Are Capability-Based Theories of Firm Boundaries Really Distinct from Transaction Cost Theory?

Boston University
School of Management

Nicholas Argyres
Todd Zenger
Abstract

A large literature has developed over the past 10 years or so that attempts to compare transaction cost and capabilities explanations of firms’ vertical boundaries. Much of this literature has treated comparative capabilities (buyers’ vs. potential suppliers’) as determinants that are independent of transaction costs, based on the idea that capabilities theories of the firm are distinct from the transaction cost theory of the firm. In this commentary, we argue that this approach is mistaken. We argue that capabilities and transaction cost determinants interact with other dynamically, and that the two theories of the firm cannot be conceptually distinguished. Our argument carries implications for theories of the firm, and for empirical research aimed at testing those theories.
Introduction

In recent years an active theoretical and empirical debate has emerged around the topic of the boundaries of the firm. While transaction cost economics (Williamson, 1975; 1985; Klein, Crawford, & Alchian, 1978) has long dominated this debate, particularly in the empirical literature, scholars have more recently advanced competing arguments based on knowledge and resource-based logic (e.g., Demsetz 1988; Kogut & Zander 1992; Langlois 1992; Argyres 1996; Conner & Prahalad 1996; Foss 1996a, 1996b; Poppo & Zenger 1998; Madhok 2002; Leiblein & Miller 2003; Jacobides & Hitt 2005). Most of this work positions capability explanations as independent of, and often competing with, more traditional transaction cost explanations for the boundaries of the firm. On the surface, these two theories do seem to propose starkly different explanations for boundary choice (e.g., Coombs & Ketchen 1999; Leiblein & Miller 2003; David & Han 2004; Jacobides & Hitt 2005; Jacobides & Winter 2005). Transaction cost logic argues that activities are internalized when the cost of governing the activity through the market exceeds the cost of governing it with the internal hierarchy of the firm. By contrast, capabilities logic explains the choice to internalize (or the persistence of this choice) as a reflection of superior capability to perform the activity within the firm relative to the capabilities of external providers.¹

Articulated in this manner, these arguments appear quite distinct. In this commentary, however, we argue that a more careful look reveals the considerable

¹ Jacobides and Hitt (2005), for example, state that, “Firms with greater productive capabilities in a stage of production will tend to perform this activity internally, and contract with another firm through the ‘market’ where they are deficient. When skills or knowledge create an advantage for only one segment of a value chain, firms will tend to specialize; if these skills apply across multiple value chain segments, there is little basis for specialization at the firm or industry level” (p. 1210).
difficulty in conceptually distinguishing them. As evidence of this difficulty, scholars from both perspectives seem to easily interpret the same empirical results as supportive of their preferred theory (David & Han 2004; Carter & Hodgson 2006). Thus, a recent review of the empirical literature in transactions cost economics concludes that the majority of empirical make-or-buy studies testing TCE hypotheses can be “reinterpreted as … consistent with a competence or resource-based perspective (Carter & Hodgson 2006).” Our contention is that this conclusion is not surprising precisely because these theories’ predictions about the basic boundary choice cannot be conceptually distinguished. Thus, our agenda in this paper is to explain the lack of distinction between the two theories of the firm, and to highlight the incompleteness of arguments that rely solely on comparative capability, as first steps toward a more complete integration of the two perspectives.

We begin by noting that transaction cost theory, as developed by Williamson (1975, 1985) and Klein, Crawford and Alchian (1978), speaks directly to the question of how firms efficiently develop capability. This point has not been fully appreciated in the literature, in part because of differences in language. Scholars in the TCE tradition including Coase (1937), Williamson (1975; 1985), and Klein, Crawford, & Alchian (1978), have generally not used the term “capabilities” in their work. Nevertheless, TCE’s treatment of the canonical make-or-buy decision is easily applied – and indeed, was meant to be applied – to this fundamental question of how to efficiently govern

---

2 The tendency to pose transaction cost and capabilities explanations for governance choices as competitive also appears in the literature on strategic alliances (for a critique, see Oxley 2002). For example, alliance structures are sometimes said to be chosen over arm’s length contracts not for transaction cost reasons, but because they facilitate the development of interfirm routines for joint learning and knowledge transfer better than arm’s length agreements (Zajac & Olsen 1993; Sobrero & Roberts 2001).
capability development. While firms constantly seek new capability, they commonly integrate as a means to efficiently form or protect it.

