SCOPE, BOUNDARY CHOICES AND PROFIT EVOLUTION
OVER THE INDUSTRY LIFE-CYCLE

CLAUDIO WOLTER
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

MICHAEL G. JACOBIDES
London Business School
Regent’s Park, London, UK NW 4SA
e-mail: mjacobides@london.edu
Phone: +44 (0)20 7000 8716; Fax: +44 (0)20 7000 7001

FRANCISCO M. VELOSO
Department of Engineering and Public Policy, Carnegie Mellon University
5000, Forbes Avenue, Pittsburgh, PA 15213-3890, USA
e-mail: fveloso@cmu.edu
Phone: + 1 412 268-4640; Fax: + 1 412 268-3757

Abstract
Recent research shows that to understand vertical scope we need to combine transaction cost and capabilities to look at all firm’s choices in terms of expanding vertically or continuing focused. The main implication is that industry competition dynamics, evolution, and vertical scope should be jointly considered. Yet existing research has not adequately studied the interplay between a sector’s evolution and firm vertical scope, and how these shape profitability. Previous models of endogenous scope have considered how profitability drives firms’ particular integration choices; yet little is said about profitability evolution as industry changes. This paper aspires to cover this gap, focusing on the profitability implications of scope changes, and how these relate to changes in firm’s underlying capability distributions. In particular, it sheds light into a sectors’ evolution dynamics, by looking at how modular and architectural innovations affect specialists and integrated firms and shape both the sector’s scope and profit distribution.

Key words: value chain, industry evolution, vertical scope, profit migration, modular innovation.

POMS 20th Annual Conference – Orlando, Florida U.S.A. May 1 to May 4, 2009
INTRODUCTION

Ever since Smith’s (1776) founding work, the understanding of specialization of productive activities and, more recently, the theory of the firm (Coase 1937) have been a central focus of economics and management research. In the recent past, several scholars in the fields of transaction cost economics (TCE) and resource-based theory have made theoretical contributions to our understanding of how transaction costs and firm capabilities jointly shape firms’ boundaries and market structure. It has, by now, been reasonably established that transaction costs and capability heterogeneity drive firms’ vertical integration decisions, and as such shape sectoral patterns (Argyres 1996, Jacobides and Hitt 2005, Leiblein and Miller 2003). More recently, work drawing on Langlois (1992) and Silver (1984) and blending evolutionary economics and institutional theory has started looking at how entire sectors change, shaping the transactional menus over time and driving firms decision in terms of scope.

Yet, for all the progress made in terms of theory and evidence, two important issues have not been tackled head-on. First, while we understand the dynamics that shape the overall choice towards integration or disintegration, both at the level of the individual firm and at the level of the sector, we do not have any direct theory that helps shape our understanding of how profits (of both integrated and disintegrated firms) evolve over time, given the changes in scope. And second, we have only a partial understanding of how vertical scope changes when technologies are dynamic – and when these technologies drive both capabilities and transaction costs. This paper aspires to build on recent research and address, using simulations from a computational model of industry evolution, both of these shortcomings.

To this end, consider first the nature of the dependent variable. Most studies of vertical scope consider integration (make or buy) as the unit of analysis. While profitability (or survival) are presumed to be driving the choices of firms in terms of the scope they adopt (with the belief that the hand of selection will weed out inefficient choices), there is much less clarity on how
profits and scope co-evolve, despite suggestions that over the life of a sector both change and both are causally connected (Christensen, Verlinden, and Westerman 2002, Jacobides, Knudsen, and Augier 2006). So, an interesting opportunity emerges there.

Thinking more seriously about profitability, though, makes us consider another issue that studies of scope have not looked at in depth – the issue of technological change. 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, Teece 1996), in particular, they fail to account for the value of being able to react flexibly to an uncertain future (Leiblein and Miller 2003).

Technological change reasonably complicates the identification and quantification of transactions costs, resources, and capabilities and the understanding of how their interactions shape individual firm boundaries and, at a higher level, market structure. Relying, for example, solely on asset specificity to explain vertical scope in innovative settings would be less than persuasive (Williamson 1985). Incorporating additional concepts from a competence perspective, such as learning (Argote and Epple 1990), dynamic capabilities (Teece, Pisano, and Shuen 1997), or time-compression diseconomies (Dierickx and Cool 1989) into a more general theory of innovation and vertical structure creates even more challenges. Moreover, technological change necessarily requires a much more extensive and robust longitudinal data collection effort on already hard to obtain constructs such as transactions and capabilities.

In spite of these difficulties, a better understanding of this relationship between technological change and firm boundaries is paramount. Not only do firms reactively adjust their boundaries as a response to exogenous innovations, but it is also important to recognize that their organizational form and associated ties to external suppliers can have a direct and endogenous impact on their innovative pursuits (Teece 1996). Understanding these issues is also critical because there is apparently conflicting empirical evidence regarding the effects of technological uncertainty on
firms’ boundary decisions, specifically whether uncertainty leads to integration or disintegration at the firm level (Hoetker 2005, Schilling and Steensma 2002, Walker and Weber 1987).

