CONTRACTUAL COMMITMENTS, BARGAINING POWER, AND GOVERNANCE INSEPARABILITY: INCORPORATING HISTORY INTO TRANSACTION COST THEORY

NICHOLAS S. ARGYRES
JULIA PORTER LIEBESKIND
University of Southern California

We extend transaction cost economics by arguing that prior contractual commitments made by a firm can limit its ability to differentiate or change its governance arrangements in the future—a condition we term governance inseparability. Changes in bargaining power between a firm and its exchange partners also can result in governance inseparability. Consequently, governance choice may be more particularistic than the current version of transaction cost economics allows. We provide several testable propositions.

Transaction cost economics (TCE) is becoming a predominant economic theory of the firm. This theory holds that transactions with particular characteristics are governed efficiently by certain types of organizational arrangements and not by others. In particular, firms are formed to govern transactions that are characterized by uncertainty combined with economic spillovers resulting from asset indivisibilities, asset specificity, and asset extensibility (Alchian & Demsetz, 1972; Coase, 1937; Grossman & Hart, 1986; Klein, 1983; Klein, Crawford, & Alchian, 1978; Teece, 1980; Williamson, 1979). Thus, in the TCE theory of the firm, an individual transaction is the unit of analysis for predicting organizational form (Williamson, 1985).

In this article, however, we argue that focusing on the characteristics of isolated transactions can be insufficient to explain the scope of the firm. The reason for this is that the governance of any new transaction in which a firm may seek to engage may become linked inseparably with the governance of other transactions in which the firm is already engaged. In these cases, what TCE might predict to be efficient modes of governance for the new transaction may, in fact, be excessively costly, or even infeasible altogether, for that particular firm. Thus, firms are subject to a condition that we call governance inseparability—a condition in which a firm’s past governance choices significantly influence the range and types of governance mechanisms that it can adopt in future periods.

Governance inseparability can constrain a firm’s governance options in two ways. First, it may constrain a firm from switching from one governance mode to another for the same type of transaction. Second, governance inseparability may obligate a firm to use an existing governance arrangement for a new transaction, even if that particular transaction would be governed more efficiently by other means. This is a constraint on governance differentiation.

Governance inseparability has important implications for the theory of the firm. It implies that there will be variance in observed governance mechanisms among firms for a given transaction, since different firms will tend to incur different levels of costs for any particular governance arrangement. Also, governance inseparability can help to explain the limits to firm scope, because it may be too costly for a given firm to internalize a given transaction, owing to its arrangements in place, even if internalization were the optimal governance solution for that transaction considered in isolation. Similarly, governance inseparability may make it more difficult for a given firm to reduce its scope.

We seek to extend TCE theory by drawing attention to two factors that serve to produce
governance inseparability. The first of these factors is contractual commitments. Firms are most efficiently engaged in long-term exchange relationships (Williamson, 1979). However, these typically require long-term contractual commitments. Yet, contractual commitments serve to engender governance inseparability because (by our definition) they are costly, if not impossible, to reverse. As a result, a firm’s outstanding set of contractual commitments may significantly restrict its governance options in the future. The second factor that serves to produce governance inseparability is changes in the bargaining power of the other parties to a firm’s contractual commitments—such as employees, suppliers, or customers. Here, these parties can use unanticipated increases in their bargaining power to obligate a firm to adopt locally suboptimal governance mechanisms in the future.

Because firms, by definition, engage in long-term transactions, no firm can entirely avoid making contractual commitments. Hence, firms must always assume some risk of governance inseparability. As a result, governance inseparability implies that most firms will become constrained over time by their existing arrangements in place, which will limit both their scope and their strategic flexibility.

Because it focuses on the individual transaction as the basis of analysis, TCE does not fully consider the possibility that parties to transactions—in particular, firms—may be constrained in their choice of governance mechanisms by past governance choices. Incorporating governance inseparability into TCE does not require major changes in the theory’s basic behavioral assumptions (i.e., bounded rationality and opportunism), nor does it imply that firms make governance choices that are inefficient in a global sense. Rather, governance inseparability calls for tempering the use of the transaction as the unit of analysis (without abandoning it entirely) in order to capture the effects of historical constraints. As we show below, governance inseparability also places somewhat more emphasis on the bounded rationality assumption, at the expense of the “far-sighted contracting” assumption employed in TCE.