Next, we argue that to fully understand the inseparable nature of these theories, we must examine the boundary choice in a dynamic context—exploring both the origin of a boundary decision and its persistence. Unfortunately, empirical explorations are typically cross-sectional and these dynamic issues are therefore largely ignored. We illustrate our points with an example from the Ford Motor Company in its early years. Following this, we ask whether an explanation for a boundary choice that is based entirely, or in large part, on capabilities considerations can provide sufficient explanation for a given make-or-buy decision. We argue that the answer is “no”, because any explanation based on capabilities alone (or on capabilities as the determining consideration) must consider why a firm doesn’t simply acquire any comparative capability it lacks, or divest itself of any comparative capability it possesses, for example, in the market for corporate control. If a firm cannot efficiently acquire capability or divest it, then we contend there are transaction cost reasons lurking as the explanation, implying that any capabilities explanation cannot stand alone.

We conclude by arguing that capabilities and transaction cost considerations, because they are so intertwined, should be integrated as part of a single theory of firm boundaries. We do not attempt a full scale theoretical integration here, but we do offer some first steps in that direction. We argue that a full integration first requires setting aside what we view as a logically unsustainable debate around the role of capability and transaction costs in explaining firm boundaries. We suggest that our efforts should instead focus on what we view as the common thread in these literatures: how varying
forms of governance facilitate capability and how firms learn to construct capability-generating governance structures over time.

**Capability Development, Firm-Specific Assets and Integration**

The resource-based view of the firm has primarily functioned as a theory of firm performance, arguing that the stock of unique, valuable, and difficult-to-imitate resources possessed by the firm determines its competitive advantage (Wernerfelt 1984; Barney 1986, 1991; Peteraf 1993). However, more recently the use of resource-based logic has extended beyond explaining firm performance to explore the configuration and boundaries of the firm. For example, Argyres (1996) found several examples from a manufacturing firm in which relative firm capabilities appeared to be important drivers of vertical scope. Other scholars have empirically linked the presence of skill sets (Poppo & Zenger 1998) and specialized experience or expertise (Leiblein & Miller 2003; Hoetker 2005) to vertical integration decisions. Jacobides and Hitt (2005), in examining the evolution of firm boundaries in the mortgage banking industry, similarly found that firms with greater productive capabilities in one stage of production performed this stage internally, and contracted with an outside firm for those stages for which their capabilities were lacking. Thus, the fundamental logic that emerges from the application of resource-based logic in this context is the simple concept that firms determine make or buy decisions through a process of comparative capability assessment. As these comparative capabilities shift, so do firm boundaries. We contend, however, that these conclusions are not based on sound causal reasoning, and therefore explain neither the origin nor the persistence of the decision to integrate.
Transactions cost economics (TCE) addresses the make or buy decision from a fundamentally different perspective, using a distinctly different theoretical language. TCE highlights the comparative efficiency of governance forms. The theory argues that hierarchies possess distinct advantages over markets in overcoming fundamental hazards in exchange, specifically hazards which arise when exchange requires exchange-specific investments by buyers and/or sellers. While this focus on exchange governance rather than comparative capability suggests a wide divergence in logic, a more careful look reveals the considerable difficulty in distinguishing them conceptually. We contend that TCE, even as it was originally formulated, is appropriately viewed as explaining the efficient governance of “unique” or “firm-specific” capability development. Thus, the inability to effectively manage the process of unique capability development through the market prompts the firm to integrate. Integration in turn promotes the co-specialized or firm-specific investments necessary to generate comparatively superior capability. Conversely, the inability to provide market-like incentives within the firm that powerfully motivate the development of broadly applicable (non firm-specific) capability constrain the boundary of the firm. As we argue below, the exchange conditions that elevate transaction costs in markets and encourage integration are precisely those that enable capability development.

