

HOLD-UP, ASSET OWNERSHIP, AND REFERENCE POINTS*

OLIVER HART

We study two parties who desire a smooth trading relationship under conditions of value and cost uncertainty. A contract fixing price works well in normal times because there is nothing to argue about. However, when value or cost is unusually high or low, one party will deviate from the contract and hold up the other party, causing deadweight losses as parties withhold cooperation. We show that allocating asset ownership and indexing contracts can reduce the incentives to engage in hold-up. In contrast to much of the literature, the driving force in our model is payoff uncertainty, rather than noncontractible investments.

I. 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 (see Hart and Moore [2008]). 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. Suppose that the parties write a rigid contract that fixes price. Such a contract works well in “normal” times because there is nothing to argue about: we assume that, in the absence of argument, the parties are willing to cooperate. However, if value or cost falls outside the normal range, one party will have an incentive to threaten to withhold cooperation unless the contract is renegotiated; that is, the party will engage in hold-up. For example, if value is unusually high, the seller will hold up the buyer to get a higher price, whereas if cost is unusually low, the buyer will hold up the seller to get a lower price. We suppose that hold-up transforms a friendly relationship into a hostile one. The consequence is that the parties withhold cooperation: they operate within the letter, rather than the spirit, of their (renegotiated)

*I am grateful to John Moore for discussions of some of the elements of this paper and to Bob Gibbons and Birger Wernerfelt for many helpful conversations. I would also like to thank Ronen Avraham, Mathias Dewatripont, Florian Englmaier, Rob Gertner, Victor Goldberg, Louis Kaplow, Josh Lerner, Bentley MacLeod, Jeremy Stein, Sasha Volokh, Abe Wickelgren, and Christian Zehnder for useful comments, and Georgy Egorov for excellent research assistance. Four referees, an editor (Larry Katz), and a second editor provided highly constructive input. Financial support from the U.S. National Science Foundation through the National Bureau of Economic Research is gratefully acknowledged.

© 2009 by the President and Fellows of Harvard College and the Massachusetts Institute of Technology.

The Quarterly Journal of Economics, February 2009

contract, causing deadweight losses. 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.

We show that there is a range of prices $[p_L, p_H]$ that depends on the state of the world such that, if the long-term contract price p (chosen before the state of the world is known) lies in this range, hold-up is avoided; whereas if p lies outside this range, hold-up occurs. One can think of $[p_L, p_H]$ as the "self-enforcing" range (see, e.g., Masten [1988] and Klein [1996]). In standard hold-up models, this range is degenerate: $p_L = p_H$. In the current model, it is nondegenerate because the deadweight costs of hold-up introduce friction: $p_L < p_H$. If there is no *ex ante* uncertainty, it is easy to find a p that lies in $[p_L, p_H]$ (exactly where it lies does not matter, because lump-sum transfers can be used to redistribute surplus). However, if the state of the world is uncertain, it may be impossible to find a price p that lies in $[p_L, p_H]$ with probability 1. Under these conditions hold-up will sometimes occur.

We establish two results. First, indexation, by which we mean tying price to a verifiable signal related to industry conditions, can improve matters. If σ is correlated with the state of the world, then indexing the contract price to σ makes it easier to ensure that the contract price lies in $[p_L, p_H]$. This can throw light on the findings on the use of indexation by Joskow (1985, 1988) and Goldberg and Erickson (1987) (see also Card [1986]).¹ Second, an appropriate allocation of asset ownership can help. (More generally, enhancing a party's outside option is useful.) *Ceteris paribus*, allocating an asset to the buyer is good to the extent that this increases the correlation between the buyer's outside option and his value from trade. This improves matters because, when value is unusually high, the buyer's outside option will also be high, which reduces the seller's ability to hold up the buyer. To put it a bit more formally, the $[p_L, p_H]$ range becomes less sensitive to the buyer's value. However, there is a countervailing effect on the seller: taking an asset away from the seller causes the $[p_L, p_H]$ range to become more sensitive to the seller's cost, which makes it easier for the buyer to hold up the seller. Thus, there is a trade-off.

1. MacLeod and Malcomson (1993) provide an alternative explanation of indexation using a model in which parties make noncontractible relationship-specific investments. See Section II.

One aspect of our approach is that, in contrast to much of the literature, it focuses on *ex post*, rather than *ex ante*, inefficiencies. Indeed, (noncontractible) relationship-specific investments play no role.² A benefit of this is that we are able to avoid (and we hope move beyond) the “Maskin-Tirole” foundational critique of models of incomplete contracts (see Maskin and Tirole [1999]). However, to do so we need to introduce a number of behavioral features.

Our model resonates with a number of papers in the empirical/applied contracting literature. A theme of this literature, which studies contractual relationships between buyers and sellers of petroleum coke, trucking services, coal, and aluminum, among other things, is that parties face problems in their trading relationships when the gains from trade become unevenly divided under an existing contract; see Goetz and Scott (1983), Goldberg (1985), Joskow (1985), Goldberg and Erickson (1987), Masten (1988, 2007), Klein (1996), and Lafontaine and Masten (2002). This will lead the disadvantaged party to engage in costly opportunistic or noncooperative behavior in order to wriggle out of, or renegotiate, the contract. To quote Klein (1996, p. 444): “Hold-ups occur when market conditions change sufficiently to place the relationship outside the self-enforcing range.”³ However, the existing literature does not explain why the parties cannot negotiate around the costs of opportunism or why they cannot write an *ex ante* contract that is flexible enough to incorporate anticipated changes in market conditions. A benefit of our approach is that we are able to explain both these things by introducing behavioral features. In our model, hold-up sours a trading relationship, causing deadweight losses from a reduction in cooperation. Moreover, the parties cannot bargain around these losses.

It is useful to compare the current paper with the existing literature on asset ownership and vertical integration. According to transaction cost economics (see, e.g., Williamson [1971] and Klein, Crawford, and Alchian [1978]), contracts between independent parties are problematic because, given contractual

2. Lafontaine and Masten (2002) and Masten (2007) argue that, in the trucking industry, relationship-specific investments are relatively unimportant, and long-term contracts are written to save on *ex post* negotiation or bargaining costs. The approach adopted here is consistent with this view.

3. Examples of opportunistic behavior include breaching the contract, performing according to the literal terms of the contract, exploiting ambiguous terms, contesting facts, or threatening to do these things. See Goetz and Scott (1983, pp. 977, 982) and Goldberg (2006, pp. 328–329). Also see Klein (1996) and Goldberg (2006, Chapter 20) for discussions of the well-known *Alcoa v. Essex* case.

incompleteness, deadweight losses will occur as parties haggle over the *ex post* division of the quasi-rents. A key factor in determining vertical integration decisions is the size of these quasi-rents. According to property rights theory (see, e.g., Grossman and Hart [1986] and Hart and Moore [1990]), parties will bargain around the deadweight losses from haggling, but *ex ante* investments will be distorted. A key factor determining vertical integration is the marginal product of quasi-rents with respect to (noncontractible) *ex ante* investments. This paper emphasizes a third factor: the variability of quasi-rents with respect to the state of the world; that is, *payoff uncertainty*. We return to the implications of our paper for the theory of vertical integration in the Conclusions.

Our paper is also connected to the relational contracts literature. Although our theory is static, it incorporates something akin to the notion of trust or good will; this is what is destroyed if hold-up occurs. In a way, our model can be seen as a reduced form of a relational contracting model. Note, however, that, in contrast to the leading formal models of relational contracts and asset ownership (see, e.g., Baker, Gibbons, and Murphy [2002] and Halonen [2002]), we emphasize *ex post* inefficiency of trade rather than *ex ante* inefficiency of investments.

Although this paper is stimulated by Hart and Moore (2008) (particularly Section III of that paper), there are several important differences. First, Hart and Moore (2008) are concerned with the trade-off between flexible contracts, represented by a price range, and rigid contracts, represented by a single price. For the most part, we focus on rigid, single-price contracts (although we show in the Appendix that all our results generalize to flexible contracts). Second, Hart and Moore (2008) ignore assets and do not allow hold-up. Parties quit in states of the world where their payoff is negative, but in the leading version of the model, there is no renegotiation, and so the parties never threaten to quit in order to get a better deal. We believe that the case of hold-up is interesting in its own right, but we have another reason for introducing it here. In any model of asset ownership, it is plausible that the parties can buy and sell assets if the relationship breaks down, that is, some renegotiation is possible. But if this is so, then a party can hold up the other party by quitting and retrading assets. In this sense, we believe that the model of this paper, in which hold-up and renegotiation are allowed, is a natural extension of Hart and Moore (2008) to study asset ownership. In

the conclusions, we will return to this issue and briefly discuss what happens when asset retrading is not possible.

