THE EFFECTS OF INNOVATION ON VERTICAL STRUCTURE: PERSPECTIVES ON TRANSACTION COSTS AND COMPETENCES

CLAUDIO WOLTER
FRANCISCO VELOSO

Program in Strategy, Entrepreneurship and Technological Change
Department of Engineering and Public Policy
Carnegie Mellon University
5000, Forbes Avenue, Pittsburgh, PA 15213-3890 USA
e-mail: cwolter@cmu.edu
Phone: + 1 412 268-5607
Fax: + 1 412 268-3757

Please do not cite or distribute without permission of the authors.
The usual caveat applies.

1The authors would like to thank Sebastian Fixson, Claudia Gonzalez-Brambila, Glenn Hoetker, Michael Jacobides, Lamar Pierce, Filipe Santos, Brian Silverman, Peter Thompson, Antonino Vaccaro, participants at the INFORMS 2005 Annual Meeting and at the Academy of Management 2006 Annual Meeting, and two anonymous referees for helpful comments on earlier drafts of this manuscript. Financial support from the International Motor Vehicle Program is gratefully acknowledged. The first author also wishes to thank Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil, for the additional financial support provided through a doctoral fellowship.
ABSTRACT

We analyze how transaction cost economics and competence arguments determine vertical organization boundaries when responding to innovation. Existing perspectives and empirical evidence have been ambiguous because of conflicting tensions between the two frameworks and simplistic views of innovation. Using Henderson and Clark’s (1990) innovation categories and a careful review of both theories, we show that it is possible to reach a consistent set of predictions on vertical integration and to reconcile apparently conflicting empirical results.
Specialization of productive activities and more generally the theory of the firm have been of interest to economists ever since Smith (1776) published his seminal work. Using the famous pin factory example, he illustrates the phenomenon of increased productivity of the manufacturing system as compared to crafts production, the result of the division of labor that comes about with growth in the extent of the market.

Building upon this initial insight, many theoretical contributions from scholars in the fields of transaction cost economics (TCE) and resource-based theory have demonstrated how both transaction costs and firm capabilities jointly shape firms’ vertical boundaries in relatively static business environments. Moreover, strong empirical support exists for the independent effects of transaction costs and capability heterogeneity on firms’ vertical integration decisions (Argyres 1996, Jacobides and Hitt 2005, Leiblein and Miller 2003).

Unfortunately, our confidence in these results cannot be extended to the determination of firm boundaries in more dynamic environments, despite the fact that the importance of doing so was long identified (Globerman 1980). Most studies of vertical structure incorporating TCE and firm competences have neglected innovation, a phenomenon that is at the core of the Schumpeterian view of economic change and life-cycle theories (Afuah 2001, Cacciatori and Jacobides 2005, Teece 1996). Leiblein and Miller (2003, p. 842), for example, note that existing studies have failed “to account for the value associated with the ability to react flexibly to an uncertain future.”

Technological change reasonably complicates the task of identifying transactions costs, relevant competences, and how the interplay between them that will shape firm boundaries. Williamson (1985), for instance, admitted that, contrary to what happens in the static case, transaction costs emerging uniquely from asset specificity would be less than persuasive explaining vertical structure in an innovative environment. Incorporating concepts from a competence perspective such as learning and dynamic capabilities into a more general theory of innovation and vertical structure creates even more challenges. Empirically, moreover, dealing with technological change means obtaining longitudinal data on the already hard to obtain constructs of transactions and capabilities.

In spite of these difficulties, a better understanding of this relationship between technological change and firm boundaries is absolutely necessary. Not only may firms need to alter their boundaries as a response to exogenous innovations, but their vertical structure and associated links to external suppliers can also have a direct influence on its innovative activities (Teece 1996). Understanding these issues is also critical because there are apparently conflicting empirical evidence regarding the effects of technological uncertainty on firms’ integration decisions (Dyer 1996), i.e., whether uncertainty leads to integration
C. Wolter and F. Veloso


Because of substantial differences across industries (e.g., services versus manufacturing), as well as in possible ways of characterizing innovations, we recognize, as do Langlois and Robertson (1989), that a general theory of vertical integration in an environment of economic change might be presently unattainable. Yet, we hope to shed some light on these issues in at least one particular setting. In this paper, we analyze how transaction costs and competence arguments interplay and relate to vertical integration decisions in the context of various product innovations in manufacturing firms.

Hence, following recent studies combining TCE with the competence perspective to understand firm boundaries (Jacobides and Hitt 2005, Jacobides and Winter 2005, Sampson 2004, Silverman 1999), our main objectives in this paper are twofold: first, we intend to extend the positive theories of firm boundaries by trying to understand how market structure evolves as a result of firms altering their degree of vertical integration in response to exogenous technological shocks; second, we want to explore firm-level competences and TCE as the underpinnings of a normative theory of vertical integration and technological change, with the aim of informing decision makers about the strategic management of value chains, with innovation and competitive advantage as targets.

Treating innovation initially as exogenous to the firms and later as endogenous simplifies the exposition of arguments, and allows us to explore possible differences in the efficiency of governance structures between leaders and followers. We also believe it is possible to reconcile theory with the previously noted conflicting empirical evidence about the effects of technological uncertainty on the direction of firm-level integration. As such, a final objective of this paper is to suggest directions for future empirical work on the aforementioned theories, through the presentation of testable propositions.

To reach these goals, a careful definition of the analysis context is critical. Part of the difficulty in understanding the effects of transaction costs and competences on vertical boundaries in dynamic settings stems from a lack of clarity in characterizing the industry and, especially, the innovation which is being analyzed. Hence, we focus on multi-technology, multi-component industries, such as the computer or automotive industry, and develop our main propositions from the viewpoint of a systems integrator (or assembler) responding to a technological shock. We characterize each type of innovation

---

2Throughout the text, “systems integrators” and “assemblers” are used interchangeably. They denote a firm manufacturing a final product which requires upstream parts that can either be produced in-house by internal suppliers, or bought in the intermediate market from independent ones. Assemblers can be more or
using Henderson and Clark’s (1990) typology: incremental (or component), modular, architectural, or radical. Then, we look at how transaction cost and competence arguments simultaneously influence make-or-buy decisions for each of them.

Henderson and Clark’s (1990) conceptualization has the advantage of combining Teece’s (1996) autonomous versus systemic dichotomy with Tushman and Anderson’s (1986) competence-enhancing versus competence-destroying. In fact, considering the various dimensions associated with an innovation, it is highly unlikely that any one type of governance mode, be it internal organization, reliance on markets, or additional hybrid forms (e.g., joint-ventures), would be able to successfully deal with all possible forms of innovation (Conner and Prahalad 1996). We also believe that while innovation typologies have provided interesting ex post rationalizations and explanations for a number of economic phenomena, their true value as a tool in the social sciences will only be proved if they can lend themselves to the construction of ex ante testable empirical hypotheses.

The following section presents the transaction cost reasoning for vertical integration when innovation is present. This is followed by an analogous presentation of arguments from the competence perspective. Next, we develop our main propositions by discussing both perspectives as applied to Henderson and Clark’s (1990) typology of innovation and from the viewpoint of a systems integrator. This is further extended to include the endogenous development of innovations. The final section concludes, discusses the limitations of this study, and points to the directions where future research might prove more fruitful.

TECHNICAL UNCERTAINTY AND VERTICAL STRUCTURE

Transaction Cost Perspective

Building upon Smith’s (1776) work, Stigler (1951) realized that although the intra-firm division of labor is indeed a matter of productivity (Langlois 1988), activities characterized by production functions displaying increasing returns to scale could be organized

---

3 An autonomous innovation is that which can be introduced without the need to modify other components in the system or product architecture. In contrast, a systemic one would require significant modifications in other parts of the product. Similarly, a competence-enhancing innovation is an improvement building on existing know-how, whereas competence-destroying innovations are “so fundamentally different from previously dominant technologies that the skills and knowledge base required to operate the core technology shift” (Tushman and Anderson 1986). Hence, an incremental innovation is defined as an autonomous, competence-enhancing innovation; modular as autonomous, competence-destroying; architectural as systemic, competence-enhancing; and radical as systemic, competence-destroying.
more efficiently in new independent firms (suppliers). Therefore, the extent of the market would also determine the degree of integration in an industry.