THE TCE THEORY OF THE FIRM

According to TCE, firms are a particular form of organization for administering exchanges, or “transactions,” between one party and another (Coase, 1937). In this conception of the firm, the firm itself is characterized as a “managerial hierarchy” and is contrasted with other forms of organization, most notably markets, in which transactions take place without managerial oversight (Williamson, 1975, 1985). The basic insight is that firms exist because they can sometimes reduce the costs of negotiating and enforcing terms and conditions of exchange relative to market contracting (Coase, 1937). This will be the case especially when uncertainty about future business conditions makes contracts incomplete and when transactions are characterized by economic spillovers, such as those that result from cospecific investments, asset indivisibility, and asset extensibility (Alchian & Demsetz, 1972; Coase, 1937; Grossman & Hart, 1986; Hart & Moore, 1990; Klein, 1983; Klein et al., 1978; Teece, 1980; Williamson, 1975, 1979, 1985). Hierarchies allow for better adaptation than contracting between autonomous parties. Adaptation is important because efficiency considerations often require that adjustments be made in the distribution of the gains from a trading relationship when trading conditions change. Such adjustments are likely to occur in a more cooperative and, hence, less costly way if the transaction occurs within a hierarchy, rather than through autonomous contracting. This is because hierarchies are able to resolve trading disputes by fiat as a last resort, whereas fiat is unavailable for governing market contracts (Williamson, 1975, 1985). The use of fiat within hierarchies is supported by contract law, since courts generally forbear from hearing cases involving disputants who are part of the same firm (Williamson, 1991). In contrast, in market contracting a party may use the law and the legal system opportunistically to “hold up” a contractual partner (Klein, 1993; Klein et al., 1978).
addition, placing the ownership of the assets in a given transaction into the hands of a single party (i.e., organizing the transaction within a firm) improves the incentives for making efficient transaction-specific investments when contracts are incomplete and the cost associated with a hold-up is significant (Grossman & Hart, 1986; Hart & Moore, 1990).

This TCE theory of the firm has been used to explain not only the existence of firms but the scope of the firm, in terms of both vertical integration and diversification. That is, the degree of vertical integration and/or diversification of a firm is determined by uncertainty and by asset specificity, asset extensibility, and asset indivisibility (Klein et al., 1978; Teece, 1980; Williamson, 1975, 1985). Moreover, the firm itself is understood to comprise not only assets that are fully internalized—in the form of wholly owned property, physical plant and equipment, and employees (Coase, 1937; Hart & Moore, 1990; Masten, 1988)—but also long-term contracts with external parties, such as buyers, suppliers, joint venture partners, and other such “hybrid” arrangements (Pisano, 1989; Williamson, 1991). Thus, the scope of the firm can be understood to be the scope of both its internal contracts—with its managers, employees, and shareholders—and its external contracts with other parties that sell to, buy from, or otherwise are parties to the firm’s non-market arrangements.

CONTRACTUAL COMMITMENTS AND GOVERNANCE INSEPARABILITY

We define a contractual commitment as an agreement between two or more parties that is binding on those parties, to the degree that to renego on the agreement will be costly. This implies that an exchange agreement (i.e., a contract) cannot be described as a contractual commitment unless enforcement mechanisms are in place that ensure both detection and punishment of reneging. With enforcement mechanisms, contractual commitments become, to a significant extent, irreversible. Without enforcement mechanisms, a contract can be changed without cost by either party. In this section we seek to demonstrate, using a number of examples, how contractual commitments can lead to a condition of governance inseparability.

The contractual commitments of a firm fall into two categories: (1) formal contracts and (2) informal or nonlegally enforceable contracts.

1. Formal contractual commitments. A firm can commit to a contract by preparing and signing a legally enforceable document that sets out the terms and conditions of future transactions between the firm and one or more other parties, be they employees, suppliers, buyers, or lenders. If a firm were to breach such a contract, it would be legally liable for damages or restitution payments. Moreover, seeking favorable renegotiating contracts can require side payment schemes that are difficult to implement (Demsetz, 1966). Examples of such legal contracts with stakeholders are union wage agreements, long-term supply contracts, exclusive dealership and franchise agreements, debt covenants, and customer warranties.

2. Informal or nonlegally enforceable contracts. A firm can also enter into informal or unwritten agreements with various other parties. To the degree that these agreements can be enforced, they also constitute contractual commitments.

One common type of informal contractual commitment is “self-bonding,” wherein a firm invests in firm-specific (i.e., sunk) assets whose value can only be recouped if it behaves in certain ways (Ghemawat, 1991; Sutton, 1991). For example, a firm may invest in assets that are cospecific to an exchange partner, such as committing irrecoverable capital to a joint venture, investing in an exclusive dealership, or financing a supplier’s firm-specific production facilities. If the firm ceases to trade with any of these parties according to agreed-upon terms and conditions, the other party can cease trading altogether, and the firm will lose the value of its investments. Williamson (1983) refers to arrangements such as these as “hostage-taking,” drawing an apt analogy between nonlegal enforcement mechanisms in medieval politics and present-day commerce.

Firms are also bound by informal “social contracts,” through which they are committed (at least to some degree) to respecting and upholding the norms of society. For instance, it may be legal for a firm to lay off many thousands of workers, but to the degree that society at large considers such actions unethical, the firm can expect to be sanctioned. Examples of such sanc-
tions include consumer boycotts, campaigns to investigate corporate actions, and refusal of social consent to other actions that firms may want to take.

In general, the contractual commitments that a firm enters into are particularistic; each contractual commitment is entered into with a particular and identifiable outside party. Thus, whereas shareholders may come and go at their own will, the parties to a firm’s contracts are tied to the firm, and the firm is tied to these other parties, because these commitments are not transferable. Moreover, contractual commitments, by their very nature, cannot be revoked at low cost. Governance inseparability is engendered because a firm’s contractual commitments tie it to specific other parties who have rights in relation to the firm. We consider two types of governance inseparability that may stem from contractual commitments: (1) constraints on governance switching and (2) constraints on governance differentiation.