The paper is organized as follows: In Section II, we lay out the basic model, and in Section III, we introduce assets. Section IV contains extensions and Section V concluding remarks. The Appendix studies a more general class of contracts and contains proofs of propositions.

II. 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 shortly before date 1, at date 1^- , say. There is symmetric information throughout, and the parties are risk-neutral and 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., if the parties cooperate at date 1), and let r_b, r_s denote B and S 's outside options. The random variables v, c, r_b, r_s are uncertain at date 0, and, although they are assumed to be observable to both parties once they are realized, they are not verifiable.

We follow Hart and Moore (2008) in supposing that, for the gains from trade to be fully realized, each party must take a number of "helpful" or "cooperative" actions at date 1. We assume that none of these actions can be specified in a date 0 contract: they are too complicated to anticipate or describe in advance. When the uncertainty is resolved at date 1^- , *some* of these actions become describable and so can be contracted on, whereas *others* are never contractible.⁴ Thus some modification or renegotiation of the contract is possible at date 1^- . We assume that all the helpful actions are chosen simultaneously by B and S at date 1. See Figure I for a timeline.

4. An example of an *ex post* contractible helpful action would be one in which one party allowed the other party to modify the characteristics of the good to be traded, e.g., the delivery time. An example of an *ex post* noncontractible (or less contractible) helpful action would be one in which one party provided useful information to the other party, consulted with the other party before making a decision, or responded to that party's phone calls or e-mails expeditiously. For evidence that the same contractible action can lead to different performance outcomes (depending on ownership structure), suggesting that noncontractible actions are important, see Januszewski Forbes and Lederman (2007).

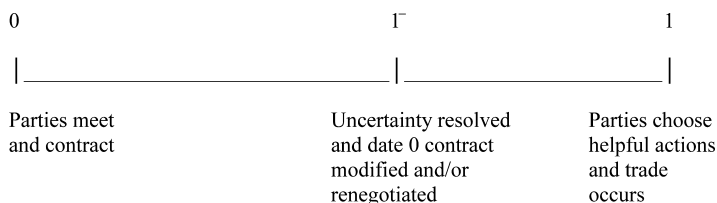


FIGURE I

We make the following assumptions:

- A1. If at date 1 all helpful actions are taken, 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 at date 1 all the contractible, but none of the noncontractible, helpful actions are taken, the value of the widget to B is $v - (1/2)\lambda(v - c)$, and the cost to S is $c + (1/2)\lambda(v - c)$, where $0 < \lambda < 1$. Hence, net surplus = $(1 - \lambda)(v - c)$ in this case.
- A3. If at date 1 none of the helpful actions (contractible or otherwise) are taken, B 's value is very low (approximately $-\infty$) and S 's cost is very high (approximately $+\infty$). In this case, each party walks away from the contract (neither party has an incentive to enforce it), and no trade occurs; that is, the parties earn their outside options.

Note that the import of (A2) is that withholding noncontractible helpful actions moves v and c in the direction of $\frac{1}{2}(v + c)$ (for example, set $\lambda = 1$). We normalize the “walkaway” or no-trade price in (A3) to be zero.

What determines whether a party is helpful? As in Hart and Moore (2008), we assume that being helpful does not cost significantly more than not being helpful: either it costs slightly more or it costs slightly less, that is, a party may actually enjoy being helpful. However, helpful actions have great value for the recipient. To simplify matters, we assume that a party is completely indifferent between being helpful and not.

Given this indifference, we take the view that a party will be willing to be helpful if (and only if) he is “well treated” by the other party.⁵ Importantly, as in Hart and Moore (2008), a party

5. There is a large amount of empirical and experimental evidence in support of reciprocal behavior. See the references in Hart and Moore (2008). Works that are particularly relevant for our context include Bewley (1999) and Kube, Maréchal, and Puppe (2007).

is “well treated” if he receives what he feels entitled to under the contract. In particular, the date 0 contract is a reference point for entitlements: among other things, it is regarded as fair, for example, because it is negotiated under competitive conditions. Implicitly, we are assuming a more competitive situation at date 0 than at date 1^- ; this could be because the parties make (unmodeled, contractible) relationship-specific investments after date 0. For further discussion, see Hart and Moore (2008).

It is useful to begin with the case where the parties write a “simple” contract at date 0 that specifies a single trading price p (recall that the no-trade price is normalized to zero). Consider what happens at date 1^- once the uncertainty is resolved. Each party has a choice. He can stick to the contract, or he can try to force the other party to renegotiate the contract—we interpret this as “hold-up.”

Let’s consider the scenario where the parties stick to the contract. Under these conditions, each party feels well-treated by the other party, since he is getting exactly what the contract said he would. Thus, each party is willing to be helpful, and all cooperative actions are undertaken. The buyer’s and seller’s payoffs are, respectively,

$$(1) \quad U_b = v - p,$$

$$(2) \quad U_s = p - c.$$

In the second scenario, one party engages in hold-up: that is, he tries to force the other party to renegotiate the contract. He does this by threatening not to undertake any helpful actions unless he receives a side-payment. We assume that such behavior is viewed as a hostile act by the victim—it is a breach of the spirit of the date 0 contract—and leads, in the first instance, to the end of all cooperation. (One can imagine that the victim reacts to the attempted hold-up with harsh words and so the dislike of the perpetrator and victim is now mutual.) The result is a Nash equilibrium where neither party cooperates. (Because parties are indifferent between cooperating and not, this *is* a Nash equilibrium.) This yields the no-trade outcome described in (A3), with payoffs r_b, r_s for B, S , respectively.

Before proceeding, we should highlight an important assumption. We are supposing that it is impossible for outsiders, for example, a court, to determine who was responsible for the hold-up and penalize that party accordingly. To put it another way, we do

not allow for damages for “breach of contract.” In this sense, the contract that the parties write at date 0 is not binding; it is more like an “agreement to agree” (see, e.g., Corbin [1993, Chapters 2 and 4]). We discuss this further in Section IV.

So, in the first instance, the outcome after hold-up is no trade. However, renegotiation is possible.⁶ Even if the relationship is soured, the parties can and will agree to undertake the helpful contractible actions at date 1. At the same time, it is assumed that neither party will provide noncontractible cooperation (again, this is a Nash equilibrium). In effect, the parties have a cold but correct relationship.⁷ Renegotiation therefore yields surplus $(1 - \lambda)(v - c)$ by (A2). Thus, if

$$\text{A4. } (1 - \lambda)(v - c) > r_b + r_s,$$

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

In summary, hold-up leads to a loss of surplus of $\lambda(v - c)$. This is a deadweight loss from a soured relationship.

We assume a 50:50 split of the surplus from renegotiation, so after hold-up, the parties’ payoffs are

$$(3) \quad U_b = r_b + \frac{1}{2}G,$$

$$(4) \quad U_s = r_s + \frac{1}{2}G,$$

where

$$(5) \quad G = (1 - \lambda)(v - c) - r_b - r_s.$$

It is easy to determine when hold-up occurs. Define p_L to be the price p such that S is indifferent between receiving p and holding B up and p_H to be the price such that B is indifferent

6. In fact, we assume that the parties cannot commit *ex ante* not to renegotiate. In Section IV, we discuss whether such a commitment would be desirable.

7. Note that we assume 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 side-payment 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. For some evidence supporting this view and suggesting that even “big” players act in this way, see the *New York Times*, August 27, 2007, p. A1 (reporting on the fact that the renegotiation of a large buy-out of a Home Depot business, triggered by the 2007 subprime mortgage crisis, created considerable bad feeling).

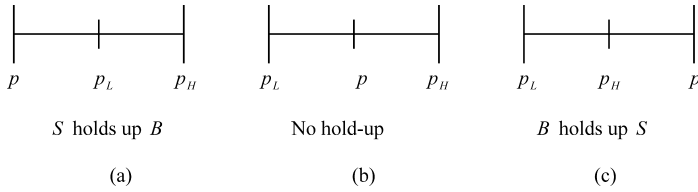


FIGURE II

between paying p and holding S up. Then from (1)–(4)

$$(6) \quad p_L - c = r_s + \frac{1}{2}G,$$

$$(7) \quad v - p_H = r_b + \frac{1}{2}G,$$

and so

$$(8) \quad p_L = c + r_s + \frac{1}{2}G,$$

$$(9) \quad p_H = v - r_b - \frac{1}{2}G.$$

Note that

$$(10) \quad p_H - p_L = \lambda(v - c) > 0.$$

Equation (10) reflects the fact that there is some friction in the renegotiation process. If hold-up did not lead to the souring of the relationship, λ would be zero, and $p_H = p_L$. However, because hold-up causes some dissipation of surplus, $p_H > p_L$: the price at which B is just ready to hold up S is strictly greater than the price at which S is just ready to hold up B .