The problem with the Smith-Stigler’s focus on production costs is its incompleteness. Without further assumptions about the costs of internal organization (Masten, Meehan, and Snyder 1989) and the costs of using markets for transacting, arguments about the division of labor do not directly translate into predictions about vertical integration (Langlois 1992). These governance costs were first identified by Coase (1937) as important determinants of firm boundaries. Thirty years later, the TCE paradigm developed and operationalized Coases’s concepts, thus emerging as a strong explanation of these boundaries (Williamson 1971, Williamson 1975, Williamson 1985).

TCE recognizes the incompleteness of all contracts, since firms can neither have perfect information nor possibly foresee all future contingencies. It also assumes that economic actors behave not only in a bounded rational form, but also opportunistically. Because contract renegotiations are costly, bounded rationality, opportunism, and relationship-specific investments expose one of the transacting parties to the hazards of ex post hold-ups (Williamson 1985). In addition, ex ante uncertainty creates strategic misrepresentation risks such as moral hazard to the parties (Williamson 1971). Either situation may lead to the non-realization of or underinvestment in otherwise lucrative relationships. The major sources of transaction costs identified in the literature are uncertainty, the frequency of contract updates, and asset specificity: physical, human, site, and dedicated assets (Williamson 1983). The higher the uncertainty, frequency, or specificity governing contracts in a relationship, the higher transaction costs will be. Therefore, firms would organize their boundaries to minimize these costs.

Let us start by examining in some detail which factors raise transaction costs in an innovative environment. As previously noted, asset specificity is one of these if the development of the innovation requires upstream or downstream firm-specific investments by one or more transacting partners. The completion of these investments, the so-called “fundamental transformation” (Williamson 1985), results in a small numbers bargaining situation that will expose at least one party to the hazards of opportunistic hold-ups or strategic misrepresentation. Vertical integration prevents unnecessary haggling, and uses fiat to harmonize interests, risk perception, expectations, and resource allocation (Williamson 1971). The frequency with which a transaction recurs is another incentive for integration, since it directly increases haggling costs.

Uncertainty in a general sense increases transaction costs because it causes contract updates and renegotiations. Even in the absence of opportunism or first-mover advan-
Innovation and Vertical Structure

Advantages resulting from the ex post small numbers bargaining, renegotiations do involve reasonable costs (Conner and Prahalad 1996, Mayer and Argyres 2004, Williamson 1971). Uncertainty also complicates the process of hostage exchange (de Figueiredo and Teece 1996). Thus, in principle, higher uncertainty will lead to higher degrees of vertical integration in an industry. Internal organization mitigates these problems by allowing firms to better cope with uncertainty through the use of adaptive sequential decisions (Langlois 1988, Masten 1984, Williamson 1975).

Nevertheless, uncertainty is not a unidimensional construct. Although sometimes it is decomposed in different ways in the literature, a common one is to consider two key dimensions: technological or product uncertainty; and demand, also called market or commercial uncertainty. To clearly separate both dimensions, demand uncertainty is considered here as that exogenous to the innovation process, i.e., uncertainty in market demand for a product due to factors other than the technology itself. Macroeconomic effects, regulatory uncertainty, and changes in consumer preferences would for example fall under this category. By the same token, technological uncertainty will include a market demand component caused not only by product price (or production costs) achieved with the given technology, but also by diffusion factors such as network externalities and bandwagon effects. In this paper, we are primarily interested in the effects of technological uncertainty.

Although empirical results have systematically supported the propositions that frequency of transactions, demand uncertainty, and especially asset specificity are positively associated with vertical integration, results are mixed to say the least when testing the influence of technological uncertainty on firms’ boundaries. Walker and Weber (1987), for instance, show that high technological uncertainty with thick upstream markets in the automotive industry led to the procurement of components in the market (contrary to TCE predictions), whereas there was no statistically significant influence with thin upstream markets. In contrast, Hoetker (2005) demonstrates that high technical uncertainty in the development of innovative notebook displays drove computer makers to integrate backwards.

Other studies have found weak or nonexistent links between technological uncertainty and vertical integration. In a study of markets for technology, Schilling and Steensma (2002), for example, divide technological uncertainty into two constructs: technological dynamism and commercial uncertainty (purely due to the technology itself). They show that whereas the former had no effect on the choice of governance mode, the latter led to licensing decisions (disintegration), rather than acquisitions (integration). Also, Poppo and Zenger (1998) find no relationship between technological uncertainty and the out-
sourcing of information services by major American companies.

One possibility for this lack of empirical agreement regarding technological uncertainty is that the competence perspective may play an important part in an innovative setting (different from that in a purely static one), diluting the influence of TCE as a driver of integration. This is to where we turn next.

**Competence Perspective**

From its original exclusive focus on transactions as the unit of analysis, TCE has since been broadened to include the minimization of both transaction and production costs. Although some of the initial efforts to combine transaction and production costs emerged within the context of corporate diversification decisions (Teece 1980, Teece 1982), the literature on vertical integration soon followed. Riordan and Williamson (1985) pioneered the incorporation of production costs into the minimization effort. This initial work considered production costs differences as a result of scale economies. But cost differences due to scale can also be seen as the result of transaction costs, since firms’ internal suppliers might find it harder than independent ones to sell parts to competing firms due to perceived hazards (Argyres 1996). Consequently, scholars began concentrating their efforts on the inclusion of the competence (or capabilities) perspective of the firm, which had more recently emerged in the strategic management literature to explain the observed heterogeneity in firms’ production costs, conduct and performance. This perspective also illuminates the concept of the market which rather than an abstract institution, is in reality comprised of other firms with their own idiosyncratic production capabilities and costs. The boundaries between a focal firm and the market is then that between the firm and all these external enterprises (Jacobides and Winter 2005, Madhok 2002).

Competences, employed here similarly as in Foss (1993), that is, in a general sense to refer to the resource-based view (RBV) (Barney 1986, Peteraf 1993, Wernerfelt 1984), evolutionary theory (Nelson and Winter 1982), and the knowledge-based view of the firm (KBV) (Kogut and Zander 1992), are much more than production costs or production functions. They involve low and higher-order routines, knowledge, skills, and learning (Langlois 1988), and are the result of technical and organizational components endogenously influencing each other (Madhok 2002). Absent capabilities, as Leiblein and Miller (2003) explain, firms facing the same type of transactions and having the same scale would end up with the same governance structure. Therefore, factoring heterogeneity in capabilities into make-or-buy decisions became the next logical step.
Heterogeneity in firm capabilities has a direct impact on make-or-buy decisions. Firms tend to specialize in activities where they have some comparative advantage (Jacobides and Winter 2005, Kogut and Zander 1992, Richardson 1972, Teece 1996). The incorporation of this competence perspective into the original TCE framework led Williamson (1999, p. 1097), for instance, to suggest that “TCE informs the generic decision to make-or-buy while competence hinges in particulars.” Langlois (1992) goes further to argue that although both transaction costs and capabilities matter, the former are short-run phenomena that tend to lose importance when learning takes place through transacting. A similar view is expressed by Foss (1993).

As previously remarked, empirical studies have confirmed the independent effects of capabilities and transaction costs on the choice of governance mode in many different industries. Studying the production of wires, cables and connectors, Argyres (1996) shows that in some cases asset specificity (human and site) is the dominant explanation of make-or-buy decisions, whereas in other situations production costs unrelated to scale are determinant. Independent effects of capabilities are also found, for example, in the semiconductor industry (Leiblein and Miller 2003), in the production of notebook displays (Hoetker 2005), and in the mortgage banking industry (Jacobides and Hitt 2005). A formal analytical model proposed by Jacobides (2006) suggests that capabilities could be the main drivers of make-or-buy decisions in certain manufacturing sectors, whereas empirical results from the information technology sector showed that transactional issues dominated these decisions in this context (Mayer and Nickerson 2005).

When dealing with innovations, however, it becomes necessary to move beyond a static representation of capabilities to a broader competence perspective of the firm including resources, knowledge, skills, routines, and learning, in order to address the effect of technological uncertainty on firms’ boundaries. According to Langlois (1992, pp. 107–108), this competence perspective should include a “real-time account of production costs where knowledge and organization have as important a role as technology.”