We define a firm’s comparative capability as its capacity to deliver a product or service at lower cost and/or with superior quality or features than alternative suppliers.

The origin of a firm-specific capability involves investments, which are of necessity

---

Note that our critique in this paper is not focused on the knowledge-based theory of the firm literature (Kogut & Zander, 1992; Conner & Prahalad, 1996; Madhok, 2002), which explains the boundary of the firm using logic distinct from the simple comparative capability argument. The argument in this literature focuses largely on the advantages that integration possesses in facilitating the learning required in capability formation.
specialized to the firm. These firm-specific investments create unique capability or, as Williamson (1975: p. 28) describes, “non-trivial cost advantages”.

Firm-specific investments, however, also create hazardous exchange when the development of capability is attempted across a market exchange. Thus, in the presence of such investments that aim to develop capability, exchange partners are “inclined to expend considerable resources bargaining over the price at which the exchange is to take place” (Williamson 1975, p. 28). Essentially, both parties will bargain over the ownership of this new-found “capability” and the stream of rents that it generates. Thus, when capability emerges in an exchange, it transforms what began as an exchange between a buyer and one of many homogenous sellers into an exchange between a buyer and a highly capable seller. This emerging capability to use Williamson’s language, “[transforms] a large-numbers exchange condition …at the outset …into a small-numbers exchange relation on account of idiosyncratic experience associated with contract execution” (p. 29). Moreover, Williamson argues that the reason “outsiders are not on parity with insiders [in term of costs] is usually because outsiders lack firm-specific, task-specific, or transaction specific experience. Such experience is a valuable resource and can be used in strategic ways by those who, by being awarded initial contracts, have acquired it” (p. 31). Thus, consistent with Williamson, integration is the means by which capability is efficiently developed and protected from appropriation.

The relationships among capability development, firm-specific assets, and hierarchical governance are clarified further in Klein, et al.’s (1978) articulation of the transaction cost argument. They focus on the development of appropriable quasi rents in an exchange that result from specialized investments. Appropriable quasi-rents are
defined as the value of using the firm’s assets in the current exchange compared to the value of using these same assets in an alternative exchange. Thus, Klein et al.’s (1978) fundamental prediction is that “as assets become more specific and more appropriable quasi-rents are created (and therefore the possible gains from opportunistic behavior increases), the costs of contracting will generally increase more than the costs of vertical integration” (p. 298).

We contend that these concepts of specific asset investments and appropriable quasi rents are very closely related to, if not synonymous with, the resource-based concepts of firm specific capability and their associated rents. In theory, a firm-specific capability can reside either within the boundaries of a focal firm or within the boundaries of a supplier. Thus, a firm specific capability exists when either a focal firm itself, or its supplier, enjoys an advantage (cost, quality, or otherwise) in providing the output of a particular activity to the focal firm. The magnitude of this firm specific capability is directly measurable as the scope of the appropriable quasi-rents from that activity—the value of using these assets to service the focal firm relative to the value of using them to service the next highest valuing buyer. Specific asset investments are made in an exchange precisely to promote the development of capability that produces these quasi-rents. Therefore, the magnitude of capability and its specificity to the focal firm is closely related to, if not precisely measured by, the scope of appropriable quasi rents. As these appropriable quasi rents increase, ex post bargaining opportunities between an external supplier and the focal firm escalate. Thus, like Williamson (1975, 1985), Klein et al. (1978) implies that the need to support the development of capability in an exchange drives the decision to integrate.
As an example, consider Wal-Mart’s unique and superior logistics capability, which is one source of the company’s historic cost advantage. Wal-Mart handles the bulk of the logistics with its suppliers of consumer goods, including trucking, warehousing, inventory management, etc. This capability involves many kinds of investments that are specific to a host of other activity choices and investments by Wal-Mart involving software, information technology, store locations, trucking assets, warehouses, and supplier locations. Because the asset investments that accompany many of these choices are highly co-specialized to a range of other activity choices by Wal-Mart, TCE logic suggests that these asset investments would be difficult to support had Wal-Mart relied exclusively on outside suppliers to promote them. Wal-Mart’s decision to develop an internal logistics capability, rather than relying on specialized logistics firms, or on the suppliers of consumer goods themselves, was therefore driven in part by transaction cost considerations. What is somewhat distinct about the logic described here is the acknowledgement of the multi-lateral nature of asset specificity. The decision to integrate (or keep integrated) a specific activity depends not on a bilateral co-specialization between two activities, but rather the multi-lateral co-specialization of a focal activity and a host of other activities involved in the production process of the firm. We address this topic more extensively below.

