On the Transaction Cost Determinants of Vertical Integration

Michael D. Whinston
Northwestern University and NBER

The transaction cost approach to the organization of firms has been among the most significant advances in industrial organization in the last 25 years. Much of this work has taken the transaction cost economics view of Williamson (1975, 1979, 1985) and Klein, Crawford, and Alchian (1978) in which high levels of quasi-rents are taken to increase the likelihood of vertical integration. More recently, however, the more formal property rights theory of Grossman and Hart (1986) and Hart and Moore (1990) has received considerable attention as a theory of integration. This article explores the predictions of property rights theory to assess what the extensive supporting evidence on transaction cost economics tells us about the property rights theory's empirical relevance. The article concludes that this evidence sheds little light on the relevance of the property rights theory and discusses how we might try to learn more.

1. Introduction

The extensive development of the transaction cost approach to the organization of firms (and industries) has been, without question, among the most significant advances in industrial organization over the last 25 years. This development began with path breaking work by Williamson (1975, 1979, 1985), Klein, Crawford, and Alchian (1978), and others that pushed forward the agenda of explaining firms' boundaries that was first laid out by Coase (1937) 40 years earlier. Their work focused attention on how ex post quasi-rents could create hazards for long-term contractual relations when contracts are incomplete and the resulting effects on firms' choices of whether to integrate distinct stages in the vertical chain of production and distribution. By doing so, transaction cost economics (henceforth "TCE") identified a pervasive—and potentially measurable—feature of long-term contracting

I thank Robert Gibbons, Oliver Hart, Scott Masten, Ilya Segal, Yossi Spiegel, Steve Tadelis, Oliver Williamson, and two anonymous referees, as well as seminar audiences at the November 1997 "Make Versus Buy" Conference at Columbia Law School, the January 2001 ASSA meetings, the May 2001 CSIO-IDEI Industrial Organization conference, and the Harvard-MIT Organizations Workshop for their comments. Adam Rosen provided excellent research assistance. I also gratefully acknowledge the financial support and hospitality of the George J. Stigler Center for the Study of the Economy and the State at the University of Chicago and financial support from the NSF.

settings as a critical determinant of the degree to which firms would choose to integrate activities.

Just as one would hope, this conceptual breakthrough was soon followed by empirical work aimed at testing the theory. In a typical study, some measure of ex post quasi-rents, such as the specificity of the product being procured or investments being made, was related to the choice of whether to integrate. Very quickly it became clear that the theory had significant predictive power. Indeed, as early as 1987, Joskow (1988) could write that "This work generally provides strong empirical support for the importance of transactions cost considerations, especially the importance of asset specificity in explaining vertical relationships." Developments since have not changed this conclusion.[1]

At the same time that empirical work was providing this confirmation, a closely related and more formal theory of vertical integration emerged, beginning with Grossman and Hart's (1986) seminal article (see also Hart and Moore, 1990, and Hart, 1995). Like the TCE, this "property rights theory" (henceforth "PRT") takes the incompleteness of contracts and development of ex post quasi-rents as critical to understanding vertical integration. It then focuses on how ownership of physical assets, which confers residual rights of control over these assets, alters the efficiency of trading relationships.

Because the two theories both focus on contractual incompleteness and ex post quasi-rents, it is often presumed that the empirical literature on transaction cost determinants of vertical integration provides support for both. The aim of this article is to consider the validity of this presumption. I show that it is unfounded. Although the two theories do share the prediction that integration decisions matter in the presence of incomplete contracts and ex post quasi-rents, in other respects their predictions differ substantially.

In Section 2 I lay out in detail how the PRT and TCE differ. For this purpose I focus on a simple linear-quadratic property rights model and examine the comparative statics predictions that emerge from it. While the TCE predicts that any increase in quasi-rents will increase the likelihood of vertical integration (a finding that is so far consistent with nearly all of the existing empirical literature), I show that the PRT offers much more refined predictions about the types of specificity that matter for integration decisions, and the direction in which various types of specificity move the likelihood of integration.

In Section 3 I discuss the findings of three of the most well-known empirical articles on the transaction cost determinants of vertical integration in light of these observations. My particular focus is on what these articles might tell us about the extent to which the PRT's predictions hold. The general lesson to be learned from this exercise is that these articles simply do not tell us enough about the economic environments under study to judge the PRT. As a

1. Masten (1996) provides a nice compilation of some of the best work on transaction cost determinants of vertical integration and vertical contracting. Shelanski and Klein (1995) survey the area.

result, it is relatively easy to construct a variety of reasonable property rights models of these settings that either do or do not agree with the observed relationships in the data. The implication of this fact is that the existing empirical literature provides little guidance regarding the applicability of the PRT.

Section 4 offers concluding remarks. There I discuss some of the issues involved in constructing tests of the PRT. I also argue that there are good reasons (highlighted by the PRT) to doubt that greater quasi-rents are universally associated with an increase in the likelihood of vertical integration. This suggests that the TCE could itself be usefully put to more demanding tests.

2. The TCE and the PRT

The TCE takes as its starting point two observations: First, that in many exchange relationships it is difficult to specify fully the contracting parties' obligations and, second, that in many cases the parties become "locked in" to one another to some extent over the course of their relationship (i.e., the parties develop "relationship-specific assets"). Using Klein, Crawford, and Alchian's (1978) terminology, there are then ex post quasi-rents, because the value of trade within the relationship comes to exceed the value of outside trading opportunities.[2] When contracts are incomplete and ex post quasi-rents are present, the contracting parties face the hazard of ex post opportunistic behavior: each party to the contract may engage in inefficient behavior in an attempt to "hold up" the other party and obtain a larger share of the available quasi-rents.[3] This reduction in the efficiency of the trading relationship, if large enough, may motivate one of the contracting parties to bring the transaction "in house"—that is, to vertically integrate—to mitigate these hazards. Doing so leads investment and trading decisions to be made in a coordinated fashion (although possibly with increased bureaucratic costs). A central prediction of the theory—and the one that the empirical literature has almost uniformly focused on—is then that when contracts are incomplete, greater levels of quasi-rents increase the likelihood of vertical integration.[4]

2. Williamson cites four kinds of specificity as giving rise to such conditions: site specificity (the parties have committed immobile physical assets to a "cheek-by-jowl" relationship), physical asset specificity (the parties must make investments in specialized physical assets), human asset specificity (the parties develop specific human capital), and dedicated assets (one or both parties build productive capacity for which there is insufficient demand absent their trading relationship).

3. As Masten (1984) puts it: "Idiosyncratic assets, because of their specialized and durable nature, imply that parties to a transaction face only imperfect exchange alternatives for an extended period. The more specialized those assets, the larger will be the quasi-rents at stake over that period, and hence the greater the incentive for agents to attempt to influence the terms of trade through bargaining or other rent-seeking activities once the investments are in place."

4. Joskow (1988) puts this prediction succinctly as follows: "Other things equal, we expect the parties to more frequently choose vertical integration or a long-term contract as the quasi-rents associated with specific investments become more important and the associated benefits of precommitment increase."