Because S is indifferent between holding B up and not at $p = p_L$, S will strictly prefer to hold B up when $p < p_L$. Similarly, B will strictly prefer to hold S up when $p > p_H$. Thus, hold-up is avoided if and only if

$$(11) \quad p_L < p < p_H.$$

Note that p_H, p_L are random variables, whereas p is not because it is chosen *ex ante*. The situation is illustrated in Figure II. The interval $[p_L, p_H]$ can be interpreted as the “self-enforcing” contractual range (see Klein [1996]). Masten (1988) derives a similar “self-enforcing” price range under the assumption that there is an exogenous cost of hold-up.

To make progress, we put more structure on the random variables r_b, r_s . We assume

$$(12) \quad r_b = \alpha_b + \beta_b v + \varphi + \gamma_b \varepsilon,$$

$$(13) \quad r_s = \alpha_s - \beta_s c + \gamma_s \eta,$$

where

$$(14) \quad 1 - \lambda > \beta_b > 0, \quad 1 - \lambda > \beta_s > 0, \quad \gamma_b > 0, \quad \gamma_s > 0.$$

Here $\alpha_b, \beta_b, \gamma_b, \alpha_s, \beta_s,$ and γ_s are constants (later they will depend on the assets the parties own), and $\varphi, \varepsilon, \eta$ are independent random variables with mean zero. Equations (12)–(14) capture the idea that B and S 's outside options covary with v, c , respectively, but not too strongly, and are also subject to exogenous noise (ε, η). The noise term φ is a smoothing device: its rationale will become clear later.

Given (12) and (13), 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^- , but it is not verifiable. (If ω were verifiable, p could be indexed to ω , and the first-best could be achieved—see below.) It is useful to rewrite p_L, p_H as functions of ω . From (8) and (9), we have

$$(15) \quad p_L(\omega) = \frac{1}{2} [\alpha_s + \gamma_s \eta - \alpha_b - \varphi - \gamma_b \varepsilon + ((1 - \lambda) - \beta_b)v + ((1 + \lambda) - \beta_s)c],$$

$$(16) \quad p_H(\omega) = \frac{1}{2} [\alpha_s + \gamma_s \eta - \alpha_b - \varphi - \gamma_b \varepsilon + ((1 + \lambda) - \beta_b)v + ((1 - \lambda) - \beta_s)c].$$

Clearly p_L, p_H are monotonic in v, c . Because hold-up occurs when $p_L(\omega) > p$ or $p_H(\omega) < p$, that is, when $p_L(\omega)$ is high or $p_H(\omega)$ is low, it follows that hold-up occurs, *ceteris paribus*, if v is unusually high or low or c is unusually high or low. This is intuitive: if v is high, S does well in the renegotiation process because there is a lot of surplus on the table (even taking into account that a fraction λ gets lost), and so S has an incentive to hold B up; similarly, if c is low, B does well in the renegotiation process, and so B has an incentive to hold S up. In addition, it is clear from (15) and (16) that the effect of unusual values of v, c is less pronounced if β_b and β_s are large, because p_L, p_H are less sensitive to v, c under these conditions.

We turn now to an optimal simple contract. Recall that a simple contract consists of a single price, where p is chosen before the state of the world is realized. Since date 0 lump-sum transfers can be used to allocate surplus, an optimal contract maximizes expected net surplus. Thus, an optimal simple contract solves

$$(17) \quad \left\{ \begin{array}{l} \text{Max}_p \int (v - c) dF(\omega) \quad + \int (1 - \lambda)(v - c) dF(\omega) \\ p_L(\omega) \leq p \leq p_H(\omega) \quad p < p_L(\omega) \\ \text{or } p > p_H(\omega) \end{array} \right\},$$

where F is the distribution function of ω .

It is clear that the first-best can be achieved in the case of certainty: just pick any price p in the interval $[p_L(\omega_0), p_H(\omega_0)]$, where ω_0 is the state of the world. However, the first-best typically cannot be achieved under uncertainty, because it is not generally possible to find a single price that lies in the intersection of a number of different $[p_L(\omega), p_H(\omega)]$ intervals.

Our analysis so far has a shortcoming. Suppose that the parties write a simple contract at date 0. Then, as we have observed, with uncertainty it is very likely that p will lie outside the $[p_L(\omega), p_H(\omega)]$ range for some ω , and so one party will hold up the other to get a better price. Why don't the parties anticipate this and build the renegotiated price into the original contract? For example, the initial contract could state that the price will normally be 10 but can rise to 20 in unusual circumstances. Or the contract might give one party the right to choose the price from a menu of prices.

To deal with (some of) these possibilities, in the Appendix we broaden the analysis to allow the parties to specify a *range* of possible trading prices $[\underline{p}, \bar{p}]$ in their date 0 contract. In each state of the world, the parties negotiate about which price to pick in the range. As long as they stay in the range, hold-up is avoided. Thus, the benefit of a large price range is that it reduces hold-up. However, as in Hart and Moore (2008), a large price range has a cost. We assume that at date 1^- each party may feel entitled to a different price in the range. Not getting what you are entitled to leads to aggrivement and "shading," that is, the partial withholding of cooperation. Note that, without shading, the first-best could be achieved with a price range $\underline{p} = -\infty, \bar{p} = \infty$. Because no restrictions are put on price, this is equivalent to not writing a contract at all at date 0. Recall that there are no *ex ante* (noncontractible)

investments in our model, and information is symmetric, so the usual arguments for long-term contracting do not apply.

However, if grievement and shading exist, then a very large price range is suboptimal, because it leads to shading in all states of the world. Thus, the parties will prefer to accept the risk of hold-up: they will choose a limited price range, a fixed price contract being an extreme example of this. In the Appendix, we show that all the results of this and the next section generalize to the case of price ranges. We also show that introducing price ranges can elucidate Goldberg and Erickson's (1987) observation that parties write shorter-term contracts in a more volatile environment.

We close this section by discussing indexation. Suppose that there is a verifiable signal σ that is correlated with the state of the world. Then, under weak assumptions, the parties can improve on the optimal contract in (17) by indexing the price to σ . To see this, suppose that σ takes on values $\sigma_1, \dots, \sigma_n$ with associated strictly positive probabilities π_1, \dots, π_n . An indexed contract consists of a price vector (p_1, \dots, p_n) , where p_i is the trading price that will rule if $\sigma = \sigma_i$. By analogy with (17), define $S_i(p)$ to be expected surplus conditional on the signal $\sigma = \sigma_i$; that is,

$$(18) \quad S_i(p) = \int_{p_L(\omega) \leq p \leq p_H(\omega)} (v - c) dF(\omega/\sigma_i) + \int_{p < p_L(\omega)} (1 - \lambda)(v - c) dF(\omega/\sigma_i),$$

$$\text{or } p > p_H(\omega),$$

where $F(\omega/\sigma_i)$ is the distribution function of ω conditional on σ_i . Then total expected surplus = $\sum_{i=1}^n \pi_i S_i(p_i)$, and an optimal indexed contract solves

$$(19) \quad \text{Max}_{p_1, \dots, p_n} \sum_{i=1}^n \pi_i S_i(p_i).$$

Obviously, a necessary and sufficient condition for (p_1, \dots, p_n) to solve (19) is that

$$(20) \quad p_i \text{ solves : Max}_p S_i(p) \text{ for all } i.$$

Assume that the solution to (20) is unique.

Now return to the case of nonindexation. An optimal nonindexed contract maximizes the objective function in (19), but

subject to the constraint that $p_1 = \dots = p_n$; that is, it solves

$$(21) \quad \text{Max}_p \sum_{i=1}^n \pi_i S_i(p).$$

Let the solution be p^* . Obviously, $S_i(p_i) \geq S_i(p^*)$ for all i , because p^* is a feasible choice in (20). Hence,

$$(22) \quad \sum_{i=1}^n \pi_i S_i(p_i) \geq \sum_{i=1}^n \pi_i S_i(p^*).$$

Moreover, the only way that (22) can hold with equality is if

$$(23) \quad S_i(p_i) = S_i(p^*) \text{ for all } i.$$

In other words, as long as the solutions to (20) are not all the same, indexation strictly dominates nonindexation. We have proved

PROPOSITION 1. Assume that the (unique) solution of (20) varies with i . Then an indexed contract is strictly superior to a nonindexed contract.