---

4 Zott and Amit (2006) discuss this example as a way to make the argument that there is indeed a distinction between the concepts of firm-specificity and asset specificity, and therefore between transaction cost and capabilities-based theories of firm boundaries. They argue that Wal-Mart’s logistics capability is an example of high firm specificity, but low asset specificity, and that because the capability is internalized, the example is explained by capabilities but not transaction cost explanations. Zott and Amit’s (2006) argument, however, only considers asset specificity in relation to exchanges between Wal-Mart and its suppliers of consumer goods. They do not consider the asset specificity involved in Wal-Mart’s exchanges with suppliers of other inputs, such as logistics software, equipment, labor, store management, real estate, etc. We would contend that there is considerable asset specificity among a wide range of activities that generate not only the logistics capability, but a wide range of other capabilities within Wal-Mart.
In summary, therefore, transaction cost logic can be understood to argue that efforts to generate unique capability through exchange with an outside supplier produce hazardous exchange conditions that promote integration. Thus, the desire to generate unique capability drives the decision to integrate.\textsuperscript{5} Efforts to generate unique capability through market exchange generate significant appropriable quasi-rents that markets have difficulty allocating efficiently. Their presence generates high contracting costs. Integration eliminates these costs (though it introduces a new set of costs) thereby promoting capability development and protecting it from expropriation.

**Interpreting Boundary Choices**

Thinking of the transaction governance problem as one of governing capability development highlights the difficulty in explaining boundary decisions from a simple comparison of capability. It also highlights a difficulty faced in interpreting recent empirical literature comparing capabilities and transaction cost explanations of firm boundaries. A common conclusion in this literature is that a past transaction is organized in a particular way (e.g., internalized or conducted through the market) at a given point in time more for comparative capabilities reasons than for transaction cost reasons. Interpreting observations of boundary choices in this static way, however, is highly problematic. The difficulty is that identifying comparative capability does not explain the origin of that capability and the role that a prior boundary decision played in

---

\textsuperscript{5} The knowledge based theory of the firm literature highlights the distinctive coordination advantages of hierarchy rather than hierarchy’s capacity to avoid market failure (Grant 1996; Conner & Prahalad 1996; Kogut & Zander 1996; Madhok 2002). But, of course, this may merely be describing two sides of the same coin. While TCE explains why markets can’t coordinate well due to market hazards and comparatively ignore why it is that hierarchies can, the KBV explains why hierarchies are superior in coordination, but it doesn’t fully explain why markets cannot replicate this (Foss 1996a, 1996b).
developing it. Moreover, as we discuss later, it fails to explain the persistence of a decision to integrate, once the capability is formed.