Even though Williamson (1971) had already recognized the coordination potential of the firm, the knowledge-based view explicitly argues that the raison d’être of firms is to create and transfer information and know-how efficiently (Kogut and Zander 1992). According to the KBV, firms in an innovative environment characterized by technological uncertainty internalize activities not because of disadvantages (transaction costs) associated with market exchanges, but rather because of the benefits of coordinating ex ante, and within the same organization, complementary (and similar) activities (Conner and
Prahalad 1996, Kogut and Zander 1992, Richardson 1972).\(^4\)

These advantages arise because organizations develop routines and language tools through training (Armour and Teece 1980, Williamson 1971) and repeated interpersonal relations\(^5\) to economize on communication costs (Hoetker 2005, Kogut and Zander 1992). These routines become highly efficient in transferring knowledge, or in tackling tasks that require quick adaptation (Argyres 1996, Langlois 1988).

Internal organization also benefits from the power of managerial fiat (Williamson 1971) to solve problems of information asymmetry, which can arise even without opportunism. Technological uncertainty might result in honest parties unintentionally keeping some knowledge private, leading to disagreements about the optimal allocation of tasks and resources (Conner and Prahalad 1996). This issue is related to the presence of high degrees of tacit knowledge.\(^6\) Because tacit knowledge can more easily be appropriated by and transferred within a single firm, vertical integration would once again be the preferred outcome to prevent knowledge leakage (Argyres 1996, Christensen, Verlinden, and Westerman 2002, Teece 1982). As a matter of fact, integration has been seen more generally as a strategic move to enhance the appropriability of knowledge in regimes of weak intellectual property (Teece 1986).

Christensen, Verlinden, and Westerman (2002) seem to agree with the KBV perspective. They argue that when customers are underserved by a technology whose future path is still uncertain, firms are constantly trying to push the possibility frontier of their products with interdependent designs. This interdependency, meaning that interfaces within the product architecture are not well specified, so that changes in design specifications of one component can easily affect additional ones, involves a lot of tacit knowledge and requires unstructured technical dialog (Monteverde 1995), which is best accomplished within the boundaries of a single firm.

Up to this point, this competence perspective of the firm has basically led to the same prediction as that from TCE regarding technological uncertainty and vertical structure.

---

\(^4\)These benefits are net of bureaucratic costs incurred by organizations (Coase 1937). Although Argyres (1996) argues that there is no reason to believe that firms have bureaucratic disadvantages when compared to markets, most scholars seem to be of the opposite opinion, taking firms to be on the higher side (Williamson 1975, Masten 1984, Balakrishnan and Wernerfelt 1986, Nickerson and Zenger 2004).

\(^5\)Williamson (1999) takes issue with the interpersonal relations argument, since these can also be developed during long-term interfirm transactions. Helper, MacDuffie, and Sabel (2000) call this interfirm development process learning by monitoring. However, empirical evidence suggests that the firm has indeed advantages in the development and efficiency of these relations (Masten, Meehan, and Snyder 1989).

\(^6\)Note that both information asymmetry and tacit knowledge can be seen to increase transaction costs due to contractual hazards in addition to benefiting internal organization (Schilling and Steensma 2002).
That is, the more uncertainty, the more vertical integration one should expect to observe in an industry. However, there are alternative reasonings that point in the opposite direction and which, as we suggested before, could be confounding empirical results on the effects of technical uncertainty on vertical integration.

In a regime of rapid technological change where the innovation trajectory is uncertain, the unique intra-organizational language and routines that are so efficient in creating and transferring knowledge might become the wrong ones. Integration will then promote an inflexibility or rigidity that might hinder the firm’s effort at successfully acquiring new knowledge and innovating (Afuah 2001, Langlois 1992, Leonard-Barton 1992, Poppo and Zenger 1998). Furthermore, investments not only in new knowledge and routines, but also in semi-specialized capital goods or other tangible resources face a high risk of quick technological obsolescence (Balakrishnan and Wernerfelt 1986, Schilling and Steensma 2002).

Integrated firms may also miss out on innovations simply by being tied to a single internal supplier (Langlois and Robertson 1989). Independent suppliers have greater incentives than internal ones to expand their knowledge and skill bases, keeping abreast of new technological developments so as to continuously expand their customer base (Nooteboom 1999, Poppo and Zenger 1998). Consequently, intermediate markets may offer a pluralism of technological alternatives in the form of capabilities and resources that any firm can access at any time. And this heterogeneity of available technological trajectories may confer evolutionary advantages to disintegrated firms, if they can indeed benefit from accessing distant and dissimilar knowledge and resources, as predicted by the Schumpeterian view of innovation as recombinations (Langlois 1992).

Hence, according to these latter arguments, vertical disintegration should be the logical outcome of technical uncertainty. Although some scholars such as Langlois (1992) believe that the evolutionary benefits provided by the market, and which function analogously to Marshallian external economies (Nooteboom 1999), do indeed outweigh the benefits of vertical integration previously discussed (both in TCE and competence terms), this is still an open empirical question.

We are not arguing that TCE is wrong when claiming that uncertainty (technological included) increases transaction costs across the board in an industry, thus promoting higher degrees of vertical integration. Much to the contrary, we agree that even holding asset specificity constant, higher uncertainty will lead bounded rational actors to spend more resources on contractual contingency clauses. And in the event these break down, time-consuming renegotiations where the possibility of opportunistic behavior is present.
will be necessary between parties. These costs could then be avoided through common ownership, where disputes can be resolved by fiat. An integrated vertical structure would, in addition, benefit the transfer and coordination of knowledge between units of an organization.

Nonetheless, we also argue that technological uncertainty by itself, and as frequently used in the empirical literature, is an insufficient construct to characterize technological change, and to predict the direction of vertical integration in an industry. Technical uncertainty might or might not lead to more integration (as illustrated by the empirical results cited on pp. 7–8), depending on whether firms’ capabilities are overturned, and how strong are the incentives to access new capabilities in the market. It thus becomes necessary to understand innovations in a more fine-grained manner.

In summary, it is possible to see that although TCE always predicts an increase in integration incentives with technological change and uncertainty, the competence perspective brings within itself an intrinsic tension to the way firms should respond. This tension arises from the conflicting needs of transferring and protecting knowledge and capabilities within the firm, and that of accessing new or alternative competences and knowledge bases in the market, thus mitigating the risk of technological obsolescence. Therefore, in some cases it is possible that the competence view will reinforce TCE call for integration, whereas in other cases it will be pulling in the opposite direction.

In order to understand when each opposing force dominates the competence perspective and how it combines with TCE to generate integration or disintegration incentives at the firm level, we proceed by analyzing the above arguments in view of the four kinds of innovations described in Henderson and Clark’s (1990). Studying firms’ responses to exogenous innovations is important as these affect the structure, conduct and performance of these firms in an industry. We also believe that it is possible to explain the conflicting empirical evidence regarding the effects of technological uncertainty on vertical integration by focusing on the particulars of an innovation.
INNOVATION AND VERTICAL STRUCTURE

Responding to Exogenous Shocks

Initially, we are interested in how the degree of integration in an industry will change due to one of the four exogenous technological shocks in the typology. In other words, the interest here is in the response of systems integrators to innovations arising outside their boundaries, i.e., innovations emerging from sources external to the industry, from independent upstream suppliers or within a competing more integrated assembler. We assume that regardless of the innovation type and where it emerged from, alternative technological solutions always emerge in the marketplace from independent suppliers. And, provided that their capabilities have not been overturned by the innovation, imitation by integrated assemblers is also possible. That is, replication occurs because the appropriability regime is weak.

The easiest case to analyze is that of incremental innovation in the value chain. This is an innovation where the core concepts of components are reinforced and there are no changes in the linkages or interfaces between these components (Henderson and Clark 1990). Therefore, upstream capabilities in the industry are not overturned by the innovation, wherever this is developed. Moreover, any investments in additional capabilities to replicate the innovation have very low risk of obsolescence, as there are no major shifts in the innovation trajectory. As a consequence, integrated assemblers have no a priori incentives to disintegrate and use the market, as they can use their internal suppliers to play catch up.

For a less integrated assembler, internalization incentives arising from incremental innovations are not significant either. First of all, the nature of the innovation implies that there are not many needs for adaptation and knowledge transfer between the assembler and its supplier. In addition, relatively low technological uncertainty translates into low transaction costs. Furthermore, long-term relationships with independent suppliers

---

7How the industry has reached its present vertical structure, i.e., through some long-term process of optimization and selection, coupled with some measure of path-dependency, is irrelevant to us.