A special case of Proposition 1 occurs when σ is perfectly correlated with ω . Then the first-best can be achieved under indexation by setting $p_i \in [p_L(\omega_i), p_H(\omega_i)]$, where ω_i is the unique state corresponding to σ_i . In other words, indexation avoids hold-up completely. However, in general, indexation will not achieve the first-best.

Our results are consistent with Joskow's (1985, 1987) and Goldberg and Erickson's (1987) findings that price indexation is a common feature of contracts between suppliers and purchasers of petroleum coke and coal. Although risk aversion is a possible explanation, these authors argue that it is more likely that price indexation is used to reduce opportunistic behavior. This is as in our model. It is also worth noting that MacLeod and Malcomson (1993) present a related explanation of these empirical findings using a model in which indexation improves parties' incentives to make noncontractible relationship-specific investments. In their model, as in ours, indexation makes renegotiation less likely; however, they consider only a restricted set of contracts.

Card (1986) studies the use of wage indexation in union labor contracts in Canada. He argues that the benefit of indexation is that it allows wages to track more closely firms' demand for labor and workers' alternative opportunities. To the extent that the

demand for labor is related to v and r_b and workers' alternative opportunities to r_s , our model seems consistent with this. At the same time, it is not so easy to explain two of Card's other observations: that wages are always indexed to the consumer price index, as opposed to some other measure of industry conditions, and that indexation is much rarer in the nonunion sector.

It should be emphasized that "simple" indexation is not the only way to avoid hold-up. We have supposed that the seller's cost is unverifiable, but in practice some measure of cost may be available. One possibility is to link price to this measure, as in a "cost-plus" contract. Such an arrangement has well-known incentive problems. Another possibility is to include a provision allowing the contract to be renegotiated if some exogenous index hits a minimum or maximum value. (The idea is that renegotiation under these conditions will not be viewed as a hostile act.) These possibilities are observed in practice. As it stands, our model is too simple to allow for them. However, we believe that they could be incorporated in an extension (the discussion of price ranges in the Appendix is a first step).

III. ASSET OWNERSHIP

In this section, we explore the idea that asset ownership can improve the parties' trading 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.⁸ This in turn affects parties' outside options and their incentives to engage in hold-up.

Denote by A the set of all assets at B and S 's disposal; we assume A is fixed and finite.⁹ Let A_b be the set of assets B owns and A_s the set of assets S owns. We assume

$$(24) \quad A_b \cap A_s = \emptyset, \quad A_b \cup A_s \subseteq A.$$

The first part of (24) 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: that is, joint ownership gives each party a veto right on its use.¹⁰

8. As in Hart and Moore (1990). See also Grossman and Hart (1986).

9. We suppose that the assets in A are already specialized and so cannot be bought or sold on the open market.

10. We confine attention to simple ownership structures. It would not be difficult to generalize the analysis to allow for options to own, etc.

We now suppose that the coefficients $\alpha_b, \beta_b, \gamma_b, \alpha_s, \beta_s, \gamma_s$ depend on asset ownership. In particular, $\alpha_b = \alpha_b(A_b), \beta_b = \beta_b(A_b), \gamma_b = \gamma_b(A_b), \alpha_s = \alpha_s(A_s), \beta_s = \beta_s(A_s),$ and $\gamma_s = \gamma_s(A_s)$. We also make assumptions similar to those in the property rights literature about how these coefficients vary with asset ownership. In particular, we assume that owning more assets increases the *marginal* payoffs of r_b, r_s with respect to v and c . That is,

$$(25) \quad \beta_b \text{ is nondecreasing in } A_b,$$

$$(26) \quad \beta_s \text{ is nondecreasing in } A_s;$$

(25) and (26) can be said to capture the idea that assets are specific to the business the buyer and seller are engaged in. Of course, one might also suppose that a party's *total* payoff increases in the assets he or she owns (α_b, α_s increasing in A_b, A_s), but we will not need this in what follows. We assume that (14) and (A4) hold for all ownership structures.

We suppose that assets can be traded at date 0. Thus, a contract is now a 3-tuple (A_b, A_s, p) , specifying an asset ownership allocation (A_b, A_s) and a date 1 price p , where A_b, A_s satisfy (24). As in Section II, given that date 0 lump sum transfers are possible, an optimal contract maximizes expected net surplus. Thus, an optimal contract solves

$$(27) \quad \text{Max}_{(A_b, A_s, p)} \left\{ \begin{array}{l} \int (v - c) dF(\omega) \qquad \qquad \qquad + \int (1 - \lambda)(v - c) dF(\omega) \\ p_L(\omega; A_b, A_s) \leq p \leq p_H(\omega; A_b, A_s) \quad p < p_L(\omega; A_b, A_s) \\ \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \text{or } p > p_H(\omega; A_b, A_s) \end{array} \right\},$$

where p_L, p_H are now indexed by the asset ownership allocation (A_b, A_s) , as well as by ω .

We begin our analysis of asset ownership by considering what happens if, *ceteris paribus*, assets are transferred at date 0 from S to B . Then, given (25) and (26), β_b rises and β_s falls. As is clear from (15) and (16), this makes p_L and p_H less sensitive to v than before, because

$$(28) \quad \frac{\partial p_L}{\partial v} = \frac{1}{2}((1 - \lambda) - \beta_b),$$

$$(29) \quad \frac{\partial p_H}{\partial v} = \frac{1}{2}((1 + \lambda) - \beta_b),$$

and these both decrease. On the other hand, $p_L(\omega)$ and $p_H(\omega)$ become more sensitive to c , because

$$(30) \quad \frac{\partial p_L}{\partial c} = \frac{1}{2}((1 + \lambda) - \beta_s),$$

$$(31) \quad \frac{\partial p_H}{\partial c} = \frac{1}{2}((1 - \lambda) - \beta_s),$$

and these both increase.

Intuitively, a reduction in sensitivity of p_L , p_H is good because, if the interval $[p_L, p_H]$ does not vary much, it is easier to find a price p that lies in $[p_L, p_H]$ for many ω . That is, hold-up is less likely. This suggests that it is optimal for B to own all the assets if only v varies, because this minimizes the sensitivity of p_L and p_H with respect to the state of the world, whereas it is optimal for S to own all the assets if only c varies. Proposition 2 confirms this.

PROPOSITION 2. (1) Suppose that $\varphi = \varepsilon = \eta \equiv 0$ and $c \equiv c_0$, where c_0 is a constant. Then there exists an optimal contract in which B owns all the assets, that is, $A_b = A$, $A_s = \emptyset$.

(2) Suppose that $\varphi = \varepsilon = \eta \equiv 0$ and $v \equiv v_0$, where v_0 is a constant. Then there exists an optimal contract in which S owns all the assets, i.e., $A_s = A$, $A_b = \emptyset$.

In the Appendix, we prove a more general version of Proposition 2 and of all the other propositions in this section. In these more general versions, a range of possible trading prices is allowed in the date 0 contract.

Note that it would not be difficult to establish uniqueness in Proposition 2 under slightly stronger stochastic assumptions. Using such assumptions one could also generalize Proposition 2 to the case where uncertainty about v or c is small.

Proposition 2 is reminiscent of the result in the property rights literature that one party should own all the assets if his investment is important. Here, the conclusion is that one party should own all the assets if his payoff is uncertain. Proposition 2 is also similar to results that have been obtained by Simon (1951) and Wernerfelt (1997), among others, showing that the party with a more variable payoff should be the boss.

Of course, in general, both v and c vary. Proposition 2 is not very helpful here, because it tells us only when one party should own everything. However, some progress can be made if we introduce the idea of an idiosyncratic asset.

Define an asset to be idiosyncratic to B if B 's owning it increases the sensitivity of r_b to v and S 's 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 $A_b \subseteq A$, $A_b \cap \{a\} = \emptyset$ 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 $A_s \subseteq A$, $A_s \cap \{a\} = \emptyset$ and $\beta_b(A_b \cup \{a\}) = \beta_b(A_b)$ for all $A_b \subseteq A$.

In other words, an asset is idiosyncratic to B if it is specific to B 's business and not to S 's, and vice versa for an asset idiosyncratic to S . Note that one reason an asset may be idiosyncratic to a party is that that party has human capital that is complementary to the asset; for example, he is the only person who knows how to operate it. Under these conditions, taking away the asset from the other party is unlikely to affect the sensitivity of that party's outside option to the state of the world.

Allocating an asset to the party to whom it is idiosyncratic would seem desirable, because it reduces the variability of the $[p_L, p_H]$ range. However, it turns out that, to establish this, one must make strong assumptions about the stochastic structure. In the next proposition, we suppose that, with high probability, v, c take on "normal" values $v = v_0, c = c_0$, whereas with low probability, v, c can each take on an "exceptional" value. Because 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 $\varepsilon = \eta = 0$.