The distribution of specialized capabilities across firms and their buyers and suppliers at a particular point in time reflects a series of past decisions by these firms to either develop or not to develop capabilities internally. Thus, the possession of a capability today reflects a choice to internally develop (or purchase) that capability yesterday. These decisions, we argue, were likely driven by comparative governance or transaction cost considerations. Consider the following example. A firm decides to internalize an activity at time 1 because performing this activity with the desired level of capability requires highly idiosyncratic investments—investments that suppliers are reluctant to make in the absence of carefully crafted safeguards. Due to the high costs of contractually creating and enforcing these safeguards, the firm chooses to integrate this capability development. As these specific investments are made over time, the firm develops the desired, superior capability to perform the activity, so that by time 2, the capability is fully developed, leaving no outside supplier with a comparable capability. Thereafter, the firm continues to be integrated.

Which theory, TCE or capabilities logic, best explains the firm’s boundary choice at time 1 in this example? A static analysis of the decision at time 2 would assign a small role to transaction costs and a large role to capabilities logic, yet this would be misleading. In fact, the capability difference between the firm and its (potential) suppliers only arose from an integration decision made at time 1, which was driven by transaction cost considerations. The integration decision at time 2 therefore cannot be interpreted as
following simply from comparative capability, and not transaction cost logic. Yet some recent studies of make-or-buy decisions promote precisely this conclusion.

The question of why a capability was originally developed internally is different from the question of why it continues to be internally governed. At some initial point in time, before a specialized capability has been formed by a buyer, a simple comparative capabilities test may well suggest the need to exchange with an external provider with superior capability. A desire to generate unique capability through co-specialized investments, however, may cause the buyer to integrate the exchange, perhaps by purchasing the more capable external provider. Attempting to generate this unique capability through a market exchange generates high transaction costs. Once the unique capability is internally generated, and we apply the comparative capability test, that test clearly predicts the persistence of integration. However, the origin of the integration decision clearly does not hinge on comparative capability, but rather on the comparative efficiency of alternative governance forms in supporting capability development. Thus, in this context, we see no logical way to distinguish between capabilities and transaction cost considerations.

Illustration: Ford Motor Company

The dynamic interaction between transaction cost and capabilities explanations of firm boundaries we have just discussed is evident in the boundary choices of Ford Motor Company during the early decades of the U.S. automobile industry. Despite a well-developed supplier base in the automotive industry, Ford chose to vertically integrate aggressively into automobile components as it introduced its moving assembly line in
In order for this new manufacturing strategy to work, interchangeable parts had to be created, which in turn required highly firm-specific, single purpose machine tools — including tools that were customized to Ford’s Model T. These tools included transmission testing devices, double-ended tools for pressing tubes in radiator fins, and the like (Nevins with Hill 1954).