Contractual Commitments and Constraints on Governance Switching

Governance inseparability in the form of a constraint on governance switching exists if a firm cannot efficiently enter into a governance arrangement of Type Y in future periods for a particular transaction because it already has a governance arrangement of Type X in place with another party for that transaction. Franchising agreements provide a very straightforward example of this type of constraint on governance modes. Outstanding contractual commitments in the form of franchise agreements and exclusive dealerships pose significant barriers to further agreements, and they also limit moves toward forward integration by the franchisor.

For example, in the past, Coca-Cola entered into a number of exclusive franchising agreements with independent bottling companies. These agreements prevent Coca-Cola from forward integrating, because they do not allow any company other than the franchisee to bottle Coca-Cola products within a given territory. As a result, when Coca-Cola decided to forward integrate in the 1980s, it could do so only by buying out its own franchisees in the open market for corporate control, where it had to pay a hefty price premium relative to the costs of simply forward integrating through internal expansion. Moreover, in some cases Coca-Cola has been unable to buy out its franchisees. In others, it has elected not to do so, since the costs of acquisition are too high relative to the benefits of internalization.

Contractual Commitments and Constraints on Governance Differentiation

Governance inseparability in the form of a constraint on governance differentiation exists when a firm is obligated to enter into a governance arrangement of Type X with one party because it already has a governance arrangement of Type X in place with another party. This type of inseparability commonly arises when firms want to differentiate their internal organizational arrangements.

For example, efforts to start and/or sustain new venture units have failed at many large firms (Burgelman & Sayles, 1986; Fast, 1978; Hlavacek, 1974; Sykes, 1986). According to TCE, this failure results from the inability of firms to make credible commitments to the “marketlike” governance arrangements often needed to support new ventures. Thus, Williamson (1985) argues that corporate management cannot sustain high-powered incentives for a venture internal to the firm, because venture managers know that corporate management can always abrogate those incentives in the future and may do so if it becomes in its economic interest. As a result, those ventures often fail or are never initiated. Hence, in this theory the commitment problem arises in the relationship between corporate management and venture managers.
Our arguments suggest that a different mechanism may be at work here. In particular, the relationship between corporate management and managers of established divisions may also be important in affecting the kinds of incentives that can be offered and sustained for internal ventures. Specifically, implicit contractual commitments to managers of established divisions may make specialized governance arrangements difficult to sustain, because these latter arrangements will be seen as violations of established commitments. For instance, high-powered incentives imply that a new venture division may be permitted to retain a higher proportion of its profits, insulating its management from profit fluctuations elsewhere in the firm and depriving other divisional managers of the opportunity to share in the venture’s success. These new incentive structures, therefore, undermine the incentive structures in place of established divisions, who will oppose them through any means available, thereby ensuring that their outstanding incentive contracts are honored. For example, in several case studies Fast (1978) documented instances of interdivisional conflict, where established divisions attempted to gain management control over new ventures and (implicitly) to suspend their special treatment.

With these arguments we do not deny that the problem of credible commitments to venture divisions is problematic for internal venturing. Rather, our arguments emphasize that a firm’s prior commitments to its other existing divisions play an important role in preventing the firm from sustaining the specialized governance arrangements needed to develop new ventures. Thus, we can conjecture that problems of governance inseparability may account, to a significant degree, for the failure of internal ventures.

Institutional theory also implies that certain types of contractual arrangements can become difficult to change over time and, hence, difficult to differentiate on a transaction-by-transaction basis. For instance, certain arrangements can become “taken-for-granted” (Zucker, 1987), mimicked in a search for various types of “legitimation” (DiMaggio & Powell, 1983), or compelled by state action (DiMaggio & Powell, 1991; Fligstein, 1990). Habits, pressures for legitimation, and mimicry do not play a role in our arguments. From our point of view, firms are profit-seeking entities, and contractual commitments are enforced by the threat of financial losses that can be imposed on defectors by their (actual or potential) private contractual partners. Thus, in our theory lack of differentiation arises because of economic realities and not because of social structures or norms per se.
State action does play a role in our argument because contract enforcement can occur through actual or potential appeal to the courts. However, in our view the state does not compel parties to adopt particular, narrowly defined kinds of contractual arrangements. Instead, it seeks to enforce the arrangements that parties have agreed upon privately, as long as those arrangements conform to certain widely held social norms, especially those concerning fairness (e.g., Michaelman, 1967). Therefore, for us, the normative environment in which contracting is undertaken remains in the background, rather than coming to the foreground in terms of directly shaping organizational arrangements, as is the case in institutional theory.

Is Governance Inseparability Associated with Contractual Commitments Avoidable?

Given that contractual commitments can engender governance inseparability, the question arises as to why a firm may not be able to foresee this eventuality and to eschew any commitments associated with this particular hazard. This would allow a firm to keep its governance options open in future periods, avoiding problems of inseparability and economizing on transaction costs.

There are at least two answers to this question. First, a firm cannot exist efficiently without commitments. Firms, as institutions, are costly both to establish and to liquidate (Williamson, 1979, 1985). Consequently, a firm that does not govern long-term transactions of one type or another is an inefficient form, yet long-term transactions typically require contractual commitments of one type or another.