PROPOSITION 3. Assume that $\varepsilon = \eta = 0$ and that φ is uniformly distributed on $[-k, k]$. Suppose that with probability $0 < \pi < 1$, event 1 occurs: $v = v_0, c = c_0$; 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

$$v_H \geq \frac{-\beta_b(A)v_0 + (1 + \lambda)v_0 - 2\lambda c_0}{1 - \lambda - \beta_b(A)};$$

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

$$c_L \leq \frac{-\beta_s(A)c_0 - 2\lambda v_0 + (1 + \lambda)c_0}{1 - \lambda - \beta_s(A)}.$$

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. See Appendix.

It is useful to understand why Proposition 3 requires such strong assumptions about the probability distribution of v and c . The reason is the following. Let p be the optimal price for the general case where v, c are uncertain. Suppose that we transfer an asset that is idiosyncratic to B from S to B . (In what follows, we suppress assets in the notation.) We know that this will reduce the variability of $p_L(\omega)$, $p_H(\omega)$ with respect to v . But this might *reduce* the probability that p lies in $[p_L(\omega), p_H(\omega)]$ if, for example, $p \in [p_L(\omega_1), p_H(\omega_1)]$, $p \notin [p_L(\omega_2), p_H(\omega_2)]$, and $[p_L(\omega_1), p_H(\omega_1)]$ moves closer to $[p_L(\omega_2), p_H(\omega_2)]$. The stochastic structure in Proposition 3 avoids this kind of situation.

A simple application of Proposition 3 is to the case of strictly complementary assets. Suppose assets a_1 and a_2 are strictly complementary, in the sense that they are valuable only when used together. Then a_2 by itself is of no use to S , whereas 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 *inalienable*, that is, 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 3, 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 3.

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 the party's payoff relative to its outside option. However, it is possible that asset ownership increases the variability of outside options relative to inside values. Under these conditions, it may be better to take assets away from people. The next proposition describes

a situation where joint ownership is optimal. In this proposition, inside values v and c are constant, whereas outside options are stochastic; in particular, φ and ε or η vary.

PROPOSITION 4. Assume γ_b, γ_s are strictly increasing in A_b, A_s , respectively, and φ is uniformly distributed on $[-k, k]$. Suppose that with probability $0 < \pi < 1$, event 1 occurs: $v = v_0, c = c_0, \varepsilon = 0, \eta = 0$; with probability $(1 - \pi)\alpha_\varepsilon$, event 2 occurs: $v = v_0, c = c_0, \eta = 0, \varepsilon$ has support $[\varepsilon_L, \varepsilon_H]$, where

$$\varepsilon_L \leq \frac{2\lambda(c_0 - v_0)}{\gamma_b(\emptyset)}$$

and $\varepsilon_H > 0$; with probability $(1 - \pi)\alpha_\eta$, event 3 occurs: $v = v_0, c = c_0, \varepsilon = 0, \eta$ has support $[\eta_L, \eta_H]$, where

$$\eta_L \leq \frac{2\lambda(c_0 - v_0)}{\gamma_s(\emptyset)}$$

and $\eta_H > 0$. 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. See Appendix.

Again similar results have been obtained in the property rights literature in situations where *ex ante* investments are used for “rent-seeking” purposes (see, e.g., Rajan and Zingales [1998] and Halonen [2002]). Just as in that literature, it is ultimately an empirical matter as to whether the conditions of Proposition 4, which argue for joint ownership, are more or less likely to hold than those of Propositions 2 and 3, which argue for unitary ownership.

IV. DISCUSSION AND EXTENSIONS

In this section, we discuss the robustness of our analysis and consider some extensions.

An important assumption that we have made is that the initial contract is sufficiently incomplete so that one party can threaten to void it by refusing to cooperate. Moreover, it is impossible for outsiders, for example, a court, to determine who the offending party is and to penalize that party accordingly. Of course, if a court could do that then the parties could always set the

damage payment for “noncooperation” or “breach” high enough so that hold-up never occurred.

Although our assumption is strong, we believe that it is defensible. Although a court may often be able to learn whether one party breached, or more generally acted in bad faith, the learning process is noisy, and so the parties may not wish the penalty for being obstructive to be that high. But this means that a party can use the threat to breach and trigger an uncertain (and possibly lengthy) judicial process to renegotiate the contract. To put it another way, we have considered the extreme case where outsiders have *no* information about who breached and so the default or noncooperative outcome (given, according to (A3), by no trade) is independent of who holds up whom. However, we believe that the model generalizes to situations where outsiders have *some* information and the noncooperative outcome (which might also be something other than no trade) can vary, but insufficiently, with who holds up whom.

It is also worth noting that there are other ways to initiate breach than those described in the model. We have assumed that a party is indifferent between being cooperative or not. However, it may be that, for trade to be efficient, the parties must modify the traded good in ways that increase the seller's costs. If everything is going well, the parties will adjust the price in some reasonable manner (e.g., the buyer pays the incremental cost). However, if the seller is dissatisfied with the division of surplus, she can use the occasion to reopen price negotiation. It may be very hard for outsiders to control such behavior.¹¹

At the same time, these considerations do help to delineate the situations to which the model applies. It is likely to be more relevant when the initial contract is highly incomplete and a party can wriggle out of it in subtle and covert ways.

There are several other assumptions that could be relaxed. First, we have supposed that, even though hold-up causes dead-weight losses, the parties prefer to continue trading with each other after hold-up rather than go their separate ways (assumption (A4)). Dropping this assumption opens up interesting new possibilities. Two cases need to be distinguished. In the first, although the trading relationship is terminated after hold-up, the parties can costlessly renegotiate asset ownership (this is similar

11. I am grateful to Birger Wernerfelt for suggesting this.

to the assumption made in Baker, Gibbons, and Murphy [2002]). Under these conditions, the surplus obtained after hold-up is independent of the date 0 assignment of assets. Because this is the critical feature of (A4), our results are unlikely to change significantly. The second case occurs when there are frictions in the asset renegotiation process (perhaps because it is difficult even to trade assets after a relationship is soured), in which case the date 0 ownership structure will affect *ex post* surplus. Under these conditions, the initial assignment of assets matters not only in determining when hold-up occurs, but also in affecting *ex post* surplus when it does. This opens the door to a richer theory of ownership.¹²

We have also assumed that the parties cannot commit not to renegotiate their contract. In most models of incomplete contracts this is crucial: if the parties can commit, the first-best is achieved. This is not true here. In the absence of renegotiation, hold-up will never occur, because it is impossible to change the terms of trade. However, the parties will have an incentive to quit by refusing to cooperate whenever their outside option exceeds their payoff inside the relationship. In particular, the buyer will quit when $v - p < r_b$, and the seller will quit when $p - c < r_s$. (This is similar to the case analyzed in Hart and Moore [2008].) There will still be inefficiency—in fact, in some cases it will be greater. Asset ownership will matter, although the effects will be different. In fact, the model is likely to resemble the second case discussed above in which (A4) doesn't hold, because, as in that case, asset ownership will affect *ex post* surplus in the event that the relationship breaks down.¹³

Finally, it would be worth relaxing our assumption that hold-up sours the parties' relationship by the same amount regardless of the reasons for it. Consider two cases of seller hold-up. In one, the seller holds the buyer up because the buyer's value has increased. In the other, the seller holds up the buyer because the seller's cost has increased. Even though price rises in both cases, it is arguable that the seller's behavior is less opportunistic in the

12. For a start in this direction, see Hart and Moore (2007).

13. An interesting possibility is that, rather than committing not to renegotiate, the parties might try to constrain the renegotiation process in their date 0 contract. For example, they could allocate the bargaining power to one party or put in place an information revelation game to be played if the relationship breaks down. Note, however, that one point of view is that constraining the renegotiation process enlarges the set of admissible, that is, "legitimate," outcomes in the initial contract and hence increases parties' feelings of entitlement. This raises grievement and shading costs, as in the model of the Appendix.

second case, and so the buyer will be less angry, perhaps because, without the price increase, the seller's profits would be negative and she would quit. We have taken the contrary view that the buyer will view the cases as the same, on the ground that, if the seller wanted the price to rise under some conditions, she should have built this into the initial contract (see the model of the Appendix). However, it seems useful to develop a model where the two cases are treated differently.

V. CONCLUSIONS

In this paper, we have studied a buyer and seller who have a long-term relationship in an uncertain environment. A fixed price contract works well in normal times because there is nothing to argue about. However, if value or cost is unusually high or low, one party will have an incentive to hold up the other, which sours the relationship and leads to the end of cooperation. We have shown that contract indexation can help by making it more likely that price remains within the "self-enforcing" range. In addition, a judicious allocation of asset ownership increases efficiency to the extent that it raises the correlation between parties' outside options and their value inside the relationship; this reduces the likelihood of hold-up.