A better analysis of technological dynamics (which will affect both transactional choices, costs, and the underlying value of capabilities) can also nicely complement the understanding of profitability evolution, so that both of these issues can help expand our knowledge in a complementary form. An illustration of this can be seen in the pattern, noted in popular business press, of profits shifting across the value chain as it evolves over time. While there are a number of conjectures on the dynamics of “profit migration” (Slywotzky 1996), little research has accompanied them. Another critical perspective is the tension between the role of movements in the boundaries of existing firms and the entry of new firms. This especially important because we know that innovative shocks often arrive from players outside established industries (Prencipe 2000).

While a number of researchers have started to confront the aforementioned issues in dynamic settings (see Wolter and Veloso (2008) for a review), much remains to be resolved. While we are not trying to come up with a definite general theory of firms’ boundaries and profits in innovative settings, we do hope, nonetheless, that this paper will greatly advance such an agenda, by uncovering the relevant mechanisms linking scope and profitability through the analysis of an evolutionary computational model.

The following section contains the theoretical background. It presents the main arguments from both transactional and competence perspectives that are thought to shape vertical scope. It then discusses these issues in more dynamic and innovative settings and reviews the more recent literature in this direction. The section concludes by highlighting how the study of profitability has been neglected and why it is relevant, particularly in this context. Next, we develop our computational model, describing its main features, and explain how we use simulations to uncover the evolutionary pattern of interactions between variables that we seek to understand. The final section concludes, discusses the results and limitations of this study and points to directions where future
research might prove more fruitful.

THEORETICAL BACKGROUND

Transaction cost and competence perspectives on vertical scope

TCE recognizes the incompleteness of all contracts due to lack of perfect information or foresight and also assumes that economic actors behave rationally and opportunistically. Consequently, in the presence of uncertainty, frequent transactions or asset specificity, firm-specific investments by one or more transacting partners results in a small numbers bargaining situation. This so-called “fundamental transformation” (Williamson 1985) exposes at least one party to the hazards of opportunistic hold-ups or strategic misrepresentation.

Firms then organize their boundaries to minimize these risks and associated costs. Vertical integration, for instance, prevents unnecessary haggling, and uses fiat to harmonize interests, risk perception, expectations, and resource allocation (Williamson 1971).

Although empirical results have systematically supported the positive effects that frequency of transactions, demand uncertainty, and especially asset specificity have on choices to internalize production, results are mixed when the dependent variable is technological uncertainty.

Walker and Weber (1987), for example, find that high technological uncertainty in the automotive industry led OEMs to outsource component manufacturing to independent suppliers (contrary to TCE reasoning), particularly when upstream markets were competitive. In contrast, Hoetker (2005) shows that high uncertainty attached to innovative notebook displays drove computer makers to internalize their development and production. Sometimes, no influence from technological uncertainty on firms’ boundary choices was observed (Poppo and Zenger 1998, Schilling and Steensma 2002).

As Wolter and Veloso (2008) discuss, one possibility for this lack of empirical agreement is
that competences may play an important role in an innovative environment where technological uncertainty is rampant (a role distinct from that played in static settings), thus diluting the influence of TCE as a driver of integration.

A long-standing emphasis on heterogeneity in capabilities or resources is expressed in the resource-based view (RBV) of the firm (Barney 1986, Wernerfelt 1984). If firms draw on dissimilar resources, this makes them differentially effective, a fact that has become ever more credible with the continuing accumulation of evidence that firms in the same line of business do, indeed, typically display wide and persistent differences in their practices, productivity and efficiency (Lieberman and Dhawan 2000). In this sense, heterogeneity does have by itself a direct and independent impact on make-or-buy decisions. Firms do tend to specialize in activities where they expect to sustain some comparative advantage (Jacobides and Winter 2005, Kogut and Zander 1992, Richardson 1972, Teece 1996), a fact empirically verified in many studies and industries (Argyres 1996, Hoetker 2005, Jacobides and Hitt 2005, Leiblein and Miller 2003).

In an innovative context, however, a static representation of capabilities does not suffice. It becomes necessary to expand the competence perspective to encompass not only the RBV, but also (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). In order to address the changes in vertical scope induced by technological uncertainty, it is imperative to move beyond the simple analysis of production costs to a more complex one that includes routines, knowledge, skills, and learning as outcomes of technical and organizational components endogenously influencing each other (Madhok 2002).