The second answer is centered more in strategy than in governance cost considerations: contractual commitments are necessary for a firm to earn rents. Broadly speaking, firms derive rents from ownership of unique capabilities (Rumelt, 1984, 1987); even arbitrage rents derive from information that is closely owned and from changes in circumstances that make some contracts more valuable than their original purchase price (Barney, 1986; Demsetz, 1966). Yet, unique capabilities cannot be purchased in markets; instead, they usually derive from specialized and/or cospecific investments whose benefits accrue over time (Dierickx & Cool, 1989; Kirzner, 1979; Lieberman & Montgomery, 1988; Rumelt, 1987; Winter, 1987). Such investments cannot be supported rationally without contractual commitments. For instance, in a competitive labor market, a firm must commit to long-term employment or make other contractual commitments to its employees, if those employees are to make significant levels of firm-specific investments of human capital, such as learning firm-specific routines. Hence, a firm that eschews making commitments typically will be unable to earn rents and, therefore, will not survive.

If governance inseparability is indeed often unavoidable, the natural question raised by TCE is why a firm cannot take steps to create safeguards against inseparability when negotiating its early agreements. This possibility is suggested by the emphasis on far-sighted contracting that TCE features. As Williamson notes, TCE concedes that comprehensive contracting is not a feasible option (by reason of bounded rationality), yet it maintains that many economic agents have the capacities both to learn and to look ahead, perceive hazards, and factor these back into the contractual relation, thereafter to devise responsive institutions (1996: 9).

He also notes that "such a concept of contract presents healthy tensions . . . for organization theory" (1996: 9).

The view we adopt is that, in many cases, these tensions may resolve themselves more in favor of bounded rationality than in favor of foresight. For example, in industries that feature rapid technological innovation, anticipating future hazards and opportunities is often very difficult (Nelson & Winter, 1982). Many other industries also experience rapid changes in trade conditions and consumer behavior. Given the possibility of unforeseeable change, entering into any commitment at one point in time will necessarily involve the assumption of some risk of governance inseparability on the part of a firm and, hence, some risk of increased transaction costs. Consequently, recognizing and incorporating governance inseparability may require greater emphasis on the effects of bounded rationality on contracting, and on evolutionary change, relative to TCE's current emphasis on far-sighted contracting.
BARGAINING POWER AND GOVERNANCE INSEPARABILITY

A second set of reasons for observing governance inseparability among a firm’s transactions relates to bargaining power, which can be defined as the ability of one party to a contract to be able to influence the terms and conditions of that contract or subsequent contracts in its own favor. In this section we seek to demonstrate that governance inseparability is engendered when the other parties to a firm’s contractual commitments have bargaining power, and particularly so when their relative bargaining power increases unexpectedly. Again, our arguments suggest that the TCE theory of isolated governance choice needs to be modified.

Bargaining Power and Constraints on Governance Switching

Recall that one effect of contractual commitments is to prevent a firm with a governance arrangement of Type X in place for a particular transaction from switching to another governance arrangement of Type Y in future periods for that same transaction. Changes in bargaining power also can produce this effect. Consider two bargaining parties—Seller S and Buyer B—who enter into an efficiently structured long-term contract that serves to support investment in specific assets by both parties so that there is no anticipated ex post asymmetry in their bargaining power. However, because there is some uncertainty about prices and/or demand in the future, the long-term contract allows some scope for renegotiation. This arrangement is rational because the parties have equal bargaining power at the time the contract is signed. However, as time passes, conditions change in such a way that Buyer B increases its bargaining power over Seller S. At the time of contract renegotiation, B may use this increased bargaining power in such a way that, in the future, the governance options for S for some of its transactions are constrained. For instance, B may prevent S from forward integrating into its own business, forcing S to continue to sell across markets rather than to integrate vertically. Thus, the changes in bargaining power of B relative to S over time have imposed a degree of governance inseparability on S that S did not anticipate, and therefore could not prevent, in its original contract with B.

One common example of this type of governance inseparability is when unionized labor uses its bargaining power to restrict outsourcing. In recent years a number of U.S. companies have attempted to outsource production or service activities in order to sustain or increase their competitive advantage, but they have been opposed by their own workers. Examples include Deere, Ameritech, McDonnell Douglas, Boeing, United Parcel Service, Chrysler, and General Motors (GM).

For instance, a major strike led by the United Auto Workers (UAW) at GM’s brake manufacturing plant crippled production of autos and trucks for over a month in 1996, costing the firm $600 million. The UAW launched the strike after GM announced its intention to increase the number of outsourced parts in its automobiles in order to reduce its labor costs and circumvent restrictive union work rules. Although GM previously had agreed with the UAW not to increase outsourcing of certain components, this agreement was stated in fairly general terms and its enforceability through the courts was in some doubt. Thus, this was not a clearcut case of union enforcement of an existing contractual commitment but, rather, a case of GM’s workers using their bargaining power to change the terms and conditions of their basic contract, with the result that GM could not change the way its parts production processes were governed. Indeed, this incident raises the question of whether outsourcing at GM was implicitly constrained in earlier periods, during which union bargaining power was even stronger than today (the late 1960s and 1970s, for example). In such a case as this, we cannot expect transaction cost logic to apply strictly to a firm’s choice between internal and external contracts.