The PRT, developed in Grossman and Hart (1986), Hart and Moore (1990), and elsewhere (see Hart, 1995, for an overview), starts from the same premises of incomplete contracts and ex post quasi-rents. It differs from the TCE, however, in at least three ways. The first is methodological rather than substantive: the PRT is substantially more formal than the (largely verbal) TCE.[5] Second, the PRT focuses explicitly on distortions in ex ante investments, in contrast to the ex post haggling and maladaptation costs that are a major focus of the TCE.[6] Third, in contrast to the TCE, which assumed that opportunism could be mitigated by bringing the transaction within the firm (with resulting bureaucratic costs), the PRT assumes that this hazard is present in all organizational modes. This is so because investment and trading decisions remain fundamentally decentralized in the PRT, regardless of the structure of asset ownership. Integrated asset ownership changes incentives, but does not result in coordinated investments as in the TCE. (Indeed, this results in a theory in which the identity of the owner of integrated assets—i.e., buyer versus seller—can matter for the performance of the "integrated firm.") Hence, in the PRT, the integration decision involves a comparison of the efficiency costs of opportunistic behavior in the various possible organizational forms. Although the "specificity" of investments is important in this theory, the types of specificity that are important and the predicted direction of their effects can differ substantially from the TCE.

To illustrate these points, in the next subsection I examine the determinants of integration in the context of a simple linear-quadratic property rights model.

2.1 A Simple Property Rights Model

Consider a simple bilateral trade setting. There are two agents who may wish to trade tomorrow, a buyer B and a seller S. The seller S uses an asset for the production of his product. In what follows, I focus on the incentives for buyer integration; that is, for B to own the upstream asset. I denote by $A_B = 1$ the case in which B owns the upstream asset (vertical integration), and by $A_B = 0$ the case in which S owns this asset (nonintegration).[7]

5. For a recent article providing a more formal model of some of the ex post adaptation issues discussed in the TCE, see Bajari and Tadelis (2001).

6. Williamson (2000) puts it this way: "The most consequential difference between the TCE and GHM [Grossman–Hart–Moore] setups is that the former holds that maladaptation in the contract execution interval is the principal source of inefficiency, whereas GHM vaporize ex post maladaptation by their assumptions of common knowledge and costless bargaining." It should be noted that the PRT's focus on ex ante investments seems mostly a matter of modeling convenience, since residual rights of control could also affect the efficiency of bargaining. For a recent article examining the effect of ownership on ex post bargaining distortions, see Matouschek (2001). It should also be noted that the TCE literature does recognize that ex ante investment distortions are a potential cost of ex post opportunism (see, e.g., Klein et al., 1978:301).

7. In principle, other ownership forms are possible, such as joint ownership, random ownership, or mechanisms that assign ownership as a function of messages sent by the parties (options to own being a simple example). I do not consider such schemes here.

Since I have in mind a situation in which there are likely to be downstream assets as well, there is also in principle a possibility of seller integration in which S owns both the downstream and the upstream assets. Implicitly, in what follows I am assuming that any observed integration is buyer integration, in which B controls both the upstream and downstream assets. In large part I do this because the empirical studies I discuss in Section 3 focus on buyers' decisions of whether to procure internally (e.g., General Motors's [GM's] decision of whether to own its own engine plant). In two of these three studies, this may be justified by the fact that the buyers are making these decisions over the procurement of many inputs so that supplier ownership of the buyer's assets is not generally a viable option.[8–10]

At time 0, the two parties can decide who will own the upstream asset (e.g., B may purchase it from S if S initially owns it). Also at that time they may agree about the levels of some contractible investments. For simplicity, I assume that for the relationship to be productive, there are certain given levels of these investments that must be made (so that this is not a choice variable for the parties, although the appropriate levels will need to be specified contractually ex ante). As in Hart (1995), no other agreements about trade are possible because the good to be traded is not describable in advance.

At time 1, each of the parties may make some noncontractible investments. We denote by $i_B \in \Re_+$ the buyer's noncontractible investment level and by $i_S \in \Re_+$ the seller's noncontractible investment level. The associated costs are $c_B(i_B)$ and $c_S(i_S)$.

At time 2, the buyer and the seller bargain over trade. Denote by $\pi(i_B, i_S)$ the profits available from the efficient trade, and by $w_B(i_B, i_S|A_B)$ the payoff to the buyer in his next-best alternative to trading with S given that investment levels i_B and i_S have been taken, and that the ownership variable is A_B; that

8. Of course, this feature would ideally be part of a more complete model of the integration decision than what I present here. (For one such model, see Bolton and Whinston, 1993.)

9. In situations in which one does not know on a priori grounds the form that any observed integration takes, an immediate issue in applying the property rights model empirically is whether this difference in integrated structures is empirically distinguishable. If so, then one would want to allow for a choice between three possible ownership structures. If not, then the property rights model can be used instead to predict the probability that *some* integrated outcome is observed. I comment briefly in note 13 on some of the implications of both B and S integration being possible.

10. Note that this also raises the question of the degree to which an integrated firm is in fact able to design the allocation of control, or authority, within the firm; that is, whether "B ownership" and "S ownership" are really the only two possible forms of authority within the firm. Indeed, one possibility is that an integrated firm can fully coordinate its behavior—regardless of the identity of its owner—as assumed in the TCE. Reality most likely lies somewhere between (see, e.g., Aghion and Tirole, 1997). This appears to be an important open issue in the PRT, but I shall follow the standard PRT framework here. Arguably this issue makes the PRT as currently formulated more plausible as a model for integration of simple firms, or for the determination of the ownership assignment of simple assets (such as in the Baker and Hubbard, 1999, article I discuss in the conclusion) than for integration decisions of complex firms.

is, B's disagreement payoffs. When $A_B = 0$, this involves procurement from another supplier (or some form of self-production) or shutting down; when $A_B = 1$, this could involve either procurement from another supplier, hiring another manager to come in and produce the input using the upstream asset owned by B, or shutting down. Likewise, let $w_S(i_B, i_S|A_B)$ denote the payoff to the seller in his next-best alternative to trading with B given that investment levels i_B and i_S have been taken, and that the ownership variable is A_B. This will involve selling to another buyer when $A_B = 0$ or shutting down; when $A_B = 1$, this could involve either selling to another buyer using a technology that does not use the asset, running another firm, or shutting down. I assume that the parties engage in Nash bargaining and (for simplicity) that they have equal bargaining power; hence they split any available surplus (the quasi-rents) evenly.[11]

Note that in contrast to Hart and Moore (1990), I allow for an agent's investment to affect not only his own disagreement payoffs but also those of the other contracting party—that is, we consider "cross-investments" (or, as Che and Hausch, 1999, call them, "cooperative" investments) as well as "self-investments." For example, S might invest in training B how to produce more efficiently. It also allows for investments in physical capital, in contrast to the human capital investments considered in Hart and Moore (1990). For example, S might invest in raising the quality of the goods that can be produced using the upstream asset. This would affect S's disagreement payoff if he owned the asset, but would affect B's disagreement payoff if instead B owned the asset. I do this because this generalization seems indispensible if the PRT is to be a serious contender in explaining patterns of integration.