But, in doing this, an intrinsic contradiction within this inclusive competence perspective arises (Wolter and Veloso 2008). One the one hand, firms in an innovative environment have incentives to internalize transactions because of the benefits of coordinating complementary (and similar) activities within the same organization (Conner and Prahalad 1996, Kogut and
These benefits emerge from intra-organizational routines and language tools that evolve from training (Armour and Teece 1980) and repeated interpersonal relations (Hoetker 2005, Kogut and Zander 1992). These tools and routines are particularly efficient in tackling tasks requiring quick adaptation (Argyres 1996, Langlois 1988). Managerial fiat, moreover, also allows organizations to quickly settle disputes arising from information asymmetry, as technological uncertainty sometimes results in honest disagreements (Conner and Prahalad 1996, Williamson 1971).

On the other hand, these unique languages and routines could end up promoting organizational rigidities that hinder the firm’s effort to acquire new knowledge in fast changing technological markets (Afuah 2001, Langlois 1992, Leonard-Barton 1992, Poppo and Zenger 1998). In these environments, firms face the risk that any investment not only in new knowledge and routines, but also in semi-specialized capital goods or other tangible resources will quickly become obsolete (Balakrishnan and Wernerfelt 1986, Schilling and Steensma 2002).

Furthermore, if independent suppliers have higher incentives to continuously expand their knowledge bases and to keep abreast of new technological developments (Poppo and Zenger 1998), then integrated firms, by being tied to its internal supplier, may find themselves at a serious disadvantage (Langlois and Robertson 1989). In other words, the possibility of accessing a pluralism of technological alternatives in intermediate markets may confer certain comparative evolutionary advantages to disintegrated firms. Especially if it allows these firms to access more distant and dissimilar resources and capabilities, as the Schumpeterian view of innovation as recombinations would predict (Langlois 1992).

To understand how and when these two opposing incentives, in addition to incentives to minimize transaction costs, affect firms’ decision making, Wolter and Veloso (2008) describe technological uncertainty in a more careful manner. They do so by analyzing these arguments in relation to the four types of innovations described by Henderson and Clark’s (1990) typology, i.e., incre-
mental, modular, architectural and radical, in the context of a multi-technology, multi-component industry. This framework leads in some cases to consistent predictions about how vertical scope will change (e.g., architectural), but not in others (e.g., modular innovation). In the latter cases, they use the extant empirical literature as guidelines to predict how and explain why vertical scope changes as a consequence of these innovations.

In the absence of new empirical evidence though, a powerful and useful tool to test and extend Wolter and Veloso’s (2008) propositions and one which has been gaining popularity is that of evolutionary computational models. Although mathematical modeling has been an integral part of the orthodox economics literature, it has not fared so frequently in the management and strategy fields. Addressing real problems that concern managers and practitioners substantially complicates models, since it means doing away with simplifications required to reach closed-form solutions. Moreover, moving beyond comparative statics to formal models of evolution dynamics significantly adds complexity, particularly when dealing with multiple vertically (or horizontally) differentiated industry segments or submarkets.

Fortunately, an alternative approach to orthodoxy has been advanced by Nelson and Winter (1982). Instead of relying too much on mathematical formalisms, firm-level behavior is modeled via simple rules (representing heterogenous capabilities and routines), so that industry-level behavior is obtained with a computational calculation of the interaction and aggregation of individual firms’ decisions and outcomes. This approach is particularly suited for evolutionary settings, where Schumpeterian competition might play an important role.

In this tradition, a few models dealing with vertical scope are worth mentioning. Jacobides (2008) developed an evolutionary model where firms comprised of two vertically linked segments choose whether to use their internal upstream segment or the market at large (or both) for the production of a component needed in the final product. Results show a long-run tendency towards specialization, where the dynamics is mostly driven by heterogeneity in capabilities, particularly
by the correlation between firms’ upstream and downstream capabilities, with transaction costs playing a moderating effect. Arora and Bokhari (2007) developed a similar model, but where firms can either make or buy the component, though not both. Moreover, in their formulation, the dynamics is primarily driven by entry and exit into both industry segments. They show that when potential entrants choose their organizational form based on independent and randomly assigned upstream and downstream capabilities, the likelihood on entry as an integrated concern decreases over time and also resulting in a tendency towards disintegration. Finally, in an update of a previous effort, Malerba, Nelson, Orsenigo, and Winter (2008) study the evolution of a sector’s vertical scope, but paying more attention to the effects that innovation may bring upon industry dynamics. Although their specific objective is to replicate the evolution of the computer industry in the United States, their effort has some similarities with our own.

In all these cases, however, although different modeling and mathematical techniques are employed, it is important to emphasize that the primarily dependent variable is the degree of vertical integration in the industry. Despite this similarity in focus, neither of these three studies mentioned above address an important piece of the industry dynamics puzzle: namely, how profitability emerges and evolves.