Franchising is another arena in which recent changes in bargaining power have served to constrain the types of contracts that can be offered by many franchisors. In the 1990s franchisees in a large number of chains have become more organized. According to Harris (1997), the American Franchisee Association grew from roughly 4,000 members to about 7,500 during the period 1993–1997, and the American Association of Franchisees and Dealers (AAFD) grew from 20 members in 1992 to about 6,000 in 1997. Chain-specific franchisee groups also have become
more widespread, doubling to 250 during 1994–1997. One of these groups, the Meineke Mufflers Dealers Association, recently won a major lawsuit against the franchisor—a victory that was apparently made possible only by the newly won financial clout of the AAFD (Harris, 1997). This legal decision could affect the ways in which rights to control communal advertising funds can be allocated in all of Meineke’s future contracts, and perhaps in future contracts offered by other franchisors as well. Thus, increased franchisee bargaining power is allowing them to impose new constraints on future contracts a franchisor might wish to offer.

Even when this increased power is partly anticipated and when contractual safeguards are included in franchise agreements, subsequent legal decisions favoring franchisees sometimes have invalidated the safeguards. In one recent case Naugles, Inc., a fast-food franchisor, specifically did not provide territorial exclusivity for its franchisees, allowing the firm to sell new franchises close to older ones as population densities increased (Harris, 1997). Naugles’ franchisees originally agreed to this contract provision. However, when Naugles attempted to open a new franchise close to an existing one, its franchisees collectively sued; they won their case in a California federal court. In this case Naugles was unable to switch from a de facto sole franchisee in a given area to multiple franchisees, despite its contractual provisions, because of increases in franchisees’ bargaining power. Note, however, that it was Naugles’ contractual commitments to its franchisees that exposed it to the deleterious effects of changes in their bargaining power.

Bargaining Power and Constraints on Governance Differentiation

In other instances a firm may be constrained in its ability to differentiate its organizational arrangements, owing to unanticipated changes in bargaining power—a case in point being the trucking industry. During the 1980s, the three major U.S. long-distance trucking firms (Yellow Corp., Consolidated, and Roadway) attempted to enter the short-haul trucking market. However, none of these companies has, as yet, been able to differentiate its organizational arrangements to suit this new market. Yellow Corp., for example, tried to establish a new subsidiary to conduct its short-haul business and to negotiate a new, more flexible union contract for this firm. According to an industry analyst, “The Teamsters basically said, ‘No way’” (Kansas City Business Journal, 1993: 1).

A parallel set of circumstances has arisen in the U.S. airline industry. Some large airlines maintain separate subsidiary firms to handle shorter routes. The pilots working for these subsidiaries earn much lower salaries, allowing (for example) American Airlines’ American Eagle subsidiary to compete with smaller, nonunionized airlines, such as Southwest. Recently, American sought to increase the proportion of flights handled by American Eagle by adding short-haul jets to its previously turboprop fleets. The pilots of American Airlines, arguing that they alone should be allowed to fly the short-haul jets, threatened a strike, rather than allow internal differentiation of jet pilot’s pay. Other airlines have avoided this problem by conducting their short-haul businesses through partially owned subsidiaries (e.g., Delta), through long-term contracts or strategic alliances with short-haul carriers (e.g., Continental), or through employee ownership (e.g., United), which may better align the incentives of the different parties.

In each of these examples, union bargaining power prevented a firm from carrying out a change in organization that it was seeking. Moreover, media reports suggest that in all of these instances, management was surprised at the union’s strength. It is likely, however, that changes in this strength were difficult to foresee. It is important to note that, in each of these cases, parties were able to obtain bargaining power because initial contractual commitments were in place. These commitments made it costly, if not impossible, for a firm to seek alternatives to dealing with the parties who had the bargaining power. Hence, contractual commitments not only expose a firm directly to a risk of increased governance inseparability, but expose it to the indirect risk of changes in bargaining power over time. Consequently, the two key

---

5 Minutes after the strike began, President Clinton intervened to avert a “national transportation crisis.” The parties went into binding arbitration, where the issue of subsidiary airliner substitution is being negotiated as of this writing. See the New York Times (1997) and the Wall Street Journal (1997).
variables that determine governance inseparability in our arguments—contractual commitments and bargaining power—can interact in their effects on governance choices. First, the presence of contractual commitments increases the effects of relative bargaining power on governance choices, especially for firms with low bargaining power. Second, and conversely, the greater a firm’s relative bargaining power, the greater its ability to negotiate in a way that limits contractual restrictions on its future activity.

Why Firms Are Unable to Anticipate and Safeguard Against Changes in Bargaining Power

Given the TCE view that hazards can, in general, be foreseen and therefore protected against, the question arises as to why contracting parties cannot foresee the buildup of bargaining power by one party or the other and construct their contractual commitments accordingly. Indeed, recognizing the potential for changes in bargaining power and adjusting governance mechanisms to offset them is TCE’s primary concern. Changes in bargaining power, however, are often difficult to foresee, because there is a large number of interrelated factors that affect the relative power of contracting parties. For example, union power may wax and wane with the state of the general economy, with government labor policies, with union leadership and organization, and with immigration rates. Thus, during the recession of the 1970s and early 1980s, union power declined in the United States, and union membership shrank. As the U.S. labor market has tightened in recent years, unions have become, arguably, less accommodating to the interests of employers.