The firm-specific investments required to generate the machine tools necessary for assembly line production implies that high transaction costs played an important role in Ford’s decision to vertically integrate the production of components. Interestingly, Langlois & Robertson (1995), who examine this piece of Ford’s history in detail, do not entertain the possibility that transaction costs, as conventionally understood, played a role in Ford’s vertical integration decisions. Instead, they argue that it was the lack of capabilities among existing suppliers to produce the specialized machine tools that accounts for these decisions to vertically integrate. They argue that, “only the men of Ford understood the uses to which such machines would be put” (p. 53). This, however, is precisely the logic that Williamson and Klein et al. articulated for why transaction costs drive vertical integration choices. Firms integrate transactions when they seek the development of a firm-specific capability that necessitates a transaction-specific investment(s) by supplier(s). By integrating the supplier(s), a firm avoids the need to contractually manage these specific investments required for capability development. In this case, the specific investments required for an external supplier to understand Ford’s products and technology, and then to design and produce the idiosyncratic machine tools Ford demanded, were likely too high to permit the efficient use of the market.
As time went on, Ford persisted in its vertical integration strategy. Indeed, through the 1920’s and early 1930’s, Ford remained highly vertically integrated even as most other companies vertically disintegrated (even while they increasingly adopted the moving assembly line). Katz (1977), for example, provides evidence that several companies, but not Ford, began vertically disintegrating after 1926. Argyres and Bigelow’s (2006) analysis of a database of all U.S. auto companies during the late 1920’s and early 1930’s also reveals a decline in the vertical integration level of the average firm. Why did Ford buck the trend? One possible answer is that Ford’s capabilities in producing the specialized machine tools necessary for assembly line production had become superior to existing suppliers’. By the mid-1920’s, Ford had at least a 15 year lead on suppliers in producing this kind of equipment, enough time to have moved down the learning curve—well ahead of available suppliers. Precision metal cutting may have been one of the key capabilities in which Ford was superior, and likely contributed to the large reductions in component costs Ford achieved after integration (Williams et al. 1993; Langlois & Robertson 1995, p. 52). It seems plausible, then, that Ford’s decision to persist in its vertical integration strategy during the 1920’s and early 1930’s, even while competitors were actively disintegrating under similar assembly line production processes, resulted in part from past decisions to vertical integrate that were themselves driven to a significant degree by conventional transaction costs. Langlois and Robertson’s (1995) account of Ford’s development of metal stamping techniques follows a similar pattern: An initial decision to internalize production, in this case by acquiring a supplier named Keim, led to the development of new stamping capabilities that reinforced this
integration decision through time (Langlois & Robertson 1995, pp. 53-54). We view these results and interpretation as quite consistent with the logic of Williamson and KCA.

Sufficiency of Capability Explanations

Even if we put aside the question of a capability’s origin or development, comparative capability alone is not sufficient to explain the boundary choice or more precisely, the persistence of a boundary choice. Thus, suppose that an outside supplier possesses capability superior to the firm’s internal capability. Does this necessarily imply that exchange should occur with this external supplier and be governed through the market? Or, suppose that an internal supplier possesses capability superior to all external suppliers. Does this necessarily imply that exchange with this internal supplier should remain within the boundaries of the firm? We contend that the comparative capability test identifies the desired provider, but does not determine the desired form of governance. The choice of an optimal provider for a given activity and the choice of how to govern that exchange are arguably quite separable decisions. This is because firms can choose to sell or spin-off an internal provider with vast capability or choose to integrate a capable external provider. The fact that firms routinely sell or spin off internal capabilities and actively integrate when they lack capability is inconsistent with the simple application of comparative capability logic to the boundary decision.

Langlois (1992) and others have argued, however, that firms often cannot spin-off or acquire a capability on the same time scale as they are making make-or-buy decisions for particular transactions, and that as a result capabilities considerations alone can drive governance choices. For example, consider a firm that is developing a new product, but
lacks the capability to produce a highly unique and highly valuable component or tool inhouse. While this capability may be easily assembled from relatively generic assets and easily obtained in the marketplace, the time to develop or acquire this capability internally may be considerable. In the short run, we might expect the firm to rely on outside suppliers, even if the unique nature of the component puts the firm at risk of hold-up. Over the medium term, however, we would expect, following TCE, that the firm will develop or acquire its own capability, assuming its demand for the unique component continues. Moreover, the initial capability difference that is affecting the short term boundary decision is once again likely itself to have been determined by past boundary decisions in which transaction costs played a role. Thus, we suggest that capabilities explanations for firm boundary choices are never sufficient on their own.

Complementarity, Co-specialization and Boundary Choices

We contend that the question of optimal governance for a given activity rests not on the scope of capability embedded in that activity, but rather the degree to which that activity is both complementary with and co-specialized to (or would benefit from being co-specialized to) other activities within the firm. It is the desire to support capability formation and protect its ensuing rents from external appropriation that prompts integration. We note that there is a close connection between the concept of co-