We make the following assumptions about these functions:

$$\begin{aligned}
\pi(i_B, i_S) &= \alpha_0 + \alpha_B i_B + \alpha_S i_S \\
w_B(i_B, i_S|A_B) &= [\beta_0 + \beta_{B0} i_B + \beta_{S0} i_S](1 - A_B) + [\beta_1 + \beta_{B1} i_B + \beta_{S1} i_S]A_B \\
w_S(i_B, i_S|A_B) &= [\sigma_0 + \sigma_{S0} i_S + \sigma_{B0} i_B](1 - A_B) + [\sigma_1 + \sigma_{S1} i_S + \sigma_{B1} i_B]A_B \\
c_B(i_B) &= \frac{1}{2}(i_B)^2 \\
c_S(i_S) &= \frac{1}{2}(i_S)^2,
\end{aligned}$$

where $(\alpha_B, \alpha_S) \geq 0$ (so that investments are [weakly] productive) and $\alpha_0 \geq \max\{\beta_0 + \sigma_0, \beta_1 + \sigma_1\}$, $\alpha_B \geq \max\{\beta_{B0} + \sigma_{B0}, \beta_{B1} + \sigma_{B1}\}$, and $\alpha_S \geq \max\{\sigma_{S0} + \beta_{S0}, \sigma_{S1} + \beta_{S1}\}$ (these three conditions imply that it is always efficient for the two parties to trade ex post; i.e., quasi-rents are always present). Let $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\sigma}$ denote the respective parameter vectors. This linear-quadratic formulation has the advantage that we can identify in a very simple way the marginal returns on investment. The assumptions adopted in Chapter 2

11. The comparative statics effects derived below (such as those illustrated in Tables 1 and 2) continue to hold (in sign) for any bargaining share $\lambda \in [0, 1]$ received by the buyer.

of Hart (1995) correspond to the case in which $\beta_{S0} = \beta_{S1} = \sigma_{B0} = \sigma_{B1} = 0$, $\alpha_B > \beta_{B1} > \beta_{B0} \geq 0$, and $\alpha_S > \sigma_{S0} > \sigma_{S1} \geq 0$.

2.1.1 The First Best. Consider first the efficient levels of investments. These solve

$$\max_{i_B, i_S} \pi(i_B, i_S) - c_B(i_B) - c_S(i_S) = \alpha_0 + \alpha_B i_B + \alpha_S i_S - \frac{1}{2}(i_B)^2 - \frac{1}{2}(i_S)^2. \tag{1}$$

The solution to this problem is $(i_B^{**}, i_S^{**}) = (\alpha_B, \alpha_S)$, and the resulting joint surplus is $W^{**} = \alpha_0 + \frac{1}{2}(\alpha_B)^2 + \frac{1}{2}(\alpha_S)^2$.

2.1.2 Equilibrium Investment Levels. Now consider the equilibrium levels of investment undertaken by the parties given the ownership variable A_B. The buyer chooses i_B to solve

$$\max_{i_B} w_B(i_B, i_S|A_B) + \frac{1}{2}[\pi(i_B, i_S) - w_B(i_B, i_S|A_B) - w_S(i_B, i_S|A_B)] - c_B(i_B)$$

or

$$\max_{i_B} \frac{1}{2}[\pi(i_B, i_S) + w_B(i_B, i_S|A_B) - w_S(i_B, i_S|A_B)] - c_B(i_B).$$

This gives first-order condition (assuming $i_B > 0$)

$$\frac{1}{2}\left[\frac{\partial \pi(i_B, i_S)}{\partial i_B} + \frac{\partial w_B(i_B, i_S|A_B)}{\partial i_B} - \frac{\partial w_S(i_B, i_S|A_B)}{\partial i_B}\right] = \frac{\partial c_B(i_B)}{\partial i_B}.$$

Substituting using our assumed functional forms and solving we get

$$i_B^*(A_B; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\sigma}) = \frac{1}{2}\big[\alpha_B + (\beta_{B0} - \sigma_{B0})(1 - A_B) + (\beta_{B1} - \sigma_{B1})A_B\big]. \tag{2}$$

In a parallel fashion we can get the seller's equilibrium investment level as

$$i_S^*(A_B; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\sigma}) = \frac{1}{2}\big[\alpha_S + (\sigma_{S0} - \beta_{S0})(1 - A_B) + (\sigma_{S1} - \beta_{S1})A_B\big]. \tag{3}$$

We can then denote by $W(A_B; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\sigma})$ the equilibrium level of welfare given ownership structure A_B and parameters $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\sigma})$ [obtained by substituting Equations (2) and (3) into Equation (1)].

2.1.3 Comparative Statics Predictions of the Theory. Now consider how changes in various parameters effect the likelihood of vertical integration. To apply the model to actual data, we want to allow for the presence of other, unobserved factors affecting the integration decision. Here I do this by assuming that the actual joint surplus under ownership structure A_B is $W(A_B; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\sigma}) + \varepsilon_{A_B}$, where ε_{A_B} is a random variable unobserved to the econometrician. With this assumption, buyer integration will be observed if

$W(1;\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\sigma})-W(0;\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\sigma})\geq(\varepsilon_0-\varepsilon_1)$. Thus, for comparative statics, we are interested in how changes in the parameters affect the difference

$$\Delta(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\sigma})\equiv[W(1;\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\sigma})-W(0;\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\sigma})].$$

Note first that the values of $(\alpha_0,\beta_0,\sigma_0,\beta_1,\sigma_1)$ are *irrelevant*: thus, changes in the level of quasi-rents that do not affect the marginal returns on investments have *no bearing* on the likelihood of vertical integration. As one implication then, we see that in contrast to the predictions of the TCE, greater levels of specificity of *contractible* investments have no bearing on the integration decision as long as they do not affect the marginal returns from any noncontractible investments.

Let us now study how changes in the marginal returns to noncontractible investments affect the likelihood of observing vertical integration. Consider first a change in the marginal productivity of B's investment when B and S trade, α_B. For this change we have

$$\begin{aligned}\frac{\partial\Delta}{\partial\alpha_B}&=[i_B^*(1;\cdot)-i_B^*(0;\cdot)]+[\alpha_B-i_B^*(1;\cdot)]\frac{\partial i_B^*(1;\cdot)}{\partial\alpha_B}-[\alpha_B-i_B^*(0;\cdot)]\frac{\partial i_B^*(0;\cdot)}{\partial\alpha_B}\\&=\frac{1}{2}[i_B^*(1;\cdot)-i_B^*(0;\cdot)]\\&=\frac{1}{4}[(\beta_{B1}-\sigma_{B1})-(\beta_{B0}-\sigma_{B0})].\end{aligned}$$

Thus if B invests more under integration, then an increase in α_B—which increases the joint return from B's investment—increases the probability of integration. Likewise, if S invests more under nonintegration (i.e., when he owns the upstream asset), then an increase in α_S, the joint return from S's investment, *reduces* the probability of integration:

$$\begin{aligned}\frac{\partial\Delta}{\partial\alpha_S}&=\frac{1}{2}[i_S^*(1;\cdot)-i_S^*(0;\cdot)]\\&=\frac{1}{4}[(\sigma_{S1}-\beta_{S1})-(\sigma_{S0}-\beta_{S0})].\end{aligned}$$

Following similar derivations, we can determine the effects of changes in the marginal returns to investments on B and S's disagreement payoffs:

$$\begin{aligned}\frac{\partial\Delta}{\partial\beta_{B0}}&=-\frac{\partial\Delta}{\partial\sigma_{B0}}=-\frac{1}{2}[\alpha_B-i_B^*(0;\cdot)]\\\frac{\partial\Delta}{\partial\beta_{B1}}&=-\frac{\partial\Delta}{\partial\sigma_{B1}}=\frac{1}{2}[\alpha_B-i_B^*(1;\cdot)]\\\frac{\partial\Delta}{\partial\sigma_{S0}}&=-\frac{\partial\Delta}{\partial\beta_{S0}}=-\frac{1}{2}[\alpha_S-i_S^*(0;\cdot)]\\\frac{\partial\Delta}{\partial\sigma_{S1}}&=-\frac{\partial\Delta}{\partial\beta_{S1}}=\frac{1}{2}[\alpha_S-i_S^*(1;\cdot)].\end{aligned}$$

Table 1. Effects of Increases in Marginal Returns on the Probability of Integration (Underinvestment Case)