Another difficulty in anticipating changes in bargaining power is that these changes may occur very gradually over long periods of time so that their future importance is difficult to perceive. For instance, according to North and Thomas’s (1973) seminal economic history of Europe, small changes in land/labor ratios in Europe accumulated over time into a large increase in the bargaining power of peasants over landlords, giving rise to the wage labor system. More generally, North (1990) emphasizes that such historic changes in relative prices have produced changes in bargaining power that lead to difficult and costly recontracting—or even full-scale institutional change.

An additional problem with foreseeing changes in bargaining power is that the enforceability of many kinds of contracts through the courts can be highly uncertain. Contract disputes, therefore, can result in sudden changes in relative bargaining power, as was the case for Naugles in its franchise contracts. Moreover, legal rulings may prevent a firm from avoiding even the most overt attempts by other parties to gain bargaining power. For instance, Toyota Motor Corporation recently sued to prevent Republic Industries from purchasing any more of its dealerships, after Republic announced its intention to purchase a significant share of Toyota’s outlets (Wall Street Journal, 1997b). In 1996 Toyota added terms and conditions to its dealership contracts, seeking to limit multiple purchases. However, the legality of these additional terms and conditions has not yet been tested in court. It is possible that Toyota will be unable to legally prevent the accumulation of bargaining power by Republic.

Even if some of these contingencies could be foreseen, however, the range of feasible governance mechanisms is unavoidably restricted in transactions involving employees and consumers. In particular, integrating parties’ interests through hierarchical governance—the mechanism of last resort when parties are highly interdependent—does not apply to relations with employees and consumers. Because human assets cannot be “owned” in the same way that physical assets can be owned, employees must, out of force, be dealt with through contractual arrangements, which will always be somewhat incomplete (Grossman & Hart, 1986; Klein, 1983).

Similarly, a firm cannot vertically integrate all of its final consumers. Hence, contracts with employees and consumers suffer from a condition of incomplete internalization, so a firm’s ability to impose hierarchical governance on transactions with these parties will always be attenuated to some degree, increasing the likelihood that it will be held up in future periods. In sum, it is difficult for firms to completely avoid hazards owing to changes in bargaining power, just as it is difficult for them to avoid contractual commitments.
Other Views of Power

In the current version of TCE, firms are generally assumed to operate in a highly competitive environment, in which sources of inefficiency, such as differences in bargaining power, are competed away relatively rapidly. As a result, there are relatively few instances in which the firm’s selection of governance arrangements is constrained by bargaining power. Moreover, in TCE theory the assumption is that changes in bargaining power can be foreseen at the initiation of a contract. For instance, when transactions involve a specific investment, the party making the specific investment might lose bargaining power after the investment is made. However, this party is assumed to be able to foresee the buildup of power of its transacting partner and to be able to act to effectively prevent it by negotiating suitable safeguards in the original contract, such as exclusive dealing provisions or holding hostage equity positions (Pisano, 1989; Williamson, 1983). According to TCE, if such safeguards cannot be written because critical contingencies cannot be foreseen, hierarchical organization will replace long-term contracting in order to forestall the emergence of asymmetrical bargaining power (Klein et al., 1978; Williamson, 1975, 1985).

While proponents of TCE generally eschew the concept of power, Williamson (1995) does admit that power can sometimes play a role in agreements between firms and workers and between firms and final consumers, especially when consumers and workers face collective action problems. Our arguments suggest, however, that rather than constituting special and unusual cases, labor and consumer transactions exert an important and pervasive influence on the boundaries of the firm and on the contractual nature of those boundaries. Furthermore, a firm’s implicit and explicit contracts with managers also can constrain the firm’s choice of organizational boundaries so that contracts with employees in general can be constraining. If so, this implies, at the very least, that evaluations and tests of TCE hypotheses about firm scope should be made contingent on the level of organization of workers and consumers involved in a given transaction and the nature of prior commitments to managers. We take the view that although competitive processes may possibly extinguish bargaining power in the long run, bargaining power differences can exert important impacts on governance choice in the short to medium term—impacts that researchers and managers cannot afford to ignore.

One theory of governance choice that explicitly features power is resource dependence theory (Pfeffer & Salancik, 1978). In this theory organizations choose governance forms, such as long-term contracts or mergers, in order to avoid dependence upon other organizations for critical resources, or to acquire the critical resources necessary to create dependence on the part of other organizations. Our theory of governance inseparability shares this emphasis on the impact of power on governance choice, but it differs from the resource dependence approach in several respects.

For instance, in resource dependence theory firms acquire power by using ploys to dupe other firms. Donaldson (1995) argues that the ploys described in resource dependence theory are often transparent, and he criticizes the theory for assuming excessive naivete. In contrast, in our theory firms are alert to power-seeking attempts by others, so power can only be generated in environments where it is difficult to predict changes in bargaining power. A second difference between our arguments and those of Pfeffer and Salancik (1978) concerns the sources of power. We follow the standard economic approach in associating bargaining power with monopolistic or oligopolistic market structures created by entry barriers and/or collective action (Bain, 1956; Scherer & Ross, 1990). In resource dependence theory, however, opportunities for power are not generated by particular market structures but simply by de facto needs for inputs or other resources. We take a more economic view of power, in which firms are assumed to be perceptive (boundedly rational) and gain power from difficult-to-anticipate changes in underlying economic structures.