---

6 Zott and Amit (2006) argue that developing a new internal capability quickly can be costly due to time compression diseconomies (Dierckx and Cool 1989), and that acquiring the capability from an external source (i.e., acquiring a supplier) can involve high acquisition costs. These considerations, they suggest, can cause capabilities considerations alone to be determinative of make-or-buy decisions. Time compression diseconomies and acquisition costs, however, are likely to be in part determined by transaction costs of various kinds. For example, internal capability development can be delayed because of haggling over specific investments by employees or other suppliers that are needed for capability development. Acquisition costs that stem from information asymmetry and related frictions in the market for corporate control can also reflect transaction costs (e.g., Williamson 1975).
specialization or asset specificity as discussed in TCE and general propositions from the resource-based literature about capability, competitive advantage, and rent generation. This latter literature argues that rent generation (or capability) arises within a firm when activities or assets are “complementary,” or “superadditive,” (Montgomery & Wernerfelt 1988; Peteraf 1993; Amit & Shoemaker 1993), or when activities have the “property that doing more of any subgroup of activities raises the marginal return to the other activities” (Milgrom & Roberts 1990). A rent generating competitive advantage itself is commonly viewed as stemming from “coherent combinations” of activities (Milgrom & Roberts 1990, p.111) or from an “integrated set of choices about activities” (Ghemawat 2005, p. 131). How do these concepts from the strategy and resource-based literature about capability and competitive advantage differ from the transaction cost literature’s concepts of asset co-specialization (Williamson 1975; Teece 1986)?

Several obvious differences in logic and framing are evident here. First, while TCE logic focuses on a dyadic or bi-lateral relationship to address optimal governance, the resource-based and broader strategy literature focuses on a multi-lateral set of relationships that combine to create rents. Second, while the resource-based and broader strategy literature focus on the concept of complementarity or fit, TCE focuses on co-specialization—actual investment that is specialized to other assets or activities within this bundle. Our contention is that while the presence of complementarity or fit among activities is essential to rent generation, it is the presence or absence of co-specialization between a focal activity and other elements in that bundle which determines whether integration of that focal activity is required. The distinction here is that one can easily have an activity or asset that is a critical part of the complementary bundle of activities
that is in no way co-specialized to these other activities. In this case, the focal asset or activity may be quite generic, with many different firms capable of providing these same assets or activities in precisely the same way. There is no transaction cost problem associated with using the market to manage this complementarity. Any effort on the part of the outside firm to appropriate rents from the complementary bundle of activity choices of which their asset are a part will result in their instant replacement by another firm possessing identical assets. By contrast, if producing rents from a complementary bundle of activities requires co-specialized investments—defined in this multilateral context as investments which have a lower value when combined with any other set of assets—then safeguards are required to protect the exchange and prompt such investments. Thus, when exchanges require substantial co-specialization to generate this complementarity, integration is more likely. In this manner, complementarity *per se* is a necessary, but not sufficient condition to explain integration.

Note that the above argument regarding boundary choice is quite independent of comparative capability. The key issue is not whether comparative capability lies within or without the focal firm, but rather the degree to which a comparatively capable activity generates rents (or could potentially generate rents) by being customized to particular other activities and assets. Thus, if a firm’s superior internal capability reflects co-specialization among activities and assets, then a decision to spin out these activities from the firm and govern their access through the market, unleashes significant transactions costs. A decision to place these exchanges in the market creates the potential for those owning or managing the newly independent assets or activities, to appropriate the quasi-rents which stem from the bundle of complements. The focal firm is likely to find
managing this hold-up problem through the market to be quite costly. By contrast, if the superior internal capability is embedded in assets or activities that are not specific to the focal firm, but are instead equally valuable to multiple firms, then contracting for access to this capability is quite easily managed through the market. In this instance, there are no appropriable quasi rents over which to bargain or haggle and the use of contracts may be quite straightforward. Indeed, outsourcing the exchange may unleash beneficial incentives that enhance the value generated by the capability. Despite superior internal capability, integration in this instance may be suboptimal. Instead, the entire future value of the rents which the activity generates should be obtainable through a simple sale of the capability or resource in competitive factor markets.