	Under:	
Investment by	Nonintegration	Integration
Self-Investments		
B	$-$ ($d\beta_{B0} > 0$)	$+$ ($d\beta_{B1} > 0$)
S	$-$ ($d\sigma_{S0} > 0$)	$+$ ($d\sigma_{S1} > 0$)
Cross-Investments		
B	$+$ ($d\sigma_{B0} > 0$)	$-$ ($d\sigma_{B1} > 0$)
S	$+$ ($d\beta_{S0} > 0$)	$-$ ($d\beta_{S1} > 0$)

Table 1 summarizes these effects for the case in which there is underinvestment under either vertical structure (so that $\alpha_j > i_j^*(A_B; \cdot)$ for $j = B, S$). In this case, an increase in the marginal effect of i_j on j's disagreement payoff under ownership structure A_B increases the likelihood of observing that ownership structure (denoted by a "+" in Table 1). In contrast, observe that an increase in i_j's effect on $-j$'s disagreement payoff under ownership structure A_B *reduces* this likelihood (denoted by a "−" in Table 1). The reason is straightforward: In a situation of underinvestment, an increase in investment levels under a particular ownership structure raises the surplus generated under that structure. When the marginal effect of i_j on j's disagreement payoff increases under an ownership structure, that increases j's optimal investment level under that structure; when the effect of i_j on $-j$'s disagreement payoff increases, however, this causes j to decrease his investment, lowering the joint surplus associated with that ownership structure. Note also that the results reverse if we are instead in a situation of overinvestment, such as occurs with rent-seeking activities, for which the marginal effect of investment on disagreement payoffs exceeds the marginal return on investment when the parties trade; that is, where (with self-investments) $\min\{\beta_{B0}, \beta_{B1}\} > \alpha_B$ and $\min\{\sigma_{S0}, \sigma_{S1}\} > \alpha_S$ (see, e.g., Rajan and Zingales, 1998).

These results can be usefully compared to Hart's (1995:49–55) discussion of the empirical content of the PRT. There he emphasizes two points: The first is that a party should be more likely to own an asset when he makes a more important investment decision. In contrast to Hart's findings, however, the model above shows that increases in an agent's marginal returns to investment have ambiguous effects, even if we restrict attention to the case of self-investments and underinvestment, as in Hart (1995). For example, it is easy to verify that in this case if we have equal-sized increases in α_B, β_{B0}, and β_{B1}, then there is *no effect* on the probability of observing integration. Thus by having β_{B0} increase by slightly more or slightly less than the other two parameters, we can move the probability of integration in either direction.

There are several reasons for this difference from Hart's result. First, when Hart makes an agent's investments "relatively unproductive" (i.e.,

less important) he actually *increases* the marginal returns of investments on disagreement payoffs (see Hart, 1995:44): in particular, because he makes the marginal returns on investments approach the marginal cost of investment, and because at the equilibrium investment level a weighted average of marginal returns equals marginal cost, the lower marginal returns on disagreement payoffs must actually be increased by this change. Arguably it is more natural to think of a more (less) important investment as involving an increase (decrease) in *all* marginal returns. Second, Hart adjusts marginal returns in a proportional way toward the marginal cost of investment rather than changing all returns equally. Finally, Hart only considers the limiting case in which an agent's investment becomes totally unproductive (because the marginal returns from it are approximately equal to its marginal cost).[12]

The analysis above does show, however, that this conclusion is warranted in the case of investments in *physical* capital, provided we are in a situation of underinvestment and ownership increases investment. For example, an increase in the importance of B's investment in improving the upstream asset would raise α_B, β_{B1}, and σ_{B0}, and each of these changes has the effect of increasing the likelihood of integration. Similarly an increase in the importance of S's investment in improving the upstream asset would reduce the likelihood of vertical integration.

The second implication that Hart focuses on is that increases in the complementarity of assets should increase the likelihood of them being co-owned. In the model above, an increase in complementarity can be viewed as reducing the marginal returns of investments on disagreement payoffs under nonintegration (where, implicitly, B and S's assets are separately owned). As can be seen in Table 1, any such change will indeed increase the likelihood of integration when all investments are self-investments, but will decrease the likelihood of integration with cross-investments.[13]

Given the focus of the TCE on specificity, it is natural to try to think about the effects of changes in marginal returns by instead considering changes in the *specificity* of marginal returns. Consider first the case in which i_B is a self-investment. Then we can think of the level of *marginal people specificity* for B of his investment i_B as being given by the difference $(\alpha_B - \beta_{B1})$, which is the amount by which the marginal return of i_B for B is reduced when B has control of the upstream assets but no longer deals with S. In contrast, the level of *marginal asset specificity* of i_B for B is given by $(\beta_{B1} - \beta_{B0})$,

12. One can show that when $\alpha_B > \beta_{B1} > \beta_{B0}$ (as in Hart, 1995) an equal *proportional* increase in these parameters raises the likelihood of integration. Thus, as in Hart (1995), the model above does deliver the result that in this case the likelihood of B owning the upstream asset must fall if we reduce all of B's marginal returns to zero. (A similar point applies to S.)

13. The same would be true if we were instead considering a situation in which both B and S integration were possible. Indeed, any parameter changes that affect marginal returns under only a single ownership structure will have the same directional effect on the probability of integration in either case. In contrast, the effect of changes that affect marginal returns under more than one structure (such as the changes in marginal people specificity in Table 2) would not necessarily be the same if both B and S integration were possible.

which is the degree to which B's returns are further reduced if he also does not have access to the upstream asset.

A change in the level of marginal people specificity of i_B for B holding the level of asset specificity fixed can therefore be captured by equal-sized reductions in β_{B1} and β_{B0} and has a differential effect on Δ equal to $\frac{1}{2}[i_B^*(1;\cdot) - i_B^*(0;\cdot)]$.[14] This is positive as long as B invests more when he owns the upstream asset. Note that increases in marginal people specificity therefore matter only if there is some marginal asset specificity present—that is, if $\beta_{B1} > \beta_{B0}$. To understand why the degree of marginal people specificity matters for the integration decision, consider first the case in which B underinvests regardless of the ownership structure (as in Hart, 1995). Then the greater the degree of marginal people specificity of B's investments, the more distortion there is in B's investment decision, and the greater is the value of the increase in its level that would be caused by integration; hence the greater the likelihood that we would observe integration. When we have, instead, overinvestment by B, greater marginal people specificity reduces the distortion in B's investment level, and so reduces the cost of the increase in B's investment that would accompany integration.

A change in the level of marginal asset specificity of i_B for B holding the level of marginal people specificity fixed is captured by a decrease in β_{B0}. This leads to a differential change in Δ equal to $\frac{1}{2}[\alpha_B - i_B^*(0;\cdot)]$, which is positive when we are in a case of underinvestment.

Finally, observe that in this (linear quadratic) model, in the case of self-investments Δ is fully determined by the levels of marginal people and asset specificity. This follows from the fact discussed earlier that in this case equal changes in α_B, β_{B1}, and β_{B0} have no effect on Δ, so we can normalize one of these three parameters (and, similarly with α_S, σ_{S1}, and σ_{S0}).

Consider, instead, the case of cross-investments. We might in a parallel fashion define the level of marginal people specificity of i_S for B by the difference $(\alpha_S - \beta_{S1})$, and the level of marginal asset specificity of i_S for B by $(\beta_{S1} - \beta_{S0})$. Unfortunately, these specificity measures are not sufficient to determine Δ in the cross-investment case. Consider, for example, an increase in the level of marginal people specificity for B of i_S. One way to create such a change is through equal-sized reductions in β_{S1} and β_{S0}. This change leads to a differential change in Δ of $\frac{1}{2}[i_S^*(0;\cdot) - i_S^*(1;\cdot)]$, which is positive as long as S invests more when he owns the upstream asset (i.e., when there is marginal asset specificity present). By way of contrast, we could instead increase the level of marginal people specificity for B of i_S by increasing α_S. This leads to exactly the *opposite* change in Δ of $\frac{1}{2}[i_S^*(1;\cdot) - i_S^*(0;\cdot)]$.

