Strategic Management Journal
Strat. Mgmt. J., **26**: 395–413 (2005)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/smj.460

THE CO-EVOLUTION OF CAPABILITIES AND TRANSACTION COSTS: EXPLAINING THE INSTITUTIONAL STRUCTURE OF PRODUCTION

MICHAEL G. JACOBIDES[1]* and SIDNEY G. WINTER[2]
[1] *London Business School, London, U.K.*
[2] *The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania, U.S.A.*

This paper proposes that transaction costs and capabilities are fundamentally intertwined in the determination of vertical scope, and identifies the key mechanisms of their co-evolution. Specifically, we argue that capability differences are a necessary condition for vertical specialization; and that transaction cost reductions only lead to specialization when capabilities along the value chain are heterogeneous. Furthermore, we argue that there are four evolutionary mechanisms that shape vertical scope over time. First, the selection process, itself driven by capability differences, dynamically shapes vertical scope; second, transaction costs are endogenously changed by firms that try to reshape the transactional environment to increase their profit and market share; third, changes in vertical scope affect the nature of the capability development process, i.e., the way in which firms improve their operations over time; and finally, the changes in the capability development process reshape the capability pool in the industry, changing the roster of qualified participants. These dynamics of capability and transaction cost co-evolution are illustrated through two contrasting examples: the mortgage banking industry in the United States, which shows the shift from integrated to disintegrated production; and the Swiss watch-manufacturing industry, which went from disintegration to integration. Copyright © 2005 John Wiley & Sons, Ltd.

Over the last 20 years, much progress has been made in the analysis of vertical scope, and in understanding what drives the governance structures observed in practice. A key figure in that development was Oliver Williamson (1975, 1985, 1999), who elaborated and, crucially, operationalized the concept of transaction costs, initially formulated by Coase (1937). That research has focused on a particular strand of the Coasean inquiry, examining the conditions under which firms choose to abandon markets in favor of integration. The potential for hold-ups and opportunistic behavior, this theory suggests, is the main determinant of vertical scope.

Independently, another stream of literature has come to have a defining impact on strategy as a field: the resource- and capability-based view of the firm. This approach, which has its roots in Penrose (1959), and more recently Wernerfelt (1984) and Barney (1991), emphasizes the importance of resources in guiding firm action, and the management of a firm's resource and capability portfolio as the central concern of strategy. Of late, this research has used principles suggested by evolutionary economists (Nelson and Winter, 1982) and the focus has shifted to dynamic capabilities (Teece, Pisano, and Shuen, 1997). That line of thinking would suggest that vertical scope is affected by the dynamics of resource

Keywords: capabilities; transaction costs; evolution; institutional structure of production
*Correspondence to: Michael G. Jacobides, London Business School, Sussex Place, Regent's Park, London NW1 4SA, U.K.
E-mail: mjacobides@london.edu

Received 14 July 2003
Final revision received 11 November 2004

management and the selection environment (Teece *et al*., 1994).

In the last few years, a convergence between these two theories has started, creating a more satisfactory account of what drives vertical scope. Transaction cost economists, in particular, now accept that we cannot fully understand choices of scope without assessing the resource bases of firms. Williamson himself recognizes that the transaction cost and internal firm perspectives 'deal with partly overlapping phenomena, often in complementary ways' (Williamson, 1999: 1098) and points out that a firm's history and capability endowments matter to boundary choices, a theme developed by Argyres (1996) and Argyres and Liebeskind (1999). Williamson also recommends that the traditional TCE query "'What is the best generic mode (market, hybrid, firm) to organize X" be replaced by the question "How should firm A—which has pre-existing strengths and weaknesses (core competences and disabilities)—organize X?"' (Williamson, 1999: 1103). This question has been recently pursued by Madhok (2002), who suggested that an individual firm's choice must depend not only on the characteristics of the transactional conditions, but also on its strategic objectives, the attributes of its own capabilities, and the governance context it has created. There is by now substantial empirical support for the proposition that considerations of transaction governance trade off against capability considerations when firms choose component suppliers (Walker and Weber, 1984; Poppo and Zenger, 1998; Schilling and Steensma, 2001; Afuah, 2001; Jacobides and Hitt, 2004; Hoetker, 2005). These contributions consider the complementary roles of transactional and capability considerations in the micro-analysis of firm decisions.

Recent progress notwithstanding, important gaps remain in our understanding of how transaction costs and capabilities combine to determine vertical scope. The preponderance of the existing literature is focused at the individual firm level, and on relatively short time spans. Valuable as this micro-analytic approach has been in understanding transaction governance at the firm level, important phenomena exist that are simply beyond its reach—as subsequent examples and cited references in this paper amply illustrate.

To address these phenomena, we propose first of all that the analysis at the individual firm level must be complemented by a *systemic* view that takes the full roster of participants into account. Indeed, this stance is required for logical completeness, when analyzing vertical structure at the industry level. Further, important causal mechanisms operate over substantial time periods; to elucidate these requires a *dynamic, co-evolutionary* view of how capabilities and transaction costs change and interact. Finally, the systemic and co-evolutionary perspectives highlight the importance of *endogenous change in transaction costs*.

In other words, to understand a firm's vertical scope we have to understand the mechanics of how transactional and capability conditions determine which of the possible choices on the menu of available alternatives will be chosen by an individual firm at a given time; this is what extant research on micro-determinants of vertical scope has largely done to date. But we also have to understand *how this menu of available choices* is being formed in the first place, both in the short run and in the long. This means understanding how that menu is shaped by the generative process at the level of the industry. To accomplish this, it is crucially important to take a systemic approach to the evolutionary dynamics of scope determination, and to identify the specific mechanisms by which transaction costs and capabilities co-evolve.