**Profitability and vertical scope**

Not that the study of profitability itself has been neglected. It is at the core of both strategic management and industrial organization research agenda. There are essentially three different intellectual traditions in regard to explaining how profit emerges. First, economists would contend that profits can emerge in an oligopolistic setting, in which firms manage to limit supply or competition, keeping prices high. Firms (and their owners) thus reap the benefits of tacit or explicit collusion. Despite some exceptions, research in this “non-competitive” branch of economics has focused much more on comparative statics than on dynamics.
Second, starting from Ricardo’s (1817) prescient analysis of the role of resources, a different branch of economic analysis maintains that profit is often due to rent generating scarce resources. Firms who own such resources may profit, even in the absence of strategic interactions, a theme followed and amplified by recent RBV research.

But despite the need to incorporate these inter-firm differences in analyses of firm profitability, most of the theoretical research on how these differences translate into profitability has focused on the conditions that allow firms to make profit (or appropriate rents) in equilibrium and on the identification of successful strategies associated with this outcome. It almost disregards profits that can accrue in the short run, when firms of differing capabilities co-exist within a and across sectors, before selection has led to the long run equilibrium.

And third, following Knight (1921) and Schumpeter (1942), a few economists, some strategists and most entrepreneurship scholars point out that differences in beliefs and expectations can lead to (mostly temporary) profits. Schumpeter stressed the role of profit as a motivator for economic activity and entrepreneurial action, whether by aspiring entrants or established firms. More importantly, though, he stressed that profits are often temporary, eroding over time, as should be expected, since imitation, excessive capacity growth and competition, or substitution, do not allow success to be particularly long-lived. Thus, even stable firm-level (or industry-level) profits, Schumpeter argued, should be seen as a succession of different “waves of creative destruction.”

While a large stream of research has built on Schumpeter, most of it focuses on understanding innovation and its determinants, or examines entry and exit over time (Jovanovic 1982, Klepper 1996, Pakes and Ericson 1998). But the question of how profit evolves over time in industries of dissimilar firms, which compete for markets and resources alike, has received limited attention (Lippman and Rumelt 1982, Lippman and Rumelt 2003, Winter 1995).

This substantial gap is nonetheless beginning to be filled. Jacobides (2006) and Jacobides, Winter, and Kassberger (2006) develop models that offer a significant elaboration of the equilib-
rium analysis, by addressing the process of adjustment from an initial profitable position towards equilibrium. The models explicitly assume firm heterogeneity, introduce resource pricing into the analysis and consider how these factors affect competitive dynamics.

These efforts are, however, focused on the equilibrium and disequilibrium conditions affecting a single industrial sector. As previously argued, a wave of creative destruction is bound to affect industry boundaries, particularly its degree of vertical integration. The present paper thus follows the lead of these two papers and extends their approach to the study of vertically-linked sectors. In other words, it has two main goals: first, to shed light on some of the empirical propositions put forth by Wolter and Veloso (2008) in regard to the effects of different types of innovations on vertical scope; and second to complement this analysis with a thorough investigation of profitability dynamics, with a special emphasis on profit migration across the value chain, in the context of modular innovations.

THE MODEL

Structure

Building upon the work of Jacobides (2008), we offer a systemic model of the evolution of vertical scope in an industry. It allows for a market for resources, in which firms compete for resources that are not in fully elastic supply to the industry. Firms also compete in the market for final product, facing a price-sensitive demand. They have varying degrees of capability and are able to extend their distinctive approaches to the competitive struggle. Capacity levels are fixed in the short term (a single period), but firms respond to profit signals by adjusting capacity over time. The extension of capacity by more capable firms is the central mechanism driving Darwinian selection.

In a nutshell, firms selling a final product are modeled as either specialized or integrated as-
semblers. In order to produce this good, they need a homogenous downstream resource and an intermediate good or component. This is manufactured internally by integrated firms or, in the case of specialized assemblers, it is bought from specialized suppliers who offer the component in an intermediate market. Both suppliers and the upstream segments of integrated firms use a second homogenous resource in the manufacturing of the intermediate component.

Explicitly, firms maximize profits. In the case of assemblers, suppliers and integrated firms, the objective functions are, respectively:

\[
\max_{QIB_i, RF_i} QF_i \times PF - QIB \times (PIT + TC_i) - RF_i \times PRF, 
\]

\[
\max_{QIS_i, RI_i} QIS_i \times (PIT - TC_i) - RI_i \times PRI, 
\]

\[
\max_{QIB_i, RF_i, RI_i} QF_i \times PF - RF_i \times PRF - RI_i \times PRI, 
\]

where \(QF_i\) is the quantity of the final good produced, \(PF\) is its industry-wide price, \(QIS_i\) is the quantity of the intermediate component sold by suppliers, \(QIB_i\) is the quantity of components bought by assemblers, \(PIT\) is the price of the component in the intermediate market, \(RI_i\) and \(RF_i\) are, respectively, the quantities of intermediate and final resources bought by the firms and \(PRI\) and \(PRF\) are the industry-wide prices of the intermediate and final resources used in the manufacturing processes. Finally, \(TC\) represent transaction (combined with coordination) costs as a per valorem tax that is payed by both buyers and sellers.