It is worth noting that, in our model, assets are equivalent to outside options. In practice, there are other ways a party can guard against hold-up than through owning assets. For example, a buyer can choose a flexible technology, so that it is easier to switch suppliers or form a relationship with more than one seller at date 0, so that it can play them off against each other. Our analysis throws light on the desirability of these kinds of strategies as well.

We have argued that our analysis can help to understand the empirical contracting literature. Several papers in this literature emphasize the idea that parties face problems in their trading relationships when the gains from trade become unevenly divided under an existing contract. However, there is no developed theory of why such a situation leads to inefficiency. We are able to provide a theory by appealing to behavioral ideas. When one party is doing badly, he or she will hold up the other party, leading to hostility and deadweight losses from the end of cooperation. Moreover, the parties cannot bargain around these costs.

Our paper also throws light on the vertical integration literature. Our model, where costly hold-up typically occurs with positive probability under uncertainty, seems relevant for understanding the costs of nonintegration and why vertical integration—interpreted as a transfer of assets—can improve matters. Furthermore, we identify different factors from the literature as being important: payoff uncertainty rather than the size of payoffs (as in the transaction cost literature) or the sensitivity of payoffs to investments (as in the property rights literature).¹⁴ It is less clear that our model can be applied directly to understand the costs of integration. Suppose that the buyer, for instance, purchases all the assets of the seller; that is, the seller becomes an employee. According to our model, in a state of the world where the seller's cost is low, her outside option will be lower than before given that she has no assets. This gives the buyer the chance to hold her up. However, it is a bit of a stretch to argue that an employee's effort costs—as opposed to a seller's production costs—vary so much exogenously that hold-up by a boss is a serious problem in practice. It seems more plausible that the seller's costs vary for endogenous reasons: an employee, fearing hold-up by a boss, is less likely to make an *ex ante* investment in cost reduction than an independent contractor. Thus, in order to understand the costs of vertical integration, it may be necessary to (re-)introduce (noncontractible) *ex ante* investments into the model.

An obvious question to ask is, to what kind of firm does the model apply? One point of view is that the feelings of hostility and grievement that we have emphasized are plausible in the case of small, owner-managed firms but less so in the case of large corporations. We are not sure that this is true. Large corporations are run by individuals who have big egos and presumably therefore can have strong emotions. In fact, the Home Depot case described in footnote 7, where contract renegotiation caused coldness, is one involving a large company. Still, it is undoubtedly the case that the effects we describe are likely to be different in a company where many decisions are delegated. In fact, one interesting trade-off is that it might be good for a subordinate to make contract decisions given that he does not care about hold-up (e.g., because he is on low-powered incentives), and hence this will reduce *ex post*

14. Lafontaine and Slade (2007) provide an excellent survey of the empirical literature on vertical integration, including that part of it that deals with payoff uncertainty.

deadweight losses, but it might be bad because he will negotiate an unfavorable deal in the first place.

Of course, the question of how decisions are delegated in large firms is a much more general one. Most models to date in the incomplete contracts literature, including this one and the companion model in Hart and Moore (2008), involve only two parties. Generalizing the model to allow for multiple parties, with an eye to studying the internal organization of firms, is an important and challenging goal for future research.

APPENDIX

We first present a discussion of price ranges. We then prove Propositions 2–4.

As we noted in Section II, anticipating the possibility of hold-up, the parties can build some price flexibility into their contract. Following Hart and Moore (2008), we therefore now allow the parties to specify a *range* of possible trading prices $[\underline{p}, \bar{p}]$ in their date 0 contract; the idea is that the parties will agree on a price in $[\underline{p}, \bar{p}]$ once the uncertainty is resolved at date 1^- .

Suppose that the parties pick the range $[\underline{p}, \bar{p}]$. What happens at date 1^- once ω is realized? There are two cases. Define

$$H(\omega, \underline{p}, \bar{p}; A_b, A_s) = [p_L(\omega; A_b, A_s), p_H(\omega; A_b, A_s)] \cap [\underline{p}, \bar{p}].$$

If $H \neq \emptyset$, the parties can avoid hold-up by choosing a price in H ; such a price is consistent with the date 0 contract and gives neither party an incentive to hold up the other. However, if $H = \emptyset$, hold-up cannot be avoided.

Start with the first case. Even though no hold-up occurs, there will be disagreement about the appropriate outcome within the contract. In the spirit of Hart and Moore (2008), we assume that each party is aggrieved to the extent that he does not receive what he feels entitled to and will respond by shading, that is, cutting back on helpful actions. We assume 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; that is, B doesn't expect to pay less than p_L or S to receive more than p_H . In other words, B feels entitled to the lowest price in H , $\text{Max}(p_L, \underline{p})$, and S feels entitled to the highest price in H , $\text{Min}(p_H, \bar{p})$. Note that the assumption

that entitlements are constrained by what is feasible simplifies the analysis but is not crucial.

To simplify matters, we assume that the parties split the difference and set

$$(32) \quad \hat{p} = \frac{1}{2} \{ \text{Max}(p_L(\omega; A_b, A_s), \underline{p}) + \text{Min}(p_H(\omega; A_b, A_s), \bar{p}) \}.$$

As part of this deal, the parties agree to undertake the contractible helpful actions. However, each party cuts back on the noncontractible helpful actions in proportion to his aggrievement. *B* is aggrieved by

$$(33) \quad a_b = \hat{p} - \text{Max}(p_L(\omega; A_b, A_s), \underline{p})$$

and shades to the point where *S*'s payoff falls by θa_b . *S* is aggrieved by

$$(34) \quad a_s = \text{Min}(p_H(\omega; A_b, A_s), \bar{p}) - \hat{p}$$

and shades to the point where *B*'s payoff falls by θa_s . The parameter θ is taken to be exogenous, and the same for *B* and *S*, and $0 < \theta \leq 1$.

Thus in Case 1, where $H \neq \emptyset$, net surplus is

$$(35) \quad \begin{aligned} W_1(\omega, \underline{p}, \bar{p}; A_b, A_s) &= v - c - \theta(a_b + a_s) \\ &= v - c - \theta \{ \text{Min}(p_H(\omega; A_b, A_s), \bar{p}) \\ &\quad - \text{Max}(p_L(\omega; A_b, A_s), \underline{p}) \}. \end{aligned}$$

In contrast, in Case 2, where $H = \emptyset$, hold-up occurs, followed by renegotiation, and net surplus is given by

$$(36) \quad W_2(\omega) = (1 - \lambda)(v - c).$$

Note that, because hold-up leads to the withdrawal of all noncontractible helpful actions, whereas shading leads to the withdrawal of only some, there is an implicit constraint that total shading costs cannot exceed total hold-up costs; that is,

$$(37) \quad W_1(\omega, \underline{p}, \bar{p}; A_b, A_s) \geq W_2(\omega).$$

Fortunately, this constraint is automatically satisfied, given $0 < \theta \leq 1$, because

$$\begin{aligned}
 (38) \quad \text{Min}(p_H(\omega; A_b, A_s), \bar{p}) - \text{Max}(p_L(\omega; A_b, A_s), \underline{p}) \\
 \leq p_H(\omega; A_b, A_s) - p_L(\omega; A_b, A_s) \\
 = \lambda(v - c),
 \end{aligned}$$

where we are using (10). In other words, however large the price range $[\underline{p}, \bar{p}]$ is, net surplus is higher if hold-up is avoided than if it occurs.¹⁵

An optimal contract maximizes expected net surplus. Thus, an optimal contract solves

$$(39) \quad \text{Max}_{(A_b, A_s, \underline{p}, \bar{p})} \left\{ \int W_1(\omega, \underline{p}, \bar{p}; A_b, A_s) dF(\omega) + \int W_2(\omega) dF(\omega) \right\},$$

$$\left. \begin{aligned}
 H(\omega, \underline{p}, \bar{p}; A_b, A_s) \neq \emptyset & \quad H(\omega, \underline{p}, \bar{p}; A_b, A_s) = \emptyset \\
 \end{aligned} \right\}$$

where F is the distribution function of ω . (We assume that F has bounded support.)¹⁶ The trade-off is the following: As \underline{p} falls or \bar{p} rises, the set H becomes larger, and so hold-up is less likely. This is good, given that hold-up reduces surplus, that is, $W_1 \geq W_2$. However, shading represented by $\theta\{\text{Min}(p_H, \bar{p}) - \text{Max}(p_L, \underline{p})\}$ rises, which means that surplus in the absence of hold-up falls; that is, W_1 is lower. This is bad.