Parallel definitions can be formulated for the marginal specificity of investments for S. Increases in specificity for S all have the opposite effects from those derived for increases in specificity for B. The results for

14. Alternatively, we can capture this change in marginal people specificity with an increase in α_B.

Table 2. Effects of Increases in Marginal Specificity on the Probability of Integration with Self-Investments (Underinvestment and Asset Specificity Case)

	Type of marginal specificity:	
Investment by	People	Asset
B	+ ($d\beta_{B1} = d\beta_{B0} < 0$)	+ ($d\beta_{B0} < 0$)
S	− ($d\sigma_{S0} = d\sigma_{S1} < 0$)	− ($d\sigma_{S1} < 0$)

self-investments in these cases and those derived above are summarized in Table 2.[15]

2.2 Comparison of the PRT with the TCE

The foregoing analysis reveals a number of significant distinctions between the PRT and the TCE. First, and most immediately, while the TCE is concerned primarily with the level of quasi-rents (and the empirical literature on the TCE is almost uniformly focused on this), the PRT's focus is on the marginal returns to noncontractible investments. These need not be closely related in the data. For example, the level of quasi-rents can vary across situations with absolutely no variation in marginal returns, or even with marginal returns moving in the opposite direction.

Even if marginal returns and the levels of quasi-rents move in the same direction, in the PRT the effects of such changes depend delicately on a number of factors. First, changes in marginal returns under different ownership structures have different effects on the probability of observing an integrated structure: in the PRT, an increase in the marginal return to a self-investment by B under nonintegration has the opposite effect on the probability of integration from the same increase under integration. In contrast, such differences do not arise in the TCE because quasi-rents are assumed to matter only under nonintegration. In addition, in the PRT, the effects of changes in marginal returns reverse completely when we focus on cross-investments

15. A few words should be said about the linear-quadratic specification I assume here. First, all of the results generalize to a case in which the returns from trade are instead $\pi(i_B, i_S) + \alpha_{B2}(i_B)^2 + \alpha_{S2}(i_S)^2$, the disagreement payoffs are instead $w_B(i_B, i_S|A_B) + \beta_{B2}(i_B)^2 + \beta_{S2}(i_S)^2$ and $w_S(i_B, i_S|A_B) + \sigma_{B2}(i_B)^2 + \sigma_{S2}(i_S)^2$, and the cost functions instead have the form $c_j(i_j) = \gamma_j i_j + \frac{1}{2}\gamma_{j2}(i_j)^2$ for $j = B, S$. An important feature of this specification (like the one in the text), however, is that the change in investment resulting from an equal change in the marginal disagreement payoff returns under integration and nonintegration are equal. This plays a role in some of the results that involve simultaneous parameter changes under more than one ownership structure. One such case is the results for marginal people specificity in Table 2. In contrast, it can be shown that it remains true that (as discussed in the text) an increase in the importance of an agent's physical capital investment in the upstream asset increases the likelihood he will own it. A second important feature of the specification is that there is no interaction between the investments of B and S. The results on parameter changes under even a single-ownership structure depend on this. They continue to hold when investments are complements (as in Hart and Moore, 1990), but may not if they are substitutes.

instead of self-investments, and when we focus on cases of overinvestment rather than underinvestment.

These observations suggest that there are in fact a rich and demanding set of theoretical predictions that in principle can be used to test the PRT. However, they also suggest that a great deal of information about the trading environment is necessary.

3. Inferences from Three Prominent Articles

In this section I examine what we can learn about the empirical relevance of the PRT from three of the more prominent empirical studies of the transaction cost determinants of vertical integration. In turn, I discuss Monteverde and Teece's (1982b) study of automobile assemblers' procurement of automobile components, Masten's (1984) examination of procurement in the aerospace industry, and Joskow's (1985) discussion of vertical integration between electric utilities and coal mines.

3.1 Monteverde and Teece (1982b)

Monteverde and Teece (1982b) (henceforth M&T) provided the first econometric study of the transaction cost determinants of vertical integration.[16] In their path breaking article, M&T sought to explain the level of internal versus external procurement for 133 components used by GM and Ford in 1976. M&T focused not on physical asset specificity, but rather on the transaction-specific know-how that would be generated for the supplier during the design development process:

> We hypothesize that assemblers will vertically integrate when the production process, broadly defined, generates specialized, nonpatentable know-how.... The existence of transaction-specific know-how and skills and the difficulties of skill transfer mean that it will be costly to switch to an alternative supplier (Teece 1977, 1980). An assembler will tend to choose vertically integrated component production when high switching costs would otherwise lock the assembler into dependence upon a supplier and thereby expose the assembler to opportunistic recontracting or to the loss of transaction-specific know-how.

M&T used an ordinal measure of the level of engineering effort that went into designing the component to proxy for the likely extent of specialized know-how acquisition. In addition, they included control variables for the specificity of the component to the manufacturer in question (specifically, whether the component can be used directly in a variety of cars), an assembler dummy variable, and subsystem dummy variables. They found statistically significant evidence that increased levels of engineering effort and

16. For a related article focusing on ownership of specialized equipment, see Monteverde and Teece (1982a).

increased component specificity both raised the likelihood of integration and that GM was more likely to procure internally than Ford, all things being equal.[17]

From the standpoint of the TCE, the results look like a resounding success: the two variables designed to capture the extent of quasi-rents both indicate that greater levels of quasi-rents increase the likelihood of integration. But what do they tell us about the PRT? This is not so clear. To begin with, M&T tell us nothing about the extent to which important investments are not contractible: in the absence of noncontractible investments, of course, the PRT would predict no effects should be observed (in contradiction to the evidence). More generally, if there is no asset specificity for these noncontractible investments, then again the PRT tells us that no effects should be observed. Thus if we are not to reject the PRT, we will need to consider models in which noncontractible investments are present whose marginal returns exhibit asset specificity.

To develop such models, we need to think first about how the payoffs for components with a higher level of "know-how" and those with a higher level of component specificity differ from those with low levels of each of these variables. In what follows, we shall think of components with a higher level of know-how compared to those with a lower level as having (i) the same payoff for B in the event he does not reach an agreement with S and does not own the upstream asset, in which case he must use a component not embodying any of this know-how; (ii) a higher level of joint payoff if B and S do trade, since they can then benefit from this know-how; (iii) possibly a higher payoff to B when he does not reach an agreement with S but owns the upstream asset, since some of the know-how may be embodied in the asset; and (iv) possibly a higher level of payoff for S when he fails to reach an agreement with B, since he may be able to use his know-how in sales to other buyers. Thus we may see increases in some or all of α_B and α_S [point (ii)], β_1, β_{B1}, β_{S1} [point (iii)], σ_0, σ_1, σ_{B0}, σ_{B1}, σ_{S0}, and σ_{S1} [point (iv)].

In contrast, we shall think of an increase in the component specificity variable primarily as reducing the payoffs to S in the event he does not reach an agreement with B, since his component cannot be sold readily to other buyers without modification (possibly leading to reductions in σ_0, σ_1, σ_{B1}, σ_{S1}, σ_{B0}, and σ_{S0}). It is also possible, however, that it may be associated with a decrease in B's payoffs in the event that he does not reach an agreement with S, since B's investment may be tailored to the specific component and a more specific component may have fewer alternative sources of supply (possibly leading to reductions in β_0, β_1, β_{B1}, β_{S1}, β_{B0}, and β_{S0}).