This optimization problem is subject to a number of constraints. The intermediate and final production functions are given by

\[
QIP_i = aI_i \times RI_i^{bl}, 
\]

\[
QF_i = aI_i \times RF_i^{bf}, 
\]

where \(aI_i \geq 0\) and \(aF_i \geq 0\) are the firm-specific upstream and downstream capabilities, and \(bl \geq 0\)
and $bF \geq 0$ the respective elasticities of production (or returns to scale).

Moreover, for integrated firms, $QF_i = QIP_i$, whereas for assemblers $QF_i = QIB_i$ and for suppliers $QIP_i = QIS_i$, implying that one component is used in each final good. Finally, $RI_i, RF_i, QIS_i, QIB_i \geq 0$.

Prices in the objective functions are endogenously determined in the model through equilibrium conditions that emerge from market clearing mechanisms. Prices for the intermediate and final resources are determined from the relationships between firm choices and two exogenous supply functions:

\[
S(PRI) = SI \times PRI^{ESI} \geq \sum_i RI_i, \quad (6)
\]

\[
S(PRF) = SF \times PRF^{ESF} \geq \sum_i RF_i, \quad (7)
\]

where $SI$ and $SF$ are constants and the exponents represent the elasticities of supply for both resources. The price for the final product is similarly obtained from

\[
D(PF) = DF \times PF^{EF} \geq \sum_i QF_i, \quad (8)
\]

Finally, the intermediate price $PIT$ arises from the market clearing equilibrium condition

\[
\sum_i QIS_i = \sum_i QIB_i \quad (9)
\]

Through these four equilibrium equations together with the Karush-Kuhn-Tucker (KKT) first-order conditions for the optimization problems and attached constraints, we arrive at a system of equations that characterize a mixed complementarity problem (MiCP). Fortunately, there are already routines implemented in GAMS for the solution of MiCP and we chose the PATH algorithm.

This concludes the static representation of the model. However, after each period, a dynamic
updating occurs. Firm-level capabilities in both segments increase due to learning-by-doing according to

\[ aI_i(t) = aI_i(1) \left( 1 + cI \times \ln(\text{cumulative } QIP_i(t)) \right), \]  

(10)

\[ aF_i(t) = aF_i(1) \left( 1 + cF \times \ln(\text{cumulative } QF_i(t)) \right), \]  

(11)

where \( cI \) and \( cF \) define the learning rates.

Moreover, firms can update their capacities (which are fixed in the short-run optimization), and decide whether they want to expand or reduce production, depending on how successful they were in the immediately previous period. Firms that reach below a minimum production threshold exit the industry in each period.

And after this happens and capabilities are updated, entry is allowed to occur. This is based on the following mechanism: a potential entrant takes two capability draws (one for \( aI \), one for \( aF \) from the same distributions that originated the initial industry heterogeneity. Given that entry entails a fixed cost \( FC \) for specialized firms and twice that for integrated ones, firms evaluate their potential profitability in the following period as either supplier, assembler or an integrated organization against these fixed costs. The higher difference between profits and fixed costs determines the entrant’s vertical boundaries.

This set-up provides a theory-informed experimental platform, from which we can explore computationally a wide variety of cases defined by alternative values of the parameters of the system.

**Modular innovation**

Despite this being a dynamic model, it does not, up to this point, address the possibility of innovations in the industry. Firms are endowed with capabilities in the beginning of the industry which pretty much determine how successful they will be. Learning only tends to exacerbate the
heterogeneity present in the beginning of the industry. There are no provisions for exogenous technological shocks, for example, to upset some of these capabilities and influence vertical scope in a different, unpredictable way. Transaction costs behave in a similar way.

Modeling the different types of innovation in Henderson and Clark’s (1990) typology entails playing with what happens with firms’ upstream and downstream productive capabilities, as well as with industry-level transaction costs. Here, we concentrate on modular innovations which seem to be a rather common feature in multi-technology, multi-component industries such as computers (Hoetker 2005), aircraft engines (Brusoni, Prencipe, and Pavitt 2001), automotive (Takeishi 2002).

Another reason to discuss modular innovations is the fact that they embody the opposing incentives discussed in Wolter and Veloso (2008). Modular innovation consists of competence-destroying technological change for upstream players in the value chain (Tushman and Anderson 1986) without a corresponding 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. 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 could be 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 sector. Therefore, assimilating it into products will include a certain degree of technical uncertainty which will accordingly raise transaction costs. 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.

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. Nonetheless, empirical studies seem to support the fact that industries move towards disintegration, that is, that specialization incentives due to the obsolescence of capabilities dominate (Wolter and Veloso 2008).

In this paper, we model this kind of innovation as an increase in the mean of the distribution from which potential entrants as suppliers draw their initial capability. In other words, at the mid-point in the simulation, firms who may choose to entry as a specialized supplier can draw substantially higher values of \( aI \).