8An implicit assumption in this paper is that integration or disintegration occur as a result of actions taken by assemblers, and that upstream suppliers do not have the necessary systems integration capabilities to integrate forward.

9Teece (1996) considers a weak appropriability regime the common situation in most industries.

10We are naturally holding other factors that could also raise transaction costs and lead to more integration such as frequency of transactions and demand uncertainty constant in this analysis. The same is true of asset specificity. We believe that there is not enough evidence to pose a relationship between the nature of innovations and the asset-specificity of new investments. Hoetker (2005, p. 78), for example, wrote that
helps to decrease these transaction and coordination costs even more through reputational effects, learning within the relationship (Mayer and Argyres 2004), and the establishment of a common inter-organizational language (Langlois 1992).

Hence, vertical scope in an industry suffering an incremental technological shock is determined primarily by capability differentials. And since these capabilities are not overturned by the innovation, industry structure should not change. In a sense, this is the closest to a static analysis. Integrated assemblers will tend to remain integrated and use their internal suppliers (a sunk investment) to replicate the innovation, whereas disintegrated ones will still turn to the market where they can access a number of technological alternatives without the costs of integrating.

Figure 1 depicts these arguments. The right side represents integration incentives. Transaction cost arguments, represented within the right-hand side dotted box, call for greater integration the more uncertain is an innovation. These are reinforced by the knowledge-based view of the firm, part of the competence perspective. We follow Globerman (1980) in considering that both transaction costs and coordination needs concurrently and monotonically increase as we move towards more “complex” innovations, i.e., from incremental to modular, architectural, and then radical. In other words, moving from one extreme to the other simultaneously increases the need of suppliers and buyers to exchange more information, as well as the hazards associated with possible hold-ups and contract renegotiations that increase transaction costs. Since both effects lead to equivalent integration incentives, they mirror each other and are represented inside the same solid box in the upper right hand side of Figure 1.

In contrast, the left side of Figure 1 illustrates what happens with upstream capabilities in the industry after an innovation is introduced, thus determining incentives to access the market, i.e., disintegrate. As explained in the previous section, this is also a dimension of the competence perspective in an innovative setting, though distinct from the KBV. The dotted box on the left represents the two opposing aspects of this perspective.

\[ \text{“interviews indicate that beyond the specificity present in any innovative [notebook] display, increasing uncertainty impedes communication and the ability to contract for technical contingencies much more than it increases asset specificity.”} \]

Moreover, if asset specificity increases with technological uncertainty, our analysis would still be valid. The same would be true of monitoring costs (Alchian and Demsetz 1972).

\[ 11\text{This assumption results from our focus on make-or-buy decisions, but should not be generalized. As Sampson (2004) reports, in the context of R&D alliances between possible competitors, transaction costs stem mainly from the hazards of proprietary information leakage. And these hazards depend on the similarity between the capability endowments of the firms in the alliance: firms with very similar or dissimilar endowments have little to fear, whereas those in the intermediate range are the most at risk. In this context, it would then be possible for transaction costs to move in the opposite direction of coordination needs.} \]
Figure 1: Framework for the evolution of industry scope after a technological shock.

When reading the figure for the case of incremental innovation, starting on the right side of the diagram one can see that transaction costs and coordination needs are relatively low. Moreover, they can be reduced by previous relationships with independent partners whose capabilities are still useful. Thus, incentives to integrate are low for more disintegrated assemblers. On the left side, one sees that these disintegrated assemblers have incentives to remain as buyers, whereas at the same time more integrated ones have no reason to disintegrate, as their upstream capabilities have mostly been preserved by the innovation. Therefore, the overall result is that industry structure tends to be stable.

Some empirical results are in line with all these arguments. Walker and Weber (1987), for example, demonstrate that in the procurement of simple automotive components, similar to incremental innovations, capabilities were the most important factors driving make-or-buy decisions. Hoetker’s (2005) analysis of the development of notebook displays also confirms that low technological uncertainty in new display technology, once again analogous to a context of incremental innovation, led to procurement decisions made on the basis of technical capabilities.
Therefore, we propose that

Proposition 1: In the case of incremental innovation, the degree of vertical integration in an industry will not change.

The second case we are dealing with is that of modular innovation, which consists of competence-destroying technological change for upstream suppliers in the value chain without an upset of the architectural knowledge of systems integrators (Henderson and Clark 1990).

Because upstream capabilities are overturned, and assuming that the market will provide a pluralism of technological alternatives, integrated assemblers have strong incentives to access the market, and thus disintegrate (see Figure 1). Arm’s-length transactions will provide the required new capabilities, and at the same time avoid the risks stemming from technological investments associated with internal development or an acquisition. It is not as much that the technology itself will quickly become obsolete in the near future, but rather that the particular trajectories that will prove successful in this industry are not completely known at the point when the innovation is first introduced.

On the other hand, integration incentives are also higher than in the case of incremental innovations. Although the technology might already be familiar to other industries, it may in fact be new to the focal industry. Therefore, assimilating it into products will include a certain degree of technical uncertainty which will accordingly raise transaction costs.\textsuperscript{12} In addition to that, changes in the specific interface between the system and the new module (or subsystem) will be required. As a consequence, closer coordination between assemblers and suppliers will be needed, especially if it is necessary to transfer tacit knowledge embedded in the new technology.

Furthermore, it is unlikely that long-term relationships will be able to play a role in decreasing these transaction and coordination costs. Because technological developments are thought to be path-dependent, with the trajectory leading to related technologies (Teece 1996), one could expect modular innovations to emerge in the market from new entrants to the industry, making use of new or distant knowledge bases (Jacobides and Winter 2005). Assemblers will then have to work on new relationships and, as Mayer and Argyres (2004) show, contractual learning is very specific to the bilateral relationship in place: there are few spillovers from one relationship to another, so that switching suppliers means starting on a new learning path.

\textsuperscript{12}We are obviously eschewing pure modularity here, which by definition would require no assimilation.
In summary, modular innovations bring together both integration and disintegration incentives. Therefore, it is not possible to theoretically predict the overall direction of vertical scope in an industry subjected to this kind of shock. In the real world, a number of external factors to our analysis, such as asset specificity (or lack thereof) and market uncertainty, might end up tipping the balance one way or another. The overall result at the industry level will also implicitly depend on the distribution of integrated versus disintegrated assemblers and the strengths of integration/disintegration incentives.

However, concentrating on the issues discussed so far also serves to enlighten us about their possible effects. Disintegration incentives would obviously tend to dominate with relatively lower transaction and coordination costs. Lower transaction costs would result, for example, from thick upstream markets. Similarly, lower coordination should be expected when the interface between the new module and the system core is relatively stable and well understood, so that the innovation consists almost of a new “plug-and-play” module. Knowledge transfer difficulties would also decrease with more codification on both supplier’s and assembler’s parts, and how much one understands the needs of the other. Moreover, the farther the new technology is from the assembler’s knowledge-base, resulting in the latter having little absorptive capacity (Cohen and Levinthal 1990), the more likely the assembler is to turn to the market.

In contrast, integration would tend to prevail with thin upstream markets, or a high degree of path-dependency in the incorporation of the new module, leading to ex post small numbers bargaining. High coordination costs would emerge from a novel interface between module and core, or from very different organizational languages spoken by the new supplier and assembler. The possibility of spillovers from this new technology to additional components or subsystems would also promote integration.

Existing empirical results have shown disintegration incentives dominating. Studies reveal that industries with modular products exhibit a relatively high stability of players at the systems integrator level, with entry and exit occurring at the supplier level as the result of modular innovations (Hobday 1998). In addition, these industries exhibit a progressive disintegration of the value chain. Examples of this phenomenon include flight simulators (Miller, Hobday, Leroux-Demers, and Olleros 1995) and aircraft engines (Prencipe 2000). In these examples, the main modular innovations were the substitution of digital or electronic components for older analog subsystems. If we relate this case to the contingencies discussed above, we can understand the empirical observations by recognizing that the technology was probably codified, and the new interface well understood, with a close to plug-and-play situation. The knowledge was also far from the assemblers’
core competences.