With these effects identified, we can imagine a number of plausible property rights models of this situation involving noncontractible investments and asset specificity. As we shall now see, some of these models produce results

17. It is somewhat unfortunate, however, that M&T provided no estimates of the economic magnitude of the engineering and component specificity effects (i.e., the degree to which changes in these variables altered the probability of internal procurement).

that are consistent with the M&T results, while some do not. Consider the following three possibilities.

Model MT1. *Know-how acquisition is exogeneous and B makes noncontractible self-investments that are complementary to* S*'s acquisition of know-how.*

As an example, we can imagine that B invests in marketing and distributing its cars, but the success of these investments depends on the quality of the components that S produces, which is in turn affected by S's level of know-how. To represent this case, assume that only B has a noncontractible investment, that $\sigma_{B0} = \sigma_{B1} = 0$, and that $\alpha_B > \beta_{B1} > \beta_{B0}$. This last assumption captures the idea that B's marginal returns to investment are increasing in the level of know-how that is incorporated into the component, and that this is highest when B has access to both S and the upstream asset, and is lowest when B has access to neither S nor the upstream asset. Note that in this situation, B underinvests under either ownership structure, and invests more when he owns the upstream asset.

Consider first the effects of increasing levels of know-how in such a setting. This would reasonably be expected to increase α_B. Moreover, it might increase β_1 and β_{B1} if S's know-how is partially embodied into the upstream asset (and so benefits B when B owns this asset and cannot reach an agreement with S). Finally, it might increase σ_0 and σ_1 if S can use its know-how in dealings with other potential buyers in the event that B and S do not trade. In the property rights model, these changes will increase the probability of observing integration, and so are consistent with M&T's findings.

Now consider the effects of an increase in the level of component specificity. An increase in component specificity is likely to decrease σ_0 and σ_1 by reducing the degree to which S can use its know-how to serve other buyers in the event of not reaching an agreement with B. This has no effect on the probability of integration. In addition, an increase in component specificity may reduce β_0, β_1, β_{B0}, and β_{B1} if B's investment returns are tied to the specific component and more specific components are less likely to be available from an alternative source. While changes in β_0 and β_1 have no effect on the probability of integration, the changes in β_{B0} and β_{B1} have offsetting effects. If they decrease equally—so that the change amounts to an increase in the marginal people specificity of i_B without any affect on its marginal asset specificity—then the probability of integration will increase in the property rights model (see Table 2). Since this reduction in outside sources is more likely to matter when B does not own the upstream asset, we might expect β_{B0} to fall by more than β_{B1}; that is, asset specificity would increase. This would also lead to an increase in the probability of observing integration.

Thus the predictions of this first property rights model are fully in accord with M&T's findings.

Model MT2. *B's investments create know-how for* S.

As an example, B may be devoting effort to helping S understand how its component can be more effective in an automobile, where some of this

knowledge is specific to B's automobiles and some applies to all manufacturers' automobiles. To represent this case, assume that only B has a noncontractible investment, that $\alpha_B > \sigma_{B0} > \sigma_{B1} > 0$ (the benefits to S's know-how are largest when B and S trade, and smallest when S serves other buyers without access to the upstream asset), and that $\alpha_B > \beta_{B1} \geq \beta_{B0} = 0$ with $\beta_{B1} > 0$ if B self-produces when he owns the upstream asset and fails to reach an agreement with S and if some of S's know-how is embodied in the upstream asset. Under these conditions, B underinvests and B invests more when he owns the upstream asset.

In this case we expect an increase in the importance of know-how acquisition to be associated with increases in α_B, in σ_{B0} and σ_{B1} if know-how can be used by S in his dealings with other buyers when he does not reach an agreement with B, and in β_{B1} if know-how is partially embodied in the asset. The increases in α_B and β_{B1} both raise the likelihood of an integrated outcome under the PRT. The changes in σ_{B0} and σ_{B1} have offsetting effects. If increases in the importance of know-how acquisition increase σ_{B0} at least as much as σ_{B1}, as might be expected if S is more able to use his know-how when he has access to the upstream asset, then the net effect of the increases in these two parameters is also to increase the probability of integration (equal increases in σ_{B0} and σ_{B1} increase the probability of integration; from Table 1, a further increase in σ_{B0} further increases this probability).

In this environment, more component specificity will limit the returns that S can achieve in the event he does not reach an agreement with B. We expect these changes to be associated with decreases in σ_0, σ_1, σ_{B0}, and σ_{B1}. It may also reduce β_0 and β_1 if there are fewer alternative sources for a more specific component (we do not expect it to decrease β_{B1}, since if this is positive it means that B finds self-production to be optimal when he owns the upstream asset and does not trade with S). Of these parameter changes, only the changes in σ_{B0} and σ_{B1} matter for integration decisions. An equal change amounts to an increase in the marginal people specificity of i_B for S, which *reduces* the likelihood of integration. Moroever, more component specificity may also be expected to decrease σ_{B0} at least as much as it decreases σ_{B1} if the know-how is tied to the specific component and if having access to the upstream asset increases S's ability to produce the specific component. If so, this would also serve to reduce the likelihood of integration.

Thus in this second property rights model, the predicted affects of more component specificity are at odds with M&T's findings.

Model MT3. *S's investments create know-how for* S.

For example, S's acquisition of know-how may be determined instead by the effort that S puts into understanding how to produce the component. To represent this case, assume that only S has a noncontractible investment and that $\alpha_S > \sigma_{S0} > \sigma_{S1}$. These inequalities reflect the fact that S's know-how has the largest return when he has access to the upstream asset and trades with B, and has the smallest return when he has access to neither the upstream asset nor B as a trading partner. In addition, $\beta_{S1} \geq \beta_{S0} = 0$, with $\beta_{S1} > 0$

if B self-produces when he owns the upstream asset and fails to reach an agreement with S and if some of S's know-how is embodied in the upstream asset. These assumptions imply that S underinvests under either ownership structure and that he invests more when he owns the upstream asset.

In this case we expect an increase in the importance of know-how acquisition to be associated with increases in α_S, σ_{S0}, σ_{S1}, and possibly in β_{S1} if know-how gets embodied in the upstream asset. The increase β_{S1} causes a decrease in the probability of integration. As we noted earlier, an equal increase in α_S, σ_{S0}, and σ_{S1} has no effect on the probability of integration. However, if the increase in α_S is at least as large as the increase in σ_{S0}, which is in turn at least as large as the increase in σ_{S1} (e.g., with equal proportional changes), then the net effect of these three changes is to *decrease* the probability of integration (the change can be thought of as starting from an equal change in the three parameters and then further increasing both α_S and σ_{S0}).

Increases in component specificity will lead to decreases in σ_0, σ_1, σ_{S0}, σ_{S1}, and possibly in β_0 and β_1 if alternative sources for B are more limited when component specificity is higher (again, we don't expect it to lead to a decrease in β_{S1} when this is positive, since if it is positive this means that B's best alternative when he does not reach an agreement with S and he owns the upstream asset is to self-produce). If σ_{S0} and σ_{S1} decrease equally this lowers the likelihood of integration. However, we may expect the decrease in σ_{S0} to be at least as large as the decrease in σ_{S1}, since S's outside alternatives are likely to be hurt more by the increase in specificity when he owns the upstream asset (since he then can produce the specific component). The net effect of this change therefore appears to be ambiguous. Thus the predictions of this third property rights model contradict M&T's findings regarding the effects of increased know-how, and possibly also their findings concerning increased component specificity.

The conclusion to be drawn from this discussion seems to be that there is simply not enough information provided in M&T's study to evaluate the predictive power of the PRT.