Moreover, we limit the capability improvement to the potential specialized upstream suppliers to represent the fact that modular innovations tend to emerge from players outside the industry (Prencipe 2000). This also models the fact that integrated assemblers have “lost” their upstream capabilities due to the competence-destroying nature of the innovation and cannot keep up with the new technology. We could also have, in principle, located the locus of innovation in incumbent suppliers, but we don’t believe that this would have had any impact on our results.

We must also consider what happens when both transaction and coordination costs increase. This can be easily accomplished by an ex-post hike in the parameter \( TC \).

**Parameterization**

The model starts with 20 firms: 10 integrated, 10 suppliers and 10 specialized assemblers. Capabilities are uniformly distributed within each group between values \([0.8, 1.2]\). When the modular shock occurs at \( t = 150 \), we center the mean of the new distribution for \( aI \) at 110% of the highest value of \( aI \) in the previous period, i.e., \( t = 149 \) (note that \( aI \) increases, as does \( aF \), due to
learning-by-doing).

Our production functions display decreasing returns to scale with \( bI = bF = 0.9 \), whereas the supply of resources and final product demand are characterized, respectively, by \( \varepsilon SI = \varepsilon SF = -\varepsilon F = 1 \). And the supply and demand constants take the following values: \( SI = SF = 200 \) and \( DF = 400 \). The fixed cost of entry is set at $0.5 for specialized and 1 for integrated firms. Firms exit the industry if their profit falls below $0.5 or their production below 1 unit.

Finally, the learning coefficient for specialized firms (suppliers and assemblers) is set at 0.20, whereas \( cF \) for integrated firms at 0.15. This represents the previously discussed higher incentives that specialized firms seem to possess or, analogously, diseconomies of scope as phrased by Arora and Bokhari (2007).

It is important to note that these values constitute a first-cut, baseline case, with many parameter values being carried over from Jacobides (2008). Nonetheless, we performed a number of sensitivity tests and found the results to be robust to a number of changes.

**RESULTS**

Figure 1 represents the natural evolution of the industry’s vertical scope as a function of transaction costs, when no exogenous technological shock takes place. In the results shown below, we do not exogenously increase \( TC \) as we could have done after shock. This is discussed later as an extension.

As expected, at low levels of transaction costs, the industry tends to complete disintegration, whereas the opposite is true for high \( TC \) levels. At mid-range, the industry attains a long-run equilibrium where it is only partly disintegrated. These results track those presented by Arora and Bokhari (2007).
In contrast, Figure 2 shows what happens when a modular innovation happens at $t = 150$. As can be seen, there is a significant shift towards disintegration at all starting levels of transaction costs.

It is perhaps easier, though, to see these differences from the base case represented by Figure 1, if they are directly plotted as treatment effects. This is done in Figures 3 and 4, where the decrease in integration after the shock, both in absolute and percentage terms, are depicted.

At a first glance, Figure 3 might give the impression that industry integration levels after the technological shock decreases more when $TC$ is high, which would seem counterintuitive (although this tendency is actually reversed at the highest level of $TC$). However, this is an illusion caused by the high levels of integration to which a high $TC$ leads prior to the shock. This can be clearly seen in Figure 4 which plots the percentage decrease in integration, demonstrating that it is indeed steeper for low $TC$ values.

Now let us turn our attention to what happens with profitability in the industry. Figures 5 and 6 illustrate the behavior of aggregate profits among specialized assemblers.
The fact that a modular innovation results in significant benefits downstream is not surprising per se, although it implies that the innovation benefits are not being completely transferred to the final consumer. More capable suppliers end up using less resources, thus producing at lower costs and, consequently, lowering the price of the component that serves as an input in the assembly of the final good. The surprising effect illustrated in Figures 5 and 6 is the fact that the higher the transaction costs, the higher is the increase in the aggregate profitability of assemblers.

This counter-intuitive results seems to be explained by the entry dynamics. With the upstream shock, entry as an assembler becomes profitable again. Although, in principle, one could expect a low $TC$ to allow a greater number of new specialized assemblers to enter and displace incumbents, thus competing profits away, this is not what is happening here. We actually do find that there is more entry at intermediate to high levels of transaction costs. This is illustrated by Figures 7 and 8, where entry in then registered by three possible values: 1 for suppliers, 2 for assemblers and 3 for integrated.

When we consider the full evolutionary path it becomes clearer why. When transaction costs are lower, there are a lot of specialized assemblers entering the industry prior to the innovation.
taking place. This makes it difficult for new ones to enter after the shock (the distribution from which their capabilities are drawn does not change). However, if transaction costs are high enough to significantly restrict entry prior to the shock, but not after, then there is considerable space for entry of specialized assemblers and that is what we observe. Moreover, all this entry does not result in a lot of incumbent exits (the worse off firms are actually the integrated ones who cannot benefit from the upstream innovation), thus leading to new capable specialized assemblers to co-exist with older low-efficient peers. This is what ends up driving the profitability increases in Figures 5 and 6 (if we increase $TC$ beyond the levels show here, then we would observe the increase in profitability being lower again, i.e., following an inverted-U shape pattern).