SOME IMPLICATIONS OF GOVERNANCE INSEPARABILITY

Governance inseparability has important implications for the transaction cost theory of the firm. These implications concern three areas. First, governance inseparability has implications for predicting the relationship between the characteristics of individual transactions and the mechanisms that are used to govern them.
Second, governance inseparability has implications for the theory of the limits to the scope of the firm. Finally, considerations of governance inseparability allow a closer theoretical relationship to be forged between transaction cost theory on the one hand and theories of competition and industry evolution on the other.

Use of Alternative Governance Mechanisms

As we have pointed out in our main argument, the condition of governance inseparability implies that, in many instances, the governance of a given transaction will not be based solely on the characteristics of that specific transaction. Instead, if the set of governance options available to one or both parties in a transaction is constrained, the transaction may be governed in ways that are suboptimal if the characteristics of that transaction alone are considered. Thus, we would expect to observe a considerable degree of empirical variation in the ways in which a given type of transaction is governed. For instance, we might observe that a given type of transaction sometimes is internalized within a firm and sometimes is governed by a long-term contract. Hence:

**Proposition 1:** Different firms may govern identical transactions in different ways, as long as each firm is also a party to other types of transactions.

We would expect these differences to be most pronounced in settings where firms of different ages are active in the same transaction set, because older firms tend to be more encumbered with past commitments, constraining their ability to switch or differentiate their internal governance mechanisms. This has two implications. On the one hand, it implies that an older firm may be obligated to externalize transactions involving firm-specific assets, whereas a younger firm with few outstanding commitments may be able to internalize such transactions efficiently according to the usual TCE arguments. This is because in the older firm it is more likely that internalizing new types of transactions will abrogate commitments in place that are shaped around existing transactions. On the other hand, older firms are more likely to be engaged in various types of contractual commitments with internal parties, limiting their ability to outsource activities that do not involve firm-specific goods or assets, if similar transactions have been internalized in previous periods. Hence:

**Proposition 2a:** Compared with younger firms, older firms more often will be obligated to use market contracting to govern transactions featuring asset specificity for the same level of firm bargaining power.

**Proposition 2b:** Compared with younger firms, older firms more often will be obligated to use hierarchical mechanisms to govern generic transactions for the same level of firm bargaining power.

Note that since younger firms tend to have less bargaining power than older ones (e.g., because of capital constraints generated by uncertainty), our propositions control for firm bargaining power.

The degree to which a firm can switch or differentiate its governance mechanisms also will depend on the legal jurisdiction in which it operates. For instance, some countries, such as Germany and France, accord greater bargaining power to labor unions than do other countries. This greater power of unions will tend to engender a greater degree of governance inseparability than where union bargaining power is weaker.

**Proposition 3:** Firms operating in jurisdictions in which labor unions are accorded more bargaining power will be obligated more often to use hierarchical mechanisms to govern generic transactions than will firms operating in jurisdictions in which labor union power is more restricted.

Limits to Firm Scope

One of the most important questions posed by Coase (1937) in his pioneering article on the theory of the firm—a question that remains—concerns the limits to the size and the scope of a firm. If firms form because markets are sometimes costly to use, why are not all transactions that are more economically governed by a hierarchy governed by one large firm? One response to this question, offered by Williamson (1985), is that incentives within firms become increas-
ingly attenuated as the size of a firm increases. Consequently, beyond a certain size, the costs of incentive attenuation within a firm will outweigh the benefits of hierarchical governance. Milgrom and Roberts (1988, 1991) offer an alternative explanation: firms, precisely because they are organized as hierarchies, promulgate influence activities by managers and thereby incur costs that markets avoid.

The argument we present in this article offers a third answer to Coase's (1937) question. Because governance inseparability limits the governance options available to any particular firm, a single firm can only engage in a limited set of transactions that can be more or less efficiently governed by its particular set of feasible governance options. Otherwise, transaction costs will outweigh the economic surplus earned. Hence, while each firm may strive over time to economize on transaction costs, this process is subject to a set of historically determined constraints that play an important role in determining future firm growth. As a result, firms will become specialized to both particular types of governance arrangements and to particular types of transactions. This implies that, ceteris paribus, a firm will face transaction cost limits on both vertical integration and diversification, if these expansion paths call for internalizing transactions that cannot be governed efficiently by its existing organizational arrangements. Hence:

**Proposition 4:** The greater the difference is between a transaction's optimal governance mechanism and a firm's governance arrangements in place, the greater the cost will be to the firm of internalizing that transaction.

A second implication of our arguments for the theory of the scope of the firm relates to the optimal scope of a firm under differing levels of uncertainty. TCE theory holds that when uncertainty is high, internalization will be more efficient than market contracting because market contracts will be excessively costly to write, govern, and enforce (Grossman & Hart, 1986; Williamson, 1975, 1979). Considerations of governance inseparability, instead, imply that firms may be more costly than markets when certain aspects of the firm's transacting environment are uncertain. In particular, internalization may be highly inefficient if there is a high degree of uncertainty about (1) whether the relative efficiency of one governance form over another will change significantly in future periods (which would call for a switch between governance forms for a given transaction) or (2) how the scope of a given firm may need to change in future periods (which would call for governance differentiation).