A complementary perspective on the direction of the vertical integration is suggested by Jacobides (2006). His analytical model shows that the correlation between firms’ upstream and downstream capabilities might be a good predictor of vertical scope in a manufacturing industry. The more positive the correlation, the more integrated the industry seems to become and vice-versa. In this sense, if the overturned upstream capabilities were previously positively correlated with the downstream (in our case, systems integration) capabilities in the industry, then by overturning them in favor of new technologies, a modular innovation may reduce this correlation to zero, hence promoting industry disintegration.

Thus, we argue that despite possible higher transaction costs and coordination needs arising from a modular innovation, the fact that upstream capabilities related to the particular module are overturned tends to dominate the reorganization of industry structure. To avoid the costs of internal development or of an acquisition, with the attached risks of technological obsolescence, previously disintegrated assemblers will likely remain so, at the same time that integrated assemblers turn to arm’s-length transactions. As a consequence of both effects,

*Proposition 2: For modular innovations, the degree of vertical integration in an industry will decrease.*

The third case is that of architectural innovation. Different authors have defined this type of innovation in somewhat different ways. Henderson and Clark (1990) define it as an innovation where the core concepts and components are reinforced whereas the linkages between these change. Langlois (1988) uses an analogous definition but calls it systemic innovation: changes in (at least) one component which requires substantial modification in components throughout the system (most often in terms of physical properties such as volume and area). In contrast, Davies (1997, p. 234) has a similar definition for systemic innovation, but at the same time defines architectural innovation as “changes in functions of components and subsystems and how they are controlled to realise a common system goal.” This is a broader definition that will not be considered here. Thus, in this paper we use Henderson and Clark’s (1990) definition of architectural innovation, and take the definitions of systemic innovation by Davies (1997) and Langlois (1988) as synonyms.

Because the core component concepts are not overturned in this innovation, the upstream capabilities that are present in the supply chain will retain their importance. And
upstream investments in modified subsystems have a low risk of becoming obsolete in the short-run. In other words, technological uncertainty resides in the new architecture which lies in the domain of the assembler’s systems integration capabilities. Therefore, similarly to the case of incremental innovations, the incentives to use arm’s-length transactions are low for integrated assemblers observing an architectural shock.

An important difference between architectural and modular innovations is that in the former case many interfaces (as opposed to only one in the latter) linking subsystems and modules have necessarily to go through some transformation and redefinition. This happens even though the working concepts of subsystems and components remain roughly the same. A new architecture can also introduce previously unknown interdependencies between components.

Consequently, this translates into significant knowledge transfers between suppliers and assemblers. The coordination required to achieve the necessary adaptation of the new architecture to customers’ requirements would probably be better accomplished within the same organization. And the possibility of contract renegotiations due to technological uncertainty concurrently increases transaction costs. As Figure 1 suggests, integration incentives in the case of architectural innovation are high enough that long-term relationships are not able to significantly decrease them.

Therefore, disintegrated assemblers can then integrate ex post to save on transaction costs and to better coordinate the development and engineering of subsystems within the new architecture without major concerns for missing technological alternatives in the market. In addition to their response, more integrated assemblers would for the same reasons tend to remain integrated after the technological shock.

In agreement with these arguments, Hoetker (2006) finds that in the development of larger notebook displays, which is akin to an architectural innovation since changes are needed to many components due to physical (spatial) and power management issues (Hoetker 2006, pp. 507), working with an internal supplier was extremely valuable for assemblers.

13 Uncertainty in architectural innovations is less about the newness of technologies and more about how the different components, modules and subsystems best fit and work together in the new product architecture.

14 Interestingly, he finds similar results when the display innovation was related to its resolution and not size, in which case it can be seen as closer to a modular innovation. But he also finds that switching suppliers was easier in the case of the modular innovation. Given the ambiguity of Proposition 2, this results is not entirely surprising. Hoetker suggests that in the context of notebook displays the costs of establishing a modular system in the first place might be preventing the outsourcing of higher resolution displays.
Hence, we propose that

**Proposition 3:** For architectural innovations, the degree of vertical integration of an industry will increase.

It is interesting to note that decisions to integrate or not can also be viewed in light of real options theory. Integration can be seen as a growth option, whereas disintegration and accessing the market can be seen as a deferral option. Experiential learning with similar technologies from a previous generation enhances the growth option of investing in the new technology. In contrast, higher risks of technological obsolescence increase the value of deferral options (Leiblein and Ziedonis 2007). Hence, in this context, modular innovations would make deferral options more attractive to firms, whereas architectural innovations would increase the value of growth options. Thus, our predictions seem to be in agreement with insights from this framework. Yet, since this is not the focus of our work, we do not pursue the link to real options theory any further.

Finally, we have the case of industries facing a radical technological change. Radical innovations are simultaneously competence-destroying for suppliers and systems integrators. They overturn not only the product architecture, but also upstream capabilities. Therefore, similarly to the case of modular innovation, previously integrated firms have strong incentives to benefit from the pluralism of new technological solutions offered by the market, thereby avoiding the risks of obsolescence and the creation of organizational rigidities. This latter point might be particularly important here because with radical innovations firms usually compete in product design, not costs. And pushing the possibility frontier to address underserved customer needs might entail following a wrong technological route, backtracking, and then following a different path (Teece 1996). Arm’s-length transactions do facilitate this trial-and-error process.

Furthermore, radical innovations signal a break from the previous paradigm and, as such, tend to emerge with interdependent rather than modular designs. Adaptation and experimentation to fulfill customers’ requirements for the new technology will be the most important actions taken by firms. Avoiding transaction and communication costs arising from the constant changing of specifications and, consequently, multiple interactions with suppliers becomes paramount. These interactions are necessary because problems posed by radical innovations are ill-structured and complex (Macher 2006), and the technical dialog between assemblers and suppliers will be unstructured, with a high content of tacit knowledge.
Complexity and tacitness also mean that an abstract and general technological representation that facilitates the division of labor is very difficult, and firms tend to rely on empirical methods (Arora and Gambardella 1994), something more easily accomplished in-house. Integration then benefits assemblers to the extent that time-to-market of the new product is critical, and working with an internal supplier can speed up product development. Thus, there are also strong incentives for the in-house organization of production.

Since both integration and disintegration incentives are high as illustrated in Figure 1, it is not possible in principle to theoretically posit how industry scope will evolve after the shock of a radical innovation. Many scholars have, however, suggested that innovations of a more systemic, interdependent, or state-of-the-art character call for backward integration into component manufacturing (Langlois 1992, Macher and Mowery 2004, Teece 1996). Jacobides and Winter (2005) also argue that industries experience reintegration phases when superior capabilities emerge from knowledge bases that are different from those supporting the displaced technologies.\footnote{Note that this argument is also true of modular innovation and yet we have posited a decrease in integration with this innovation.}

Afuah’s (2001) study of the introduction of RISC microprocessors in the workstation market is in agreement with this view. RISC was not only an architectural innovation from the point of view of workstation manufacturers, but also a radical, competence-destroying upstream innovation to chip producers. Afuah (2001) finds evidence that manufacturers that integrated into RISC microprocessor production after the technological shock did significantly better than those that remained disintegrated.

In line with this result, we propose that for radical innovations the benefits of in-house organization will outweigh its costs. In other words,

**Proposition 4: For radical innovations, the degree of vertical integration of an industry should increase.**

It is useful now to revisit the question of when, in the presence of technological uncertainty, the competence perspective reinforces TCE’s pledge for more vertical integration and when it pushes the industry in the other direction. In a nutshell, the last three propositions above suggest that: for modular innovations, the competence perspective not only suggests that the benefits of accessing market capabilities might be greater than the costs of knowledge transfers, but this difference might be strong enough to offset the effects of higher transaction costs; in contrast, for architectural and radical innovations, the benefits
of internal knowledge transfer and coordination seem to be greater than the value presented by foregone technological alternatives provided by the market. Consequently, in these cases the competence perspective ends up reinforcing the TCE push towards higher industry integration.

Table 1 summarizes these results. Next to the characteristics of each innovation, it first establishes whether there are increased ex post incentives to either disintegrate and access the market to avoid obsolescence, or increased incentives to internalize the transaction due to TCE and KBV arguments. When one type of incentives dominates, we indicate whether one should expect more or less industry integration. Question marks indicate when theory does not give a clear prediction, with effects within brackets corresponding to the direction of integration posited in our propositions, according to the limited empirical evidence available in the literature.