Analogous logic governs the circumstance in which an external provider possesses superior capability. Again, the governance choice depends on the nature of the external provider’s capability. If the external provider’s capability is broadly applicable to multiple buyers, then the exchange will likely be easily managed through contracts. If by contrast the superior capability of the external supplier reflects significant buyer specific investments, then this implies the presence of appropriable quasi rents and the potential for holdup in exchange. Market governance under these conditions is presumed costly and integration is likely.

We contend, then, that it is not really comparative capability that determines the boundary choices of a focal firm. Comparative capability only determines the desired provider. The preferred form of governance with respect to that provider is determined by the degree to which the capability in question is complementary or specific to other
activities within the focal firm. If the degree of such specificity is sufficiently high, integration will be the efficient choice.

**Governing for Capability and Capability for Governing**

Our arguments thus far imply that transaction cost and capabilities considerations do not have independent effects on firm boundaries. To the contrary, these two logics are tightly interwoven. Capabilities affect boundary choices due to transaction cost considerations, and transaction costs emerge as firms seek to form capability. We are convinced that to improve our understanding of boundary choice, we need a more complete understanding of the complex relationship between transaction costs and capabilities development. We contend that further explorations of this relationship should focus on two organizing concepts: “governing for capability”, and “capability for governing”.

“Governing for capability” identifies research that attempts to understand how governance choices impact the rate and efficiency of capability formation. The knowledge-based theory of the firm literature takes up this agenda directly, highlighting the virtues of the firm in supporting knowledge flows and knowledge creation (e.g., Grant 1996; Conner & Prahalad 1996). There are many other specific research questions to which the notion of “governing for capability” points. For example, with respect to alliances, what kinds of contractual safeguards facilitate the transfer of tacit knowledge and encourage the development of desired routines (e.g., Oxley 2002)? With respect to internal organization, what kinds of authority and incentive structures facilitate the development of new knowledge within the firm (e.g., Nickerson & Zenger 2004)? How
does informal organization reduce opportunism in ways that stimulates the emergence of new knowledge within the firm (e.g., Argyres & Mui 2006)? We suggest that research aimed at these kinds of questions will turn up insights into the relationships between governance choice and capabilities development.

“Capability for governing”, on the other hand, identifies research that aims to understand how firms learn to govern their internal and external relationships over time. Whereas capabilities-based theories of the firm have tended to emphasize that the key firm capabilities are technological in nature, recently scholars have begun to explore the possibility that firms can develop capabilities for governing activities in ways similar to those in which they develop production capabilities (e.g., Child 1999; Anand & Khanna 2000; Azoulay & Shane 2001; Dyer & Singh 1998; Mayer & Argyres 2004). Research on capability for governing thus aims to uncover insights about how firm boundaries evolve as firms, particularly those in emerging industries, hone their capabilities for governing internal and external exchange relationships.

**Concluding Remarks**

In our view, the growing strategy literature on the determinants of firm boundaries contains a misleading thread. A false dichotomy between transaction cost and capabilities theories of firm boundaries has emerged. As contributors to this early literature, we acknowledge our own contributions to the misunderstanding. If capabilities differences between a firm and its potential suppliers play a key role in determining the firm’s governance choices, it is more than likely that transaction costs lie somewhere at the roots of these differences. Therefore, we contend that scholars should cease to assume
the existence of a meaningful distinction between transaction cost and capabilities-based
theories of firm boundaries – one that can be used to drive empirical research on the
determinants of those boundaries. Instead, scholars should treat capabilities
considerations as inextricably intertwined with transaction cost logic, and should seek to
analyze aspects of this complex interaction. We hope that our comments will help clarify
thinking about the boundaries of the firm.
References


Argyres, N. and Mui, V-L. 2006. Rules of Engagement, Credibility, and the Political Economy of Organizational Dissent”, manuscript, Boston University