3.2 Masten (1984)

Masten (1984) provided the second econometric study of procurement integration decisions. Masten's focus was on the procurement decisions of a large aerospace company over 1,887 components. Masten included as explanatory variables the degree of component complexity (as a measure of the difficulty of complete contracting), the degree to which the component was specialized to this aerospace firm (similar to M&T's component specificity variable), and the degree to which colocation of facilities or processes was thought to be important (as another measure of quasi-rents). He found that the first two variables had positive and statistically significant effects on the likelihood of integration, while the third variable had a positive but statistically insignificant effect on the likelihood of integration. He also found that the positive effect of complexity only seemed to occur conditional on there being a high

level of component specificity, and that the effect of component specificity was much higher when the part was complex.[18]

Once again, the results appear to be a success for the TCE, although perhaps a less complete success than the M&T results, given the insignificance of the site specificity variable (although the documented importance of the complexity variable adds a distinct supporting piece of evidence). As we have already seen, however, we could readily describe reasonable property rights models of this setting to get the effect of increased component specificity on the probability of integration to go in either direction. Moreover, as I shall discuss in detail when considering Joskow's (1985) study in the next subsection, a similar conclusion can be drawn about the effects of colocation. Without more detail about the contracting environment, little can be said about the ability of the property rights model to explain the integration patterns found in Masten's data.

3.3 Joskow (1985)

Joskow (1985) studies the coal procurement decisions of electric utilities. In contrast to the previous two articles, Joskow (1985) is not an econometric exercise, but it does offer its readers a more detailed view of the investment and procurement process than do those articles. Its primary finding regarding vertical integration is that vertical integration appears to be much more likely for "mine-mouth" electric generating plants than for others. This finding is interpreted as being in accord with the TCE, since such colocation raises the level of quasi-rents.

What does the PRT have to say about the effect of colocation (which, it seems clear, is itself a contractible decision) on the probability of integration? To fix ideas, consider Figure 1. There are five potential coal mines, labeled S and S_1 to S_4. The buyer, an electric utility plant, serves customers in the city, and the closest coal mine to the city is S. All roads go through the city. Now consider the effect of B moving the location of its plant from F (Far) to N (Near), which is nearer to S and further in road distance from all of the other coal mines.[19] Note that this change has no effect on S's payoff when S cannot reach an agreement with B (S's distance to other electric utilities has not been affected).[20] Thus, in contrast to the component specificity variables in Monteverde and Teece (1982b) and in Masten (1984), the colocation variable here affects only B's disagreement payoffs.

A second difference from the component specificity variable used in those studies is that we can reasonably put much more structure on the effects

18. One concern, however, is that this finding could be driven by the strong functional form assumptions embodied in the probit model Masten estimated.

19. It is worth noting that the effects derived below would reverse if moving closer to S also moved B closer to the other coal mines.

20. Note that in actual data it could be that colocation decisions do (on average) affect the distance of the coal mine to other electric utility plants, in contrast to our maintained assumptions here.

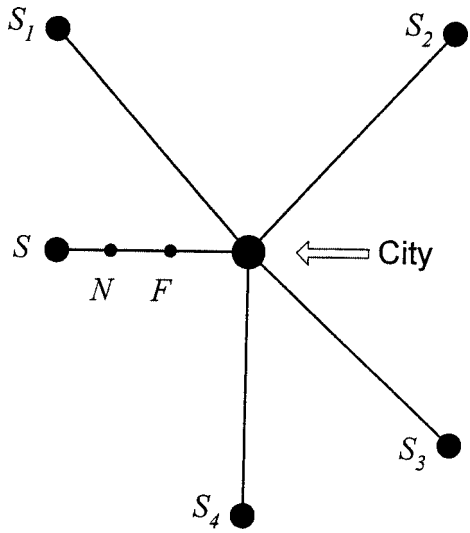

Figure 1. Two possible electric utility plant locations (*N* and *F*).

of colocation. In particular, for given investment choices by the utility and the mine, the effects of colocation on the payoffs when *B* and *S* trade, and on *B*'s payoffs from each alternative trading possibility in the event he does not reach an agreement with *S*, arise solely through the changes it induces in transportation costs (of coal and electricity). As a result (assuming transportation costs are unaffected by investments), any effects of colocation on the marginal returns to investment arise solely through the changes colocation induces in the identity of *B*'s next-best alternative in the event that he does not reach an agreement with *S*. For example, moving closer to *S* may mean that *B*'s next-best alternative when it owns *S*'s asset is to bring in another manager rather than to procure from another coal mine, which it might find is its next-best alternative if it is located at *N*.

With this in mind, we can see that the effects of colocation on the probability of observing vertical integration in a property rights model are ambiguous. To start, note that if *B*'s next-best alternative under each ownership structure is independent of its location (e.g., if—regardless of his location—*B* would procure from another mine if he does not own the upstream asset but would bring in a manager to run the mine if he did own it), then colocation has no effect on the probability of integration: in this case, colocation changes transport costs, but not the marginal returns to noncontractible investments. On the other hand, to see that colocation can have effects in either direction, consider the following three models:

Model J1. Suppose that *B* invests (noncontractibly) in improving his efficiency from using *S*'s coal and that this has no effect on his efficiency from

using coal from mines $S_1, \ldots, S_4$. Suppose also that when B is at F his next-best alternative in the event he does not reach an agreement with S is to procure from another mine, regardless of whether he owns the mine. However, suppose that when he is located at N his next-best alternative is procuring from another mine if he does not own the mine, but is bringing in a manager to run the mine if he does own it (that is, colocation lowers the payoff to outside procurement, and thus leads B to optimally run the mine when located at N). Hence $\beta_{B1} = \beta_{B0} = 0$ if B is at F, but $\beta_{B1} > \beta_{B0} = 0$ if B is at N. In this case, Table 1 tells us that colocation increases the probability of integration since it raises the marginal return to investment for B's disagreement payoff under integration.

Model J2. Suppose that S invests (noncontractibly) in quality improvements that are embodied in the mine. B's next-best alternatives are the same as in Model J1. Formally we would have $\beta_{S0} = \beta_{S1} = 0$ when B is located at F, but we would have $\beta_{S1} > \beta_{S0} = 0$ when B is located at N. In this case, S always underinvests, and reduces his investment when B is located at N and owns the mine. Table 1 tells us that in this case colocation reduces the probability of integration: intuitively, by keeping the mine out of B's hands when B is located next to the mine, nonintegration keeps S from worrying that by investing in coal quality he is improving B's disagreement position.

Model J3. Now suppose that B invests (noncontractibly) solely in developing a capability for burning coal from mines $S_1, \ldots, S_4$. Maintain the same assumption about B's next-best alternatives as in Model J1. In this case, the social value of his investment is zero: $\alpha_B = 0$. Then we have $\beta_{B0} = \beta_{B1} > 0$ when B is located at F, while when B is located at N, $\beta_{B0} > \beta_{B1} = 0$. Here, B overinvests if he is located at F or under nonintegration if he is located at N, while he invests efficiently (i.e., does not invest) when he is located at N and owns the mine. Here, colocation is associated with an increase in the likelihood of integration.

Once again, despite the relatively greater informativeness of Joskow's article about the contracting and production environment he studies, there is insufficient information to evaluate the PRT.

4. Conclusion

The discussion above indicates that the PRT's predictions differ in important ways from those of the TCE, and that the existing empirical evidence that is supportive of the TCE sheds little light on the empirical relevance of the PRT.