<table>
<thead>
<tr>
<th>Innovation</th>
<th>Characteristics</th>
<th>Disintegration incentives</th>
<th>Integration incentives</th>
<th>Predicted direction of industry integration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incremental</td>
<td>autonomous competence-enhancing</td>
<td>Low</td>
<td>Low</td>
<td>No change</td>
</tr>
<tr>
<td>Modular</td>
<td>autonomous competence-destroying</td>
<td>High</td>
<td>High</td>
<td>? (Decrease)</td>
</tr>
<tr>
<td>Architectural</td>
<td>systemic competence-enhancing</td>
<td>Low</td>
<td>High</td>
<td>Increase</td>
</tr>
<tr>
<td>Radical</td>
<td>systemic competence-destroying</td>
<td>High</td>
<td>High</td>
<td>? (Increase)</td>
</tr>
</tbody>
</table>

Table 1: The predicted effects of both TCE and competence arguments on vertical integration decisions as leading to our four propositions. Results within parentheses indicate effects expected to dominate according to previous empirical studies.

It is important to note that although heterogeneity in capabilities does not enter our model explicitly, it does so implicitly in two ways: first, by considering that there are integrated and disintegrated assemblers in an industry prior to a shock, we account for the response of firms with and without upstream capabilities to formulate our propositions. Second, these propositions are at the industry-level, so that firm-level competences could lead a few individual firms to act in the direction opposite to that predicted on the average. This would account for the high degree of heterogeneity in vertical structure observed in many industries (Simon 1991).

**Developing Endogenous Innovations**

So far we have discussed what the industry-wide response in terms of its vertical structure would be in the case of different kinds of exogenous technological shocks. In reality,
because more innovative firms are usually in a better position to sustain a competitive advantage, managers are probably more interested in a normative theory of vertical structure with innovation as an ex ante goal, rather than a positive theory explaining how firms respond ex post to exogenous technological advances.

For systems integrators, the internal development of innovations is contingent on the degree of vertical integration of its value chain. For example, a firm needs an upstream segment to independently develop incremental and modular innovations. Moreover, the vertical scope chosen by a systems integrator shapes its incentives and cognitive scope regarding capability improvement and innovation (Jacobides and Winter 2005). Therefore, the interest here lies in the management of the vertical structure of the value chain, from the point of view of the assemblers, as a means of generating the different types of innovations in Henderson and Clark’s (1990) typology.

Before proceeding to this analysis, it is important to note that an efficient value chain structure does not necessarily have to be the same for firms responding to an exogenous innovation and for those trying to develop new products. Being an innovator or first-mover entails facing higher risks and costs (e.g., higher R&D expenditures), but also higher potential payoffs than being a follower or second-mover.16 Therefore, while incentives, competences, transaction and bureaucratic costs considerations influence the structure of the value chain for firms trying to introduce innovations into their products, it is possible that these considerations affect vertical scope in a different manner from the previous discussed case of firms simply responding to innovations. High-powered incentives, for example, tend to be by definition much more important ex ante than ex post, so that firms can overcome cognitive and bureaucratic barriers that support the status quo and hinder innovation (Teece 1996).

To investigate these issues, we proceed in the same fashion as in the previous section, moving from one type of innovation to the next. Let us start once again with the case of incremental innovation. The question then is: are firms better off trying to develop incremental product innovations in-house or relying on upstream suppliers?

Naturally, this can only be answered with certainty if one knows a firm’s capabilities. Capabilities not in a static production sense, but rather those associated with the continuous effort to innovate. An internal supplier that can provide a constant stream of incremental innovations in components and subsystems is a potential source of competitive advantage, even in low appropriability regimes, and as such should be kept in-house.

---

16This is true even absent strong appropriability, but provided the innovator has access to complementary assets (Teece 1986).
However, developing and maintaining capabilities superior to those available from independent suppliers might prove difficult in practice. Independent suppliers are thought to have higher incentives to exploit new knowledge bases (Poppo and Zenger 1998), and to accumulate knowledge faster (Jacobides and Winter 2005), because the potential payoff of capturing market share in the form of orders from many assemblers is larger. In contrast, internal suppliers have a captive customer, but will lack the scale to take full advantage of the stream of innovations, if assemblers fear the hazards associated with buying upstream components from a competitor (de Figueiredo and Teece 1996). This translates into lower innovation incentives and higher inertia against change (Nooteboom 1999).

Since incremental innovations tend to proceed through directional or local search, these arguments are also in line with Nickerson and Zenger’s (2004) claim that markets are ideal to govern this search. Furthermore, as discussed in the previous section, incremental innovations are characterized by both low transaction costs and coordination needs, reinforcing the advantages of disintegration.

In summary, our initial conjecture of an innovative internal supplier might prove hard to encounter in practice. Although superior capabilities in incremental innovations might at some point in time reside with integrated assemblers, maintaining this technological edge in-house might prove to be difficult and costly. A strategy of using an internal supplier to replicate external incremental innovations might be feasible in the beginning, but if incentives to innovate are really that different, capabilities will eventually diverge in the long-run. Therefore, in a mature industry with an established dominant design and whose future technology trajectories seem to be dominated by incremental innovations, firms seeking to be innovative might be better off avoiding the bureaucratic costs of an in-house supplier and relying on independent ones with higher incentives.

However, by relying on independent suppliers, assemblers are not able to extract competitive advantage from these upstream innovations per se, as they are available to all firms. Having given up its upstream capabilities, assemblers need to develop new capabilities in supply chain management to substitute for the lost ones. Not only do they need to concentrate on their core competences (Prahalad and Hamel 1990) in product architecture and systems integration, but also to learn how to take advantage of the pluralism of market alternatives by pitting different suppliers against each other for contract bidding, thus guaranteeing a future provision of innovative effort on their part (Fine and Whitney 1999).

The case of modular innovation is somewhat similar. If developed internally, it could give a systems integrator enough performance differential to provide some first-mover advantage. The problem is that in addition to the incentive issues discussed in the con-
text of incremental innovations, the knowledge base behind modular ones tends to reside farther from a firm’s current capabilities. And because path-dependency and cognitive biases usually limit the search for new technologies to areas related to present competences (Teece 1996), integrated firms might find it very difficult to internally develop this kind of innovation.

Therefore, unless an assembler has an organizational culture and the right capabilities to search for distant technological solutions, it might be better off waiting for a modular innovation to emerge in the marketplace and then quickly moving to incorporate it into its product. As discussed in the previous section, this strategy will incur higher coordination and transaction costs, especially if innovators are new entrants to the industry.

In practice, one way assemblers deal with these costs is to acquire new knowledge associated with the innovation. Takeishi (2002), for instance, finds that in the Japanese automotive industry, automakers develop overlapping knowledge boundaries with their suppliers for the development of more innovative components and modules (closer in concept to modular innovations), even though their task boundaries clearly define them as buyers and not makers. They also develop a somewhat hybrid vertical structure (Dyer 1996, Helper and Sako 1995), where they work closely and simultaneously with a selective small number of competing independent suppliers. This decreases transaction and coordination costs, while at the same time providing the necessary incentives for suppliers to innovate by keeping them at arm’s-length and in check through the constant threat posed by competition. This group of firms formed by an assembler and its preferred supplier is named a vertical group by Shanley and Peteraf (2004), who also argue for their usefulness in more innovative settings.

The caveat is that for these hybrid governance structures to be successful, the assembler must ensure that suppliers in its vertical group are committed to being highly innovative, since modular innovations by definition entail the obsolescence of current capabilities. If these innovations always emerge from players outside the industry or simply outside the vertical group, the additional costs of maintaining this governance form might not pay-off. This would be equivalent to using the group only for incremental innovations, which can arguably be accomplished in the market at lower cost.

Hence, similarly to the case of incremental innovations, assemblers trying to incorporate modular innovations into their products might find disintegration beneficial and decide to concentrate their capabilities into value chain management, but with an additional twist: knowledge management rises in importance, with firms extending their knowledge beyond their task boundaries (Brusoni, Prencipe, and Pavitt 2001, Fixson, Ro,
and Liker 2005, Lee and Veloso 2007). In other words, assemblers seek to learn about novel and distant technologies that are not part of their current endowment (and perhaps not even part of their current suppliers’), but might provide significant benefits if incorporated into the product. 

