, where $\eta_L \leq \frac{2\lambda(c_0 - v_0)}{\gamma_s(\phi)}$ and $\eta_H > 0$, and φ is uniformly distributed on $[-k, k]$. Here $\alpha_\varepsilon \geq 0$, $\alpha_\eta \geq 0$, $\alpha_\varepsilon + \alpha_\eta = 1$, $k > 0$, and φ is independent of ε and η in events 2 and 3 respectively. Then for small enough k the following is true: if π is close to 1 it is uniquely optimal for all assets to be jointly owned by B and S.

Proof. We sketch the proof since the argument is very similar to that of Proposition 2. Suppose joint ownership is not optimal. For small k choose a sequence of optimal contracts as $\pi \to 1$. The limiting contract is optimal for event 1. Hence (3.11) is satisfied. Consider a new sequence of contracts where all assets are jointly owned and $\underline{p}_r$, $\overline{p}_r$ are adjusted to reflect the new ownership structure, i.e.,

$$
\begin{aligned}
\underline{p}_r' - \underline{p}_r = \overline{p}_r' - \overline{p}_r \\
&= \tfrac{1}{2}\,[\alpha_s(\phi) - \alpha_s(A_{sr}) - \alpha_b(\phi) + \alpha_b(A_{br}) \\
&\quad - \beta_s(\phi)\,c_0 + \beta_s(A_{sr})\,c_0 - \beta_b(\phi)\,v_0 + \beta_b(A_{br})\,v_0].
\end{aligned}
$$

Then surplus does not change in event 1. Since the initial contract is optimal surplus must weakly fall in events 2 or 3. Wlog suppose it falls in event 2. Take limits as $r \to \infty$. The limiting joint ownership contract has the property that p_L, p_H vary less with ε than under the original contract. But this makes hold-up less likely. Hence the joint ownership contract creates higher net surplus. Contradiction. Q.E.D.

Again similar results have been obtained in the property rights literature (see, e.g., Halonen (2002) and Rajan and Zingales (1998)). Note that Proposition 3 depends on the assumption that owning more assets increases the variance of the outside option. Although plausible, one can certainly imagine other possibilities. For this reason we are inclined to put less weight on Proposition 3 than on Propositions 1 and 2.

Let us conclude by discussing what happens if (A4) does not hold. In this case, hold-up leads to the end of the relationship: the parties prefer to go their separate ways than trade under hostile terms. Thus hold-up is equivalent to quitting, $p_L = c + r_s$, $p_H = v - r_b$, and the model becomes similar to that of Hart and Moore (2007, Section 3). Asset ownership still has a role to play to the extent that it reduces the variability of p_L and p_H with respect to c and v.

The above discussion ignores a subtlety. If (A4) does not hold, and the asset ownership structure is inefficient given that the parties are going their separate ways, one might expect them to renegotiate the structure (as in Baker et al. (2002)). This leads to interesting new hold-up possibilities. We leave an analysis of these to future work.

4. Conclusions

In this paper we have studied the trade-off between contractual flexibility and rigidity in a buyer-seller relationship. A flexible contract is good because it allows the parties to adjust the terms of trade to uncertain events, but bad because it leads to argument, aggrievement and shading. A rigid contract avoids argument in normal times but has the undesirable feature that if value or cost falls outside the normal range one party will attempt to force renegotiation on the other, damaging the relationship and causing deadweight losses. We have shown that a judicious allocation of asset ownership can improve this trade-off. For example, allocating assets to the buyer will cause the value of the buyer's trade inside the relationship and his value outside the relationship to move together in such a way that the seller's ability to engage in hold-up is reduced. Thus, if the buyer's value varies a lot relative to the seller's cost, the range of parameters over which hold-up is avoided is expanded.

It is useful to compare our analysis to that found in the property rights literature (see Grossman and Hart (1986) and Hart and Moore (1990)) and the transaction cost literature (see Williamson (1985) and Klein, Crawford, and Alchian (1978)). Some of our results are similar. As in the property rights literature, we find that under plausible conditions strictly complementary assets should be owned together and an asset that is idiosyncratic to a party should be owned by that party. (However, we also find that in some cases joint ownership is optimal.) The transaction cost literature has emphasized, among other things, that parties will vertically integrate when quasi-rents are large, but use the market when they are small. This conclusion is consistent with our analysis in the following sense. If quasi-rents are small under nonintegration, then (a) even in the absence of a contract there is little to argue or be aggrieved about since each party can easily switch to another trading partner at any moment; (b) for the same reason neither party is vulnerable to hold-up. Hence nonintegration, in combination with short-run contracts, will do a reasonable job of regulating the relationship. (Note, however, that integration may also perform well under these conditions.)

Although our model shares similarities with the existing literature, there are also important differences. Our model emphasizes the variability of the gains from trade rather than their magnitude (or the sensitivity of their magnitude to investment). This suggests that there will be ways to distinguish our model empirically from the transaction cost and (standard) property rights theories.

As we have emphasized, our theory ignores ex ante considerations. In future work it would be interesting to reintroduce these. For example, suppose that the seller can take an action – an ex ante relationship-specific investment, say – that reduces her cost. To the extent that this moves cost outside the "normal" range, this will give the buyer an incentive to hold the seller up. Anticipating this, the seller's incentive to engage in such an investment is reduced. A model that includes ex ante as well as ex post inefficiencies is likely to provide a richer understanding of the costs and benefits of asset ownership and more generally of vertical integration. Developing a model along these lines is an interesting and challenging goal for future research.

REFERENCES

Baker, George, Robert Gibbons, and Kevin Murphy (2002), "Relational Contracts and the Theory of the Firm," *Quarterly Journal of Economics* 117: 39-84.

Grossman, Sanford, and Oliver Hart (1986). "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy* 94: 691-719.

Halonen, Maija (2002), "Reputation and the Allocation of Ownership," *The Economic Journal* 112: 539-558.

Hart, Oliver, and John Moore (1990), "Property Rights and the Nature of the Firm," *Journal of Political Economy* 98: 1119-58.

________, _______ (2007), "Contracts as Reference Points," *Quarterly Journal of Economics*, forthcoming.

Klein, Benjamin (1996), "Why Hold-ups Occur: The Self-Enforcing Range of Contractual Relationships," *Economic Inquiry* 34: 444-63.

______ (2007), "The Economic Lessons of Fisher Body-General Motors," *International Journal of the Economics of Business* 14: 1-36.

______, R. Crawford, and A. Alchian (1978), "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process," *Journal of Law and Economics* 21: 297-326.

______, and Kevin M. Murphy (1997), "Vertical Integration as a Self-Enforcing Contractual Arrangement," *American Economic Review* 87: 415-420.

Rajan, Raghuram G., and Luigi Zingales (1998), "Power in a Theory of the Firm," *Quarterly Journal of Economics* 113: 361-386.

Williamson, Oliver (1985), *The Economic Institutions of Capitalism.* New York: Free Press.