We omit in the interest of space the plots representing the profitability of integrated firms, but suffice it to say that they follow very similar patterns as those shown in Figures 3 and 4. In other words, there are no surprises here as integrated firms suffer from their intrinsic organizational rigidity in the face of a modular innovation, and this is more pronounced, in percentage terms, at low transaction costs. Figures 9 and 10 do show, however, the combined effect downstream of changes in the profitability of both specialized and integrated assemblers.
Perhaps unsurprisingly, as $TC$ act as a tax on profitability of firms trading in the intermediate market, total profits downstream only increase, relative to the base case without innovation, when transaction cost levels are low.

The aggregate change in the profitability of suppliers is depicted in Figures 11 and 12. As we can see, the direction of change in profits for specialized suppliers after a modular innovation depends significantly on the level of transaction costs. However, a low $TC$ curiously leads to profits lower than those existing before the shock, implying that innovation is not paying off.

Again it seems that this dynamics is driven by entry and exit. Differently from what happens with specialized assemblers, a lower $TC$ value allows a larger number of suppliers to enter. Because the new distribution from which specialized suppliers draw their capability is significantly higher in the case of a modular innovation, all incumbents active prior to the shock end up exiting the industry. Consequently, the fiercer competition among innovators fostered by lower transaction costs ends up diluting away profits. Higher levels of transaction costs act as an isolating mechanism by limiting entry, therefore allowing innovative suppliers to enjoy higher profits.

But if innovative suppliers are not reaping the benefits of their prowess, at least in some cases,
it begs the question, who is then? To answer this question, we plot in Figures 13–16 treatment effects for payments made to intermediate and final resource holders as well as the endogenously determined prices of the final good and the intermediate one used in its production.

As expected, the modular innovation substantially decreases the production cost for the component and, consequently, the price it ends up fetching in the intermediate market. More importantly, though, is the fact the part of the benefits of the innovation ends up being passed onto the final consumers, as can be seen in Figure 16. This also happens at the expense of intermediate resource holders (Figure 13). The new upstream efficiency results in a lower demand for this resource, thus depressing its price (which in relative terms falls less with higher transaction costs). Perhaps a more elastic demand for the final product would be able to counter this through a market deepening effect (Jacobides 2006). At last, we can see in Figure 14 that owners of the final resource reap significant benefits from the upstream innovation through an increase in demand for the final product.

All results discussed above proved to be robust to a number of parameter changes. For ex-
ample, there was a substantial shift towards disintegration in the industry even when the modular shock was centered at only 90% of the highest upstream capability at $t = 149$ (as opposed to 110%). There were no appreciable qualitative changes in the results either when capacity constraints, i.e., limits to growth and shrinkage were introduced. Even allowing some degree of imitation or increasing or decreasing demand and supply elasticities did not seem to significantly alter our results.

Finally, as previously mentioned, we did not increase transaction costs after the shock in the simulations depicted above. However, when we did it in extensions we found that this effect wasn’t substantial. $TC$ had to double to significantly affect the results. This seems to be caused by the fact that the innovation, by increasing upstream efficiencies, ends up reducing the price of the intermediate good $PIT$ (Figure 13). And since, in our formulation, $TC$ is fixed and added or subtracted from $PIT$, the decrease in component price means that transaction costs are actually endogenously increasing as a percentage of this price.
FIGURE 13: Treatment effect on intermediate resource payments.

FIGURE 14: Treatment effect on final resource payments.

FIGURE 15: Treatment effect on price of component.

FIGURE 16: Treatment effect on price of final good.

DISCUSSION

In the quest for algebraic solutions, most work in the industrial economics or RBV tradition cannot escape the ceteris paribus; that is, it takes the supply of resources as “given”, or only focuses on the interactions between the competition and supply structures. Economic systems, however, adapt, and systemic analysis can provide insights that help us move beyond “partial equilibrium” analysis, and occasionally reverse received wisdom.
After all, insights into strategic dynamics must necessarily emerge from the industry system as it evolves. It is particularly interesting to consider how heterogeneity in firms’ capabilities interacts with transaction costs, resource usage and payments in a competitive setting, and with the structure of demand, in order to improve our understanding of the dynamics underpinning a Schumpeterian wave.

Important research in evolutionary economics (Nelson and Winter 1982), while setting the stage for a careful analysis of profitability analysis over time, does not go into much detail on the factors that determine how profits evolve in an industry under selective pressures. It lays much of the foundations for this paper, but does not focus on the precise dynamics and drivers of profit evolution as the sector shifts from its initial variety of capabilities, to the long run equilibrium. It also does not incorporate an analysis of the role of upstream and downstream conditions, or of the systemic elements.