Architectural innovations pose more interesting questions. A new architecture can not only provide a substantial advantage to an innovator in an established market (Henderson and Clark 1990), but also open new submarkets. Assemblers have then high incentives to come up with architectural innovations. The main issue rests once again on whether more or less integrated assemblers are in a better position to innovate.

On the one hand, integration benefits the transfer of information and ideas with internal suppliers whose capabilities and skills will not be much affected by the technological change. Integration will also save on transaction costs due to uncertainty and prevent leakage of information on the new architecture to competitors via independent suppliers. On the other hand, internal suppliers may create cognitive difficulties stemming from the information filters and communication channels embodied in the current architecture (Henderson and Clark 1990, Nooteboom 1999), thus delaying or preventing the innovation.

Which forces prevail in practice is currently an open empirical question. One would think that to break this inertia caused by in-house suppliers, an integrated firm would need an organizational culture that promotes outside-the-box thinking and innovation at every level. Technical employees would need to feel part of the larger organization as opposed to only of an upstream subunit. This would give them a stake in the technological development of the whole product and discourage the myopic protection of their own turf. Rotating employees through different units or component areas might help to give them a bigger architectural picture (Takeishi 2002). A firm that is capable to implement and maintain these organizational features would be better positioned to develop architectural innovations through close coordination with its internal suppliers.

Nevertheless, even if an integrated firm accomplishes this, it might still be at a disadvantage in the long run. If innovation at the architectural level occurs with much less frequency than the previous two types of innovation, as argued by Teece (2000, ch. 1), disintegration previous to an architectural shift might then be more beneficial in the long-run, since the firm can take advantage of the market for incremental and modular innovations in this period, as suggested above.

Yet, architectural innovation might be the one that most benefits from a hybrid supply
chain. Avoiding full vertical integration should decrease the cognitive problems discussed above, giving assemblers a better chance to innovate. And since upstream capabilities are not overturned, long-term relationships would be able to significantly decrease transaction and coordination costs between the assembler and suppliers. As suggested by Takeishi (2002) and Lee and Veloso (2007), an important condition for this arrangement to be successful is not only for assemblers to extend their knowledge boundaries toward outsourced components, but also for suppliers to extend their knowledge boundaries in the direction of architectural knowledge.

Finally, radical innovations. As has previously been discussed, product interdependencies require close coordination between assemblers and suppliers, at least for later stages of product development. Integration will provide this coordination as well as economize on transaction costs and prevent information leakage.

However, what is necessary in the very early stages of product development and whether more or less integrated assemblers are at an advantageous position is less obvious. Integration might have an edge if internal suppliers would be able to better promote the exchange of new ideas about emerging technologies (Hoetker 2005). Or because interdependency and complexity need heuristic or cognitive search, something better accomplished by hierarchies (Nickerson and Zenger 2004). Nevertheless, because upstream capabilities are overturned by radical innovations, internal suppliers might be in a weak position to support this new product development. And they could even be resistant to change (Teece 1996), becoming a serious liability and leading the assembler to miss technological opportunities in intermediate markets. Hybrid structures might be able to somewhat mitigate these hazards. Nevertheless, the obsolescence of many upstream capabilities, which would tend to render a long-term vertical group of not much use, and the infrequency with which radical innovations tend to come about would lead one to be skeptical of this solution in the long-run.

Based on the empirical results of his study of RISC workstations, Afuah (2001, p. 1225) suggests that firms “may be better off being vertically integrated into the major components that drive the discontinuity and better off not being vertically integrated into the components that drive the incremental change.” The problem is that it is very difficult to know a priori which knowledge bases and capabilities will drive the radical or discontinuous change. A possible solution to this conundrum is for firms to establish strong R&D organizations, capable of supporting current operations and, perhaps more importantly, of constantly scanning the environment for novel technological opportunities, and absorbing this valuable knowledge into the organization (Argyres and Silverman 1999).
This represents a type of superior dynamic capability that can lead to successful searches for solutions of the ill-structured and complex problems posed by radical innovations (Macher 2006).

Recapitulating, this section addresses the interesting notion that innovation and vertical structure might also be endogenously determined, as for example when more disintegrated firms get better at introducing incremental and modular innovations (Langlois and Robertson 1989), or when more integrated firms get better at introducing more systemic or interdependent advances (Sorenson 2003). As always, there is no free lunch. Firms intending to become innovators need to develop specific capabilities suited for each kind of innovation.

Disintegrated assemblers are required to improve the management of supplier relationships to extract maximum value from incremental or modular innovations introduced by independent suppliers. In the case of incremental innovation, it is simply a matter of putting requirements up for competitive bidding. For modular innovations, though, firms need to know more than they make to be successful (Brusoni, Prencipe, and Pavitt 2001). Only then, can they scour the market for new technologies and efficiently integrate them into their products. In contrast, although at new systemic trajectories disintegrated assemblers could have a cognitive advantage, they would then be at a disadvantage coordinating the implementation of a new architecture with a myriad of independent suppliers.

Integrated assemblers, on the other hand, would suffer from higher bureaucratic costs and cognitive limitations arising from their in-house suppliers in the case of autonomous innovations. They would also miss out on the high-powered incentives that flow to independent suppliers. Nonetheless, integrated assemblers can take advantage of their organizational form to avoid transaction costs and to extract the benefits of internal coordination associated with architectural and radical innovations. However, they need additional competences not only to deal with the cognitive biases and inertial forces posed by their structure, but also to absorb external knowledge and quickly change technological trajectories if the initial path in the innovation development proves to be the wrong one.

These arguments are in agreement with Jacobides and Winter’s (2005) claim that the process of capability development is conditioned on changes in vertical scope, something that deserves more attention in the literature. Comparative advantages would lead different types of firms to specialize and get better at specific types of innovation.

It is important to note, though, that hybrid or quasi-vertical integrated structures such as a Keiretsu would be able to mitigate some of the problems faced by disintegrated (in-
Innovation and Vertical Structure

tegrated) assemblers trying to develop systemic (autonomous) innovations. Hybrid structures have the ability to provide better coordination mechanisms and decrease transaction costs when compared to pure outsourcing, and also to present market incentives that are absent in hierarchies.

DISCUSSION

The determination of the vertical structure of industries is contingent not only on transaction costs, but also on firm-specific competences and the ability to adapt to technical change. While technological uncertainty existent in innovative environments always raise transaction costs, thus providing incentives for integration, the competence perspective brings in itself an intrinsic tension. On the one hand, the need for close coordination of activities and constant knowledge transfers reinforces TCE’s incentive for more integration. On the other hand, the dangers of following the wrong technological route, of asset obsolescence, and the possibility of accessing many alternative technological solutions in intermediate markets provide an incentive for disintegration.

One of the contributions of this paper is to carefully analyze these forces in the context of the innovation typology offered by Henderson and Clark (1990). We propose a theory and framework to investigate the direction of industry integration after an exogenous shock of the four kinds described in the typology. This direction will indicate which forces dominate make-or-buy decisions at the firm level in the context of technological change.

Another important contribution of this paper is to clarify the observed conflicting empirical results regarding the effect of technological uncertainty upon the division of labor. Because uncertainty is contextual in nature, different technical developments that are result of, say, similar incremental innovations, may be seen by an engineer responding to a survey as being characterized by different levels of technological uncertainty. Analogously, engineers’ perceptions about the magnitude of uncertainty across different innovation types might be deceptively similar. Both situations might lead to the contradictory empirical results previously described. If this is the case, tying technological uncertainty to the type of innovation will prove to be a better proxy.

Our four propositions explain the little empirical evidence available and imply that keeping this bigger innovation picture in mind may help to better analyze the effect of technological uncertainty on the evolution of an industry’s vertical structure. Consider,
for instance, Walker and Weber’s (1987) study of simple automotive components. As predicted by the discussion leading to Proposition 1, they find that existing capabilities are the most important drivers of make-or-buy decisions. Moreover, with higher technological uncertainty, indicating a transition from incremental to modular innovations, and when upstream markets are thick, i.e., there are plenty of technological alternatives, firms find it better to buy than make. This contradicts TCE and KBV predictions that uncertainty should lead to integration, but is accounted for in Proposition 2 by the benefits of accessing the market. A failure to consider the intrinsic characteristics of the different types of innovations may then explain, for example, the overall lack of effect of technological uncertainty on licensing decisions described by Schilling and Steensma (2002).