It is useful to analyze the optimal contract in some simple cases. We have already seen in Section II that a simple contract can achieve the first-best if there is no uncertainty. It turns out that a nonsimple contract can achieve the first-best if there are just two states: $\omega = \omega_1$ or ω_2 . (In what follows, we suppress assets.) To see why, note that there are two possibilities. Either $[p_L(\omega_1), p_H(\omega_1)] \cap [p_L(\omega_2), p_H(\omega_2)]$ is nonempty, or it is empty. In the first case, choose any price \hat{p} in the intersection and set $\underline{p} = \hat{p} = \bar{p}$. 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 $\underline{p} = p_H(\omega_1)$, $\bar{p} = p_L(\omega_2)$. In state ω_1 , $p = p_H(\omega_1)$, and in state ω_2 , $p = p_L(\omega_2)$. Hold-up is avoided, and there is no aggrievement, because the set H is a singleton in both states.

15. In fact, (A2) implies the further constraint that each party's shading costs cannot exceed $(1/2)\lambda(v - c)$. This is automatically satisfied given (32).

16. It is easy to show that an optimal contract exists, because the objective function is upper semicontinuous in $[\underline{p}, \bar{p}]$.

Once there are three states, it is easy to show that the first-best typically cannot be achieved even with a nonsimple contract.

One obvious question that can be asked is, can the parties do better than specify a price range? For example, would it help for them to play a message game that revealed the state of the world? Or would it be useful for them to agree to an informal state-contingent contract based on observable but unverifiable information? We refer the reader to Hart and Moore (2008), who argue that, at least under some assumptions, the answer is no.

A price range has several interpretations. One is the following. Suppose date 1 trade is not instantaneous: rather the parties trade over a period of time, for example, between dates 1 and 2. (For simplicity, suppose uncertainty is still resolved at date 1⁻.) Then the parties might agree at date 0 on the *length* of their contract. For example, they might agree on a trading price that will operate for a fraction τ of the period between dates 1 and 2. The parties recognize that when they get to date 1⁻, they will renegotiate the price for the remaining fraction $(1 - \tau)$ of the period. A small value of τ then corresponds to a flexible contract—there is a wide range of possible (average) prices over the period from an *ex ante* perspective—whereas a large value of τ corresponds to a rigid contract.

This interpretation can throw light on Goldberg and Erickson's finding (1987) that parties in the petroleum coke industry tend to write shorter-term contracts in a more volatile environment. The parallel result in our model is that an increase in uncertainty leads to a larger price range $[\underline{p}, \bar{p}]$. Although further assumptions would be required to prove a general result along these lines, the result is at least true at the extremes. With no uncertainty, the optimal contract is a single price $p \in [p_L(\omega; A_b, A_s), p_H(\omega; A_b, A_s)]$, whereas with sufficiently large uncertainty, the optimal contract will be a nondegenerate price interval, because the chance that a single price will lie in $[p_L(\omega; A_b, A_s), p_H(\omega; A_b, A_s)]$ becomes vanishingly small. Thus, in a broad sense, our model seems consistent with Goldberg and Erickson (1987). Note that Gray (1978) also explains why parties write shorter-term contracts in a more volatile environment; however, in her model contracts are not optimal.

Proof of Proposition 2. We prove (1). Let $(A_b, A_s, [\underline{p}, \bar{p}])$ be an optimal contract. The proof proceeds in two steps. We first replace $[\underline{p}, \bar{p}]$ by another price interval and show that shading

costs fall (weakly). We then allocate all the assets to B , make another change in the price interval, and show that shading costs fall again and that the hold-up region becomes (weakly) smaller. Thus, the new contract in which B owns everything must also be optimal.

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

$$(40) \quad [p_L(v), p_H(v)] \cap [\underline{p}, \bar{p}] \neq \emptyset$$

for $v = \underline{v}$ and $v = \bar{v}$. Because $p_L(v), p_H(v)$ are increasing in v , (40) must also hold for $\underline{v} \leq v \leq \bar{v}$; that is, hold-up does not occur for intermediate v 's. Note that (40) implies that $p_H(\underline{v}) \geq \underline{p}, p_L(\bar{v}) \leq \bar{p}$.

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

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

Clearly, $\underline{p}' \geq \underline{p}$. Also, either $\bar{p}' \leq \bar{p}$ or $\bar{p}' = p'$. 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,

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

for $\underline{v} \leq v \leq \bar{v}$. Hence, the new price interval avoids hold-up for $\underline{v} \leq v \leq \bar{v}$, just like the old one. In addition, grievement 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 ; that is, set $A_b = A, A_s = \emptyset$. Call this the new ownership structure. Define a new price interval $[\underline{p}'', \bar{p}'']$ given by

$$(41) \quad \underline{p}'' = p_H^N(\underline{v}), \quad \bar{p}'' = \text{Max}(p_L^N(\bar{v}), p_H^N(\underline{v})),$$

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

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

lower than those under the old ownership structure and $[p, \bar{p}]$. That is, we demonstrate that

$$(42) \quad \begin{aligned} & \text{Min}(p_H^N(v), \bar{p}'') - \text{Max}(p_L^N(v), \underline{p}'') \\ & \leq \text{Min}(p_H(v), \bar{p}') - \text{Max}(p_L(v), \underline{p}'). \end{aligned}$$

There are several cases to consider. Note first that if $\underline{p}' = \bar{p}' = p_H(v) \geq p_L(\bar{v})$, that is, the right-hand side (RHS) of (42) is zero, then

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

where we are using the fact that $p_H^N(\bar{v}) - p_H^N(v) \leq p_H(\bar{v}) - p_H(v)$ because $\partial p_H / \partial v$ falls the more assets B owns (by (29)), and $p_H^N(\bar{v}) - p_L^N(\bar{v}) = p_H(\bar{v}) - p_L(\bar{v})$, that is, $p_H - p_L$ is independent of the ownership structure (see (10)). Hence, $p_H^N(v) \geq p_L^N(\bar{v})$. It follows from (41) that $\underline{p}'' = \bar{p}'' = p_H^N(v)$, and so the left-hand side (LHS) of (42) is zero. Therefore, (42) holds.

Consider next the case where $\underline{p}' = p_H(v) < \bar{p}' = p_L(\bar{v})$. If $\underline{p}'' = \bar{p}'' = p_H^N(v) \geq p_L^N(\bar{v})$, (42) again holds. So suppose $\underline{p}'' = p_H^N(v) < p_L^N(\bar{v}) = \bar{p}''$. We must show that

$$(43) \quad \begin{aligned} & \text{Min}(p_H^N(v), p_L^N(\bar{v})) - \text{Max}(p_L^N(v), p_H^N(v)) \\ & \leq \text{Min}(p_H(v), p_L(\bar{v})) - \text{Max}(p_L(v), p_H(v)). \end{aligned}$$

We can rewrite (43) as

$$(44) \quad \begin{aligned} & \text{Min}\{p_H^N(v) - p_H^N(v), p_L^N(\bar{v}) - p_H^N(v), \\ & \quad p_H^N(v) - p_L^N(v), p_L^N(\bar{v}) - p_L^N(v)\} \leq \text{Min}\{p_H(v) - p_H(v), p_L(\bar{v}) - p_H(v), \\ & \quad p_H(v) - p_L(v), p_L(\bar{v}) - p_L(v)\}. \end{aligned}$$

To establish (44), one shows that each component in the min formula on the LHS of (44) is no greater than the corresponding component on the RHS of (44). This follows from the facts that $\partial p_H / \partial v, \partial p_L / \partial v$ are nonincreasing in the assets that B owns and that $p_H - p_L$ is independent of ownership structure for a given v . Hence, (43) holds, and so does (42).

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

Proof of Proposition 3. 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 $(A_{br}, A_{sr}, \underline{p}_r, \bar{p}_r)$ such that $a \in A_{sr}$ for all r ; that is, S owns asset a , and $\pi_r \rightarrow 1$ as $r \rightarrow \infty$. Without loss of generality (WLOG), suppose that $A_{br} \rightarrow A_b(k)$, $A_{sr} \rightarrow A_s(k)$, $\underline{p}_r \rightarrow \underline{p}(k)$, and $\bar{p}_r \rightarrow \bar{p}(k)$. Then $(A_b(k), A_s(k), \underline{p}(k), \bar{p}(k))$ 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, because there is almost no uncertainty. A necessary condition for this is that there is a single trading price $p(k)$ in the limit, that is, $\underline{p}(k) = p(k) = \bar{p}(k)$ (so that shading costs are zero), and

$$(45) \quad \begin{aligned} & p_L(\omega, A_b(k), A_s(k)) \\ & \leq \underline{p}(k) = p(k) = \bar{p}(k) \leq p_H(\omega, A_b(k), A_s(k)) \end{aligned}$$

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 (15) and (16) and are indexed by the limiting ownership structure.