This is particularly true of two most recent efforts to model the evolution of a sector’s vertical scope. Despite having a model focused on the specific replicability of the evolution of the computer industry, (Malerba, Nelson, Orsenigo, and Winter 2008) neglect investigating any profitability implications of their dynamics. Neither do Arora and Bokhari (2007), whose model does not address any possible destabilizing consequences of exogenous changes in capabilities or in transaction costs. This is important because such changes affect the evolution of industry not only through their impact on incumbents, but also on potential entrants. And understanding how these two effects interact is critical to a full comprehension of industry evolution under technological change.

In this paper we advance this agenda by taking a first look at these issues in the context of modular innovations. Our results show some familiar patterns, but also some surprising new ones. Modularization and outsourcing has recently been a strong characteristic of many multi-technology and multi-component industries (Fixson, Ro, and Liker 2005) and has led to an ob-
served decrease in vertical integration in these industries. Our model replicates this pattern and points to a possible reason why (Wolter and Veloso 2008), based on a review of empirical evidence, expect modular innovations to lead to disintegration, even in the presence of opposing transactional incentives. Because entry in one sector (upstream in our discussion) makes it easier for new and more capable specialized assemblers to enter at a later point in time, it reinforces the already strong forces pushing towards disintegration (Arora and Bokhari 2007, Jacobides 2008, Jacobides and Winter 2005). And making it much harder for any contemporary increase in transaction costs to have a significant effect in the other direction.

Furthermore, the model also illustrates why firms in certain industries would be willing to cede control of innovative activities to suppliers. As Helper (1991) observed in the US automotive industry, “a reduction in vertical integration at a time when both the technology and the market structure of the industry are in flux seems to contradict the predictions of organization theorists that vertical integration should rise with increasing uncertainty.” Under a modular innovation regime, specialized assemblers seem to be the ones who most benefit in terms of profitability.

Despite being represented in our formulation by incumbents at the start of the simulation and later by new entrants, in the real world we could mainly see OEMs becoming specialized assemblers by divesting parts divisions (e.g., GM’s spin-off of Delphi or Ford’s of Visteon), particularly in more mature industries. And to avoid ex post transaction and coordination costs emerging from these decisions, Helper (1991) notes that “In the new system, only a few suppliers provide each type of part, and information is interchanged extensively between buyer and supplier.”

Although this type of arrangement could be seen to create a fair amount of hazard to suppliers, our research, by moving beyond the RBV and combining elements from both the Neo-Schumpeterian and Porterian traditions, also hints at why this could work to their advantage. As we showed, too much entry results in profit erosion, even for innovative suppliers, with profits trickling down the value chain to assemblers and, ultimately, consumers. Hence, in order to profit
from the innovation, suppliers need an isolating mechanism to decrease upstream competition (Lippman and Rumelt 1982). In our model, this mechanism was provided by transaction and fixed setup costs which decreased the profitability of entry as specialized suppliers. In practice, this raising of entry barriers seems to also be accomplished by these longer-term contracts between suppliers and assemblers which, as a compromise, end up being written in the three- to five-year range.

Since our approach seems promising in explaining some counter-intuitive but nevertheless real patterns in the evolution of vertical scope and industry profitability, we are currently working on extending it in a number of ways. Our results presented here consider one initial heterogeneity distribution, but it is important to understand which effects more or less heterogenous firms at the beginning of the industry could have on its evolution. Moreover, we will extend the analysis of firm-level profitability to consider not only its instantaneous trajectory, but also its value throughout the full path in net present value terms (Jacobides, Winter, and Kassberger 2006). We are also interested in understanding potential changes to the dynamics if we allow integrated firms to sell surplus upstream production in the intermediate market or if we introduce additional exogenous intermediate submarkets as in Malerba, Nelson, Orsenigo, and Winter (2008) (e.g., when automotive suppliers sell parts to tractor OEMs). Finally, we are also exploring additional robustness tests such as allowing demand and supply to grow or varying returns to scale and the elasticities of supply and demand.

More importantly, we will investigate additional types of innovation, with a particular emphasis on architectural innovations, since these have the potential not only to considerably affect the existing market structure, but also to open additional submarkets to new firms.
REFERENCES

AFUAH, A. (2001): “Dynamic boundaries of the firm: Are firms better off being vertically inte-
grated in the face of a technological change?,” Academy of Management Journal, 44(4), 1211–
1228.

920–924.

Strategic Management Journal, 17(2), 129–150.


ARORA, A., AND F. A. S. BOKHARI (2007): “Open versus closed firms and the dynamics of


BARNEY, J. (1986): “Strategic factor markets: Expectations, luck, and business strategy,” Man-
agement Science, 32(10), 1231–1241.

coupling, and the boundaries of the firm: Why do firms know more than they make?,” Admin-
istrative Science Quarterly, 46(4), 597–621.