At the same time, the discussion above suggests that the PRT offers a rich set of predictions that should lend themselves to testing. One can imagine two different approaches to this task. One is a "natural experiment" approach in which we identify some variation whose effect on marginal returns to investments in various ownership structures seems very clear a priori. This

approach is much like that followed in essentially all of the empirical work testing the TCE, in which some measure of specificity (e.g., product specificity, investment specificity, or colocation) is taken as a proxy for the level of quasi-rents in the relationship. Like the literature testing the TCE, such an approach faces the challenge of finding convincing natural experiments; that is, ones that clearly affect marginal returns in particular ways and are exogenous sources of variation. Admittedly, finding such natural experiments may be a greater challenge for testing the PRT than for testing the TCE. A small number of recent articles (Hanson, 1996; Baker and Hubbard, 1999; Woodruff, forthcoming) have attempted to do just this. Baker and Hubbard (1999), for example, study the ownership of tractor-trailer trucks in the trucking industry. Their "natural experiment" is the introduction of on-board computers, which (they argue) essentially made driver care contractible, a change that can be viewed as equivalent to a change in marginal returns.[21,22]

It is worth noting that good natural experiments for testing the PRT may also provide more stringent tests of the TCE. For example, increases in disagreement-payoff marginal returns to cross-investments under nonintegration raise the likelihood of integration in the PRT (see Table 1), but—since such changes reduce quasi-rents—they would be predicted to reduce integration in the TCE. In fact, there are reasons to suspect that the relation between quasi-rents and integration may not be as clear as suggested by the current empirical literature on the TCE. Indeed, it appears that trading partners sometimes deliberately increase the extent of lock-in in their relationship by, for example, writing exclusive contracts or agreeing to joint ventures requiring unanimity. This suggests that firms may not always find such lock-in undesirable, and so may not always integrate to avoid it, in contrast to the predictions of the TCE (but explainable with the PRT).[23]

A second possible approach is more structural in nature. In particular, one might consider using data on outcomes to measure the marginal returns

21. Specifically, imagine that S is the driver, and B the trucking firm, and i_S is driver care of the truck. Since the owner of the truck receives its full future value, before the introduction of monitoring we have $\beta_{S0} = \sigma_{S1} = 0$ and $\beta_{S1} = \sigma_{S0} = \alpha_S$. Monitoring permits the parties to write a contract that induces efficiency when the firm owns the truck by paying the driver $\alpha_S i_S$ when the (now observed) care level is i_S. This contractual payment effectively changes σ_{S1} to α_S and β_{S1} to 0, a change which, according to Table 1, raises the likelihood of ownership of the truck by the trucking firm. (Baker and Hubbard, 1999, actually consider a multitask property rights model in which drivers may exert effort both to take care of the truck and to rent seek by looking for alternative hauls.)

22. Hanson (1996) and Woodruff (forthcoming) argue that the nature of buyer and seller investments varies across different segments of the industries they study (textiles for Hanson, shoes for Woodruff). Specifically, Hanson argues that the importance of buyer versus seller investments differs across segments (recall, however, that our discussion above noted that such a prediction is problematic for human capital investments); Woodruff makes a similar argument regarding the importance of investments, but also argues that there is greater marginal asset specificity in some industry segments.

23. Williamson (1983) makes a similar point about the potential value of lock-in in his discussion of hostages to support exchange, but does not connect this point to the question of whether greater lock-in might therefore sometimes lead parties to remain unintegrated.

to investments, and then testing the PRT given these observed returns. But this approach encounters two serious difficulties. First, how can we measure investments that are assumed to be noncontractible in the theory? Second, how can we measure the effects of investments on disagreement payoffs when in the theory these payoffs are never observed?[24] Overcoming these conceptual challenges seems essential for any structural effort to test the PRT.

Although existing empirical work on the TCE may not shed much light on the applicability of the PRT, the time seems right to do so, and also to subject the TCE to more stringent tests. This would go far toward advancing our understanding of the transaction cost determinants of firms' boundaries.

References

Aghion, Phillippe, and Jean Tirole. 1997. "Formal and Real Authority in Organizations," 105 *Journal of Political Economy* 1–29.

Bajari, Patrick, and Steven Tadelis. 2001. "Incentives versus Transaction Costs: A Theory of Procurement Contracts," 32 *RAND Journal of Economics* 387–407.

Baker, George, and Thomas Hubbard. 1999. "Contractibility and Asset Ownership: On-Board Computers and Governance in U.S. Trucking," mimeo, University of Chicago.

Bolton, Patrick, and Michael D. Whinston. 1993. "Incomplete Contracts, Vertical Integration, and Supply Assurance," 60 *Review of Economic Studies* 121–48.

Che, Yeon-Koo, and Donald B. Hausch. 1999. "Cooperative Investments and the Value of Contracting," 89 *American Economic Review* 125–47.

Coase, Ronald H. 1937. "The Nature of the Firm," 4 *Economica* 386–405.

Grossman, Sanford J., and Oliver D. Hart. 1986. "The Costs and Benefits of Ownership: A Theory of Vertical Integration," 94 *Journal of Political Economy* 691–719.

Hanson, Gordon H. 1996. "The Decline of Vertical Integration in English Textiles: An Empirical Application of the Theory of the Firm," mimeo, Univeristy of Texas, Austin.

Hart, Oliver. 1995. *Firms, Contracts, and Financial Structure*. Oxford: Oxford University Press.

———, and John Moore. 1990. "Property Rights and the Nature of the Firm," 98 *Journal of Political Economy* 1119–58.

Joskow, Paul L. 1985. "Vertical Integration and Long-Term Contracts: The Case of Coal-Burning Electric Generation Plants," 1 *Journal of Law, Economics, & Organization* 33–80.

———. 1988. "Asset Specificity and the Structure of Vertical Relationships: Empirical Evidence," 4 *Journal of Law, Economics, & Organization* 95–117.

Klein, Benjamin, Robert G. Crawford, and Armen A. Alchian. 1978. "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process," 21 *Journal of Law and Economics* 297–326.

Masten, Scott E. 1984. "The Organization of Production: Evidence from the Aerospace Industry," 27 *Journal of Law and Economics* 403–17.

———, ed. 1996. *Case Studies in Contracting and Organization*. New York: Oxford University Press.

Matouschek, Niko. 2001. "Information and the Optimal Ownership Structure of Firms," mimeo, Northwestern University.

Monteverde, Kirk, and David J. Teece. 1982a. "Appropriable Rents and Quasi-Vertical Integration," 25 *Journal of Law and Economics* 321–28.

———. 1982b. "Supplier Switching Costs and Vertical Integration in the Automobile Industry," 13 *Bell Journal of Economics* 206–13.

24. The measurement of disagreement payoffs would also be an issue for any structural attempt to test the TCE.

Rajan, Raghuram G., and Luigi Zingales. 1998. "Power in a Theory of the Firm," 113 *Quarterly Journal of Economics* 387–432.

Shelanski, H. A., and P. G. Klein. 1995. "Empirical Research in Transaction Cost Economics: A Review and Assessment," 11 *Journal of Law, Economics, & Organization* 335–61.

Williamson, Oliver E. 1975. *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: Free Press.

———. 1979. "Transaction Costs Economics: The Governance of Contractual Relations," 22 *Journal of Law and Economics* 233–62.

———. 1983. "Credible Commitments: Using Hostages to Support Exchange," 73 *American Economic Review* 519–40.

———. 1985. *The Economic Institutions of Capitalism*. New York: Free Press.

———. 2000. "The New Institutional Economics: Taking Stock, Looking Ahead," 38 *Journal of Economic Literature* 595–613.

Woodruff, Christopher. "Non-Contractible Investments and Vertical Integration in the Mexican Footwear Industry," forthcoming in *International Journal of Industrial Organization.*