Although knowing how to respond to exogenous technological shocks, a form of ad hoc problem solving (Winter 2003), is important to managers and their firms, it is perhaps even more relevant to learn how to organize their firm’s vertical structure in order to become the innovator. Understanding what kinds of products, activities, and services to outsource prior to a determined innovative regime coming into place, and understanding how to extract maximum value from suppliers in terms of product performance and technical innovative solutions can possibly lead to long-term competitive advantages and, consequently, survival.

Our analysis indicates that the disintegration called for in Propositions 2 may also be beneficial to firms, if future innovations remain in the incremental or modular categories. The case for integration called for by Propositions 3 and 4 as a response to exogenous architectural and radical innovations seems to be at a minimum weaker when dealing with the ex ante effort to be an innovator. Unfortunately, the only empirical work dealing with similar issues is limited to Afuah’s (2001) study of RISC microprocessors, which interestingly finds ex ante disintegration benefits. The scholarship on vertical integration would greatly benefit from a more thorough investigation of this area.

In all four endogenous cases addressed, though, the right scope does not seem to be a sufficient condition for success in innovating. Firms also need to develop the right competences (e.g., value chain management, R&D organizations) associated with the related vertical structure. Since there is not a single governance mode capable of optimally dealing with all four types of innovation, these ancillary competences are paramount and need to be developed and protected overtime if a firm is to try to engage in all or most types of innovation.

The discussion about what happens with upstream capabilities after an innovation is another contribution of this paper. It complements Teece’s (1996) analysis of verti-
Innovation and Vertical Structure

Innovation as a function of autonomous and systemic innovations by dealing with the competence-destroying features of modular and radical innovations. Although our propositions still agree that autonomous innovations will be related to disintegration and systemic innovations with integration, Table 1 indicates that this might not be the case in practice and only future empirical studies will be able to verify this.

In summary, this paper makes an explicit effort to link the evolution of product architecture through its innovations to the vertical scope of an industry, with this relationship being moderated by transaction costs and firm-specific capabilities.

Interestingly, despite the fact that technological change is at the core of life-cycle theories, the literature explicitly linking innovation to the division of labor and industry evolution is limited. We know that specialization is a major feature of industry life-cycle (Stigler 1951), but one which cannot be explained solely in terms of economies of scale resulting from growth in the extent of the market. Cycles of specialization and reintegration are also observed in different industries (Macher and Mowery 2004).

Reintegration, for instance, seems to emerge when consumers become underserved by a certain technology and/or related capabilities (Cacciatori and Jacobides 2005, Christensen, Verlinden, and Westerman 2002). But these arguments are not sufficient and need to be tied to the innovations driving the evolution of product technology. Although we agree with Macher and Mowery (2004) that there is a lack of generalizable findings in the literature, we believe that a better understanding of disintegration and reintegration cycles can be obtained by studying the types of innovations that emerge in an industry. For example, to the extent that incremental and modular innovations account for the bulk of innovations after the establishment of a dominant design (Utterback and Abernathy 1975), the combination of Propositions 1 and 2 with the endogenous discussion would signal a pattern of decreasing vertical integration, as observed in most industries (Stigler 1951). Similarly, according to Propositions 3 and 4, the introduction of architectural or radical innovations would lead to reintegration phases.

This study is not without its limitations. Although we have mostly talked about make-or-buy decisions as a dichotomous choice, vertical scope can vary between these two extremes in a much more fine-grained manner (Jacobides and Billinger 2006). This is the reason, however, why our propositions speak broadly about the degree of vertical integration in an industry as increasing or decreasing. Future empirical work should be able to consider hybrid governance modes such as alliances and joint ventures as a raise or decline in integration depending on the starting point. Variations in tapered integration could also be included.
This paper also sidesteps issues related to knowledge spillovers between upstream and downstream segments (Macher and Mowery 2004). Spillovers might be important because they can bridge knowledge gaps, alter the distribution of capabilities and even allow upstream suppliers to integrate forward into assembly, a possibility not considered here. Similarly, we focus only on the response of assemblers and first-tier suppliers and ignore what happens to lower-tiered ones. Interestingly enough, a modular innovation, for example, could lead an assembler to disintegrate, and at the same time increase the level of integration of first-tier suppliers.

Testing our propositions will require careful empirical work and the use of panel data. One would need to obtain data on one product-system from an industry and observe its evolution through time, determining the types of innovations that have shaped its technological trajectory. Simultaneously, it would be necessary to know whether parts and subsystems were made or bought by assemblers.

Recent efforts to separate transaction costs from firm-level effects on assemblers’ make-or-buy decisions (Leiblein and Miller 2003, Poppo and Zenger 1998, Schilling and Steensma 2002) should be carefully expanded to a context of technological change, as the strength of these effects change with different types of innovation. For example, some constructs obtained from past survey instruments such as component complexity interdependency are thought to influence firm boundaries through both knowledge-based mechanisms and transaction costs effects. If data is somehow available on the performance or quality of components as in Takeishi (2002), then it would be possible to use a procedure similar to that in Poppo and Zenger’s (1998) study, running separate regressions for components manufactured in-house and those outsourced, before and after a technological shock, in order to understand how these constructs affect market and organizational costs, and influence integration.

Data on firms’ competences in manufacturing (patents have, for instance, been used in the literature) would also come useful to disentangle capabilities and resources from asset specificity effects on make-or-buy decisions. While the importance of asset specificity in static settings is well established, it would be useful to measure its influence as compared to the other effects analyzed in this paper for different kinds of innovation. Using constructs such as firm and industry standardization to proxy for asset specificity, a longitudinal data set might allow the use of simultaneous equation models (Schilling and Steensma 2002) to disentangle both effects. Great care should be taken trying to separate KBV arguments from human asset specificity. Learning and investments in specific skills will tend to improve knowledge transfer and coordination between suppliers and assem-
blers. However, language and communication barriers may persist due to cognitive issues and organizational cultures even after these investments are made.

As firm-specific capabilities will also emerge in time as the result of a past governance mode choices (Murmann, Aldrich, Levinthal, and Winter 2003), it would also be interesting to obtain data on which assemblers were the leaders in introducing the innovations for the components of the data set. This would enable us to differentiate between the vertical structure of leaders and that of followers for different types of innovations, shedding light on the endogenous innovation issues discussed in the second part of this paper.

Another interesting avenue for research would be to explore the differences between outside or independent procurement, and the use of vertical groups or a Keiretsu. In principle, transaction costs in general and fear of opportunistic behavior in particular would be much lower or non-existent between assemblers and K-suppliers who constantly interact to develop new products and technologies. On the other hand, it is not clear how Keiretsu and independent suppliers differ from assemblers in their respective organizational languages and codes. If K-suppliers enjoy no advantage in communicating with assemblers as compared to independent suppliers, then make-or-buy decisions between assemblers and K-suppliers would be solely based on knowledge-based mechanisms, whereas with independent suppliers these decisions would involve transaction costs considerations.

Obviously, this would not hold if K-suppliers are able to develop similar organizational language and codes as their affiliated assemblers. Or if long-term relationships and reputational effects between assemblers and independent suppliers are able to significantly decrease transaction costs as well. Using surveys to ask managers about the risks of opportunistic behavior might be a way to illuminate this issue.

Finally, any empirical model would need to control for additional factors or contingencies outside TCE and the competence perspective that might also influence the choice of firms’ boundaries in an innovative environment such as entry barriers (Williamson 1971), diversification strategies (Kogut and Zander 1992, Leiblein and Miller 2003), institutional isomorphism (Lamoreaux, Raff, and Temin 2003), the strategy of leading firms (Macher and Mowery 2004), path-dependency (Langlois 1988, Simon 1991, Teece 1996, Williamson 1999), and the need to pace innovations when network externalities are present (de Figueiredo and Teece 1996).

This careful work should nevertheless pay off by broadening our understanding not only of how capability distributions influence vertical scope, but also of how technological change affects the distribution of governance modes in an industry as a consequence of
the response of second-movers to the new challenges, and the ex ante organization efforts of the innovators.

REFERENCES