Balakrishnan and Wernerfelt (1986) have argued that in environments characterized by frequent technological obsolescence, the incentive to vertically integrate is reduced because the expected returns to investments that are irreversible and transaction specific are lower. Teece (1992) argues that internalization of activities may be excessively costly in the presence of technological uncertainty because it is costly per se for a firm to absorb new activities. Our arguments suggest that internalization is also costly, because it may be impossible to undo in later periods and may engender commitments that spill over to other activities. Hence, following point 1 above, we would expect that firms operating in environments where (for example) production technologies are changing rapidly in terms of firm specificity, inseparabilities, or other types of economic spillovers would be cautious about internalizing production activities that might engender governance inseparability in later periods. Similarly, following point 2 above, we would expect to observe that firms that possess highly extensible resources, such as firms that carry out discovery research, would be more cautious about entering into early commitments than firms with fewer diversification prospects. For instance, a new R&D-based firm might prefer to outsource complementary functions until it can identify which activities would be most profitably internalized in terms of both rent generation and governance costs, taking inseparability into account. This is because a firm that internalizes a relatively low-valued activity early on, and shapes its contractual commitments around that activity, may later find it difficult to differentiate its internal arrangements to accommodate other more highly valued activities. Thus, for reasons of governance inseparability alone, we would expect the following:

**Proposition 5:** Greater uncertainty will reduce the vertical and horizontal scope of the firm.
Competition and Industry Evolution

It is important to note here that our predictions do not depend on an assumption that competitive forces are attenuated. Although we expect governance inseparability to engender suboptimal governance, the total value of a given transaction is the net value of (1) the economic surplus yielded by that transaction and (2) the costs of conducting the transaction. Thus, even in a competitive environment, a profitable transaction will endure, even if it is burdened to some degree by excess transaction costs. Of course, over time we would expect to observe a trend toward increased economizing on transaction costs, as argued by Williamson (1991). However, it may well be that the sources of surplus (i.e., rents) may be eroded long before a firm is able to optimize its governance arrangements.

Our view of the firm as a system of commitments provides some institutional underpinning to Lippman and Rumelt’s (1982) model of competition. In this model the authors assume firms to be “locked in” to a particular cost function that is drawn randomly from a given distribution. Our arguments here suggest that this lock-in results, in part, from contractual commitments. For instance, in the early stages of an industry, some firms may make choices to produce a given input internally, whereas other firms may choose to outsource this input. Over time one or the other choice may prove to yield a higher net surplus. However, it may be extremely difficult for any firm to change its governance arrangements over time, if its initial governance choices involved entering into commitments that engender governance inseparability. Thus, because commitments may foreclose future opportunities to adjust to more efficient governance modes, the governance choices that firms make early in their development may well determine their long-run competitive success.

Our arguments are also relevant to addressing the phenomenon of organizational inertia. Population ecologists argue that organizations are inherently restricted in their ability to adapt to changes in circumstances so that processes of organizational birth and death will contribute more to industry change than will processes of organizational adaptation (Hannan & Freeman, 1977, 1984). Evolutionary economists have taken a similar view (Nelson & Winter, 1982). We take a somewhat more moderated view of inertia: organizations will be inert according to the degree that the contractual commitments they entered into in earlier periods constrain their subsequent governance options, either by directly engendering governance inseparability or by exposing the firm to changes in bargaining power that result in governance constraints. We assume here that these early commitments are made in at least a boundedly rational way and are not merely the result of random processes. However, when circumstances change, the constraints on governance choices that follow these commitments may restrict a firm’s ability to adapt to changes in circumstances.

Hence, contractual commitments place limits on Lamarckian adaptation but do not preclude it. Accordingly, we would expect that change in organizational arrangements will take place only very slowly within surviving firms. Alternatively, when environments are highly selective, change will take place through the death of firms that are burdened with inescapable governance inseparabilities, to be replaced by other less-constrained firms. Nonetheless, it is important to note that inefficiency is not necessarily part of this account. Rather, our theory suggests that organizations in general, and firms in particular, will demonstrate “weak-form path dependency,” a condition in which history matters because it is too costly to reverse—not because it generates inefficiencies per se (Leibowitz & Margolis, 1995). But, whatever the efficiency implications, we submit that governance inseparabilities often have an important impact on governance choices and must therefore be accounted for in a positive theory of governance.

REFERENCES


Scherer, F. M., & Ross, D. 1990. *Industrial market structure*...


Nicholas S. Argyres received his Ph.D. in economics from the University of California at Berkeley. He is an assistant professor in the Department of Management and Organization, Marshall School of Business, University of Southern California. His research interests include strategy and organization structure, management of technology, contracting, institutions, and organizational politics.

Julia Porter Liebeskind received her Ph.D. in policy and organization from the Anderson Graduate School of Management, UCLA. She is an associate professor in the Department of Management and Organization, Marshall School of Business, University of Southern California. Her research interests include firm scope, corporate governance, restructuring, and intellectual property.