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

$$(46) \quad \begin{aligned} \underline{p}'_r - \underline{p}_r &= \bar{p}'_r - \bar{p}_r \\ &= \frac{1}{2} [\alpha_s(A_{sr} \setminus \{a\}) - \alpha_s(A_{sr}) \\ & \quad + \alpha_b(A_{br}) - \alpha_b(A_{br} \cup \{a\}) \\ & \quad + \beta_b(A_{br})v_0 - \beta_b(A_{br} \cup \{a\})v_0] \\ &= \Delta_r(k). \end{aligned}$$

In other words, we adjust $\underline{p}_r, \bar{p}_r$ by an amount equal to the change $\Delta p_L, \Delta p_H$ in p_L, p_H that occurs as a result of the shift in ownership structure, where $\Delta p_L, \Delta p_H$ 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 $\underline{p}_r, \bar{p}_r, p_{Lr}(\omega), p_{Hr}(\omega)$ 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.

Because 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 $r \rightarrow \infty$. WLOG $(A'_{br}, A'_{sr}, \underline{p}'_r, \bar{p}'_r) \rightarrow (A'_b(k), A'_s(k), \underline{p}'(k), \bar{p}'(k))$ and $\Delta_r(k) \rightarrow \Delta(k)$. Given (45) and (46), we must have

$$p_L(\omega, A'_b(k), A'_s(k)) \leq \underline{p}'(k) = p'(k) = \bar{p}'(k) \leq p_H(\omega, A'_b(k), A'_s(k)) \tag{47}$$

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

We show that this conclusion is false. Because 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. Because we are in event 2, index the state by (v, φ) . Then, from (46),

$$\begin{aligned} p'(k) - p(k) &= p_L((v_0, \varphi), A'_b(k), A'_s(k)) - p_L((v_0, \varphi), A_b(k), A_s(k)) \\ &= p_H((v_0, \varphi), A'_b(k), A'_s(k)) - p_H((v_0, \varphi), A_b(k), A_s(k)) \end{aligned} \tag{48}$$

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 (47) 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 $v > v_0$,

$$\begin{aligned}
 & p_L((v, \varphi), A'_b(k), A'_s(k)) - p_L((v_0, \varphi), A'_b(k), A'_s(k)) \\
 &= \frac{1}{2}[(1 - \lambda) - \beta_b(A_b \cup \{a\})](v - v_0) \\
 (49) \quad &< \frac{1}{2}[(1 - \lambda) - \beta_b(A_b)](v - v_0) \\
 &= p_L((v, \varphi), A_b(k), A_s(k)) \\
 &\quad - p_L((v_0, \varphi), A_b(k), A_s(k)),
 \end{aligned}$$

because a is idiosyncratic to B . From (48) and (49), we may conclude that $p'(k) < p_L((v, \varphi), A'_b(k), A'_s(k)) \Rightarrow p(k) < p_L((v, \varphi), A_b(k), A_s(k))$, that is, 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 3, which ensures that $p_L((v, \varphi), A_b(k), A_s(k)) > p_H((v_0, \varphi), A_b(k), A_s(k))$ 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. QED

Proof of Proposition 4. We sketch the proof, since the argument is very similar to that of Proposition 3. Suppose joint ownership is not optimal. For small k , choose a sequence of optimal contracts as $\pi \rightarrow 1$. The limiting contract is optimal for event 1. Hence, (45) is satisfied. Consider a new sequence of contracts where all assets are jointly owned and $\underline{p}_r, \bar{p}_r$ are adjusted to reflect the new ownership structure, that is,

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

Then surplus does not change in event 1. Because the initial contract is optimal, surplus must weakly fall in events 2 or 3. WLOG suppose it falls in event 2. Take limits as $r \rightarrow \infty$. 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. QED

HARVARD UNIVERSITY

REFERENCES

- Baker, George, Robert Gibbons, and Kevin Murphy, "Relational Contracts and the Theory of the Firm," *Quarterly Journal of Economics*, 117 (2002), 39–84.
- Bewley, Truman F., *Why Wages Don't Fall during a Recession* (Cambridge, MA: Harvard University Press, 1999).
- Card, David, "An Empirical Model of Wage Indexation Provisions in Union Contracts," *Journal of Political Economy*, 94 (1986), S144–S175.
- Corbin, Arthur L., *Corbin on Contracts* (St. Paul, MN: West Publishing Co., 1993).
- Goetz, Charles J., and Robert E. Scott, "The Mitigation Principle: Toward a General Theory of Contractual Obligation," *Virginia Law Review*, 69 (1983), 967–1024.
- Goldberg, Victor, "Price Adjustment in Long-Term Contracts," *Wisconsin Law Review*, 3 (1985), 527–543.
- , *Framing Contract Law* (Cambridge, MA: Harvard University Press, 2006).
- Goldberg, Victor, and John R. Erickson, "Quantity and Price Adjustment in Long-Term Contracts: A Case Study of Petroleum Coke," *Journal of Law and Economics*, 30 (1987), 369–398.
- Gray, Jo Anna, "On Indexation and Contract Length," *Journal of Political Economy*, 86 (1978), 1–18.
- Grossman, Sanford, and Oliver Hart, "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy* 94 (1986), 691–719.
- Halonen, Maija, "Reputation and the Allocation of Ownership," *The Economic Journal*, 112 (2002), 539–558.
- Hart, Oliver, and John Moore, "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98 (1990), 1119–1158.
- , "Incomplete Contracts and Ownership: Some New Thoughts," *American Economic Review*, 97 (2007), 182–186.
- , "Contracts as Reference Points," *Quarterly Journal of Economics*, 123 (2008), 1–48.
- Januszewski Forbes, Silke, and Mara Lederman, Does Vertical Integration Affect Firm Performance? Evidence from the Airline Industry, unpublished, UCSD, 2007.
- Joskow, Paul L., "Vertical Integration and Long-Term Contracts: The Case of Coal-Burning Electric Generating Plants," *Journal of Law, Economics, & Organization*, 1 (1985), 33–80.
- , "Price Adjustment in Long Term Contracts: The Case of Coal," *Journal of Law and Economics*, 31 (April 1988), 47–83.
- Klein, Benjamin, "Why Hold-Ups Occur: The Self-Enforcing Range of Contractual Relationships," *Economic Inquiry*, 34 (1996), 444–463.
- Klein, Benjamin, R. Crawford, and A. Alchian, "Vertical Integration, Appropriate Rents, and the Competitive Contracting Process," *Journal of Law and Economics*, 21 (1978), 297–326.
- Kube, Sebastian, Michel André Maréchal, and Clemens Puppe, Do Wage Cuts Damage Work Morale? Evidence from a Natural Field Experiment, unpublished, University of Karlsruhe, Germany, 2007.
- Lafontaine, Francine, and Scott Masten, "Contracting in the Absence of Specific Investments and Moral Hazard: Understanding Carrier-Driver Relations in U.S. Trucking," NBER Working Paper No. W8859, 2002.
- Lafontaine, Francine, and Margaret Slade, "Vertical Integration and Firm Boundaries: The Evidence," *Journal of Economic Literature*, 45 (2007), 629–685.
- MacLeod, Bentley, and James Malcolmson, "Investments, Holdup and the Form of Market Contracts," *American Economic Review*, 83 (1993), 811–837.

- Maskin, Eric, and Jean Tirole, "Unforeseen Contingencies and Incomplete Contracts," *Review of Economic Studies*, 66 (1999), 83–114.
- Masten, Scott E., "Equity, Opportunism and the Design of Contractual Relations," *Journal of Institutional and Theoretical Economics*, 144 (1988), 180–195.
- , Long-Term Contracts and Short-Term Commitment: Price Determination for Heterogeneous Freight Transactions, unpublished, University of Michigan, 2007.
- Rajan, Raghuram G., and Luigi Zingales, "Power in a Theory of the Firm," *Quarterly Journal of Economics*, 113 (1998), 361–386.
- Simon, H., "A Formal Theory of the Employment Relationship," *Econometrica*, 19 (1951), 293–305.
- Wernerfelt, Birger, "On the Nature and Scope of the Firm: An Adjustment-Cost Theory," *Journal of Business*, 70 (1997), 489–514.
- Williamson, Oliver, "The Vertical Integration of Production: Market Failure Considerations," *American Economic Review*, 61 (1971), 112–123.