Our paper's specific contribution, then, is to provide a theoretical framework that explains how capabilities co-evolve with transaction costs to set the menu of available choices that firms face in an industry. This is done through the identification of the specific evolutionary mechanisms that determine vertical scope over time. The paper also highlights other important considerations that have been underemphasized to date. First, it points out that transaction costs themselves are not fully exogenous; their magnitude depends on the conscious actions undertaken by firms. If industry participants stand to benefit from transaction cost reduction, they will actively try to reduce them. Furthermore, changes in vertical scope involve a significant feedback loop in the capability development process; changing scope changes how capabilities evolve, and this also changes the roster of 'qualified' industry participants. Vertical disintegration in particular opens up an industry to entirely different participants and new pools of capability. To illustrate, vertical disintegration in sectors such as financial services have led outside, specialized firms such as IBM and EDS to become

significant players, and have also changed the identity of financial institutions (how firms conceive of themselves; what their points of reference and comparison are) and, as a result, the way capabilities develop.

Thus the paper reaches beyond the question of vertical scope to deepen our understanding of industry evolution itself. It also sheds light on the process of capability development, providing a new angle on important phenomena such as outsourcing and industry convergence. Of course, it is only a stepping-stone for further research. Ultimately, we believe that such research can yield a satisfying response to Ronald Coase's (1992) call for understanding of the 'institutional structure of production' (ISP)—a *joint* analysis of how transaction costs and capabilities shape the division of labor across the vertical divide and between different firms.

We begin by examining how transaction costs and capabilities interact to determine vertical scope in the short term; this illustrates our industry-level approach and contrasts it with the micro-analysis of individual firm decisions. We next move into the paper's core contribution, and consider the four key evolutionary mechanisms. First, we look at how the selection process drives scope, amplifying capability differences in the industry; second, we consider how capability differences set the incentive context for endogenous reduction of transaction costs over time.[1] We then complete the loop, discussing how vertical scope affects the process of capability development, and how changing scope reshapes the identity of existing industry participants, as well as the roster of potential entrants. To illustrate these dynamics we examine how industries shift from integration to disintegration, and vice versa, using examples from U.S. mortgage banking and Swiss watch manufacturing, respectively. A concluding discussion considers how our research informs institutional economics and the resource/capabilities-based view, and suggests paths for empirical and theoretical research in strategic management.

[1] We use the term 'selection process' to refer to the rising relative importance (share of industry output) of successful firms, due to differential growth and survival as a result of competition, as well as the additional force raising the relative importance of successful practices (including a firm's vertical structure), imitation.

THE SHORT TERM: COMPARATIVE ADVANTAGE IN THE INSTITUTIONAL STRUCTURE OF PRODUCTION[2]

We start our analysis with the short-term, relatively static view, focusing in particular on how the effect of transaction cost reductions is mediated by the capability distribution in the industry. At the heart of our argument is the premise that productive capabilities are generally quite heterogeneous across firms and across stages in the production process. Here, the term 'productive capabilities' embraces the underlying determinants of the efficiency with which firms manage to carry out their productive activities. Productive capabilities rest on the firm's general and specific knowledge of how to do things (Richardson, 1972; Teece *et al.*, 1997), and also involve specific investments in equipment, training, and retention of key personnel, etc., required to put that knowledge to work. They are what Winter (2003) describes as 'zero-level' capabilities, i.e., the determinants of a firm's efficiency or effectiveness in engaging in its current business activities.

Heterogeneity in productive capabilities is typical because the capability to carry out a complex activity is typically developed in an organization through a long, path-dependent learning process (Winter, 1988), in which there is abundant opportunity for various contingencies to shape the way of doing things that ultimately emerges (cf. Levinthal, 1997). Particularly important contingencies are the different 'bets' that actors make, in the face of great uncertainty, as to what will prove to be the most effective way of doing things. Even if the correct recipes become clear, their diffusion is limited by complexity, often due to interactions among activities (Porter, 1996; Rivkin, 2001; Siggelkow, 2001). The force of imitation is also weakened by the path dependence associated with the fact that investments in capabilities are so often durable and/or sunk. Correction of past mistakes is not necessarily economic at the relevant margin. Thus, even in environments where primary resources are quite homogeneous, different organizations are likely to display significantly different ways of accomplishing approximately the same thing, displaying different efficiencies as a result

[2] This discussion draws extensively on a formal model of how capabilities and transaction costs interact to shape vertical scope by Jacobides (2004).

(Lieberman and Dhawan, 2001; Collis and Noda, 2001). Firms may thus differ either in terms of cost efficiency, or of 'effectiveness' (level of quality for any given set of inputs, as discussed by Barney, 1991, and Peteraf and Barney, 2003.)

The second key part of our analysis is the deceptively simple observation that behind the façade of 'the market' lies another firm. 'The market' does not produce anything; it is the thin interface through which the product or service of another firm is purchased. This means that there is an 'adding-up constraint' on the vertical structure of the industry as a whole: every purchase in the intermediate market is also a sale, and total purchases therefore equal total sales. This constraint is lost from view when vertical scope is addressed only at the firm level and the choice framed as 'make' vs. 'buy' or 'firm' vs. 'market.'

This simple observation has important implications. Firms, in deciding whether they want to be integrated or not (in a given historical context), compare their own abilities with those of other firms—as signaled by the price and quality terms on which those others are prepared to deal. This is a fundamental point from which most of the subsequent economic and TCE analysis shies away for analytical convenience; TCE focuses on the conditions of exchange, to the neglect of the conditions of production (Demsetz, 1988; Coase, 1988, 1992; Langlois and Foss, 1999; but see Riordan and Williamson, 1985). The implication of this point is that it is necessary to look at the *distribution of productive capabilities*— the source of any potential 'gains from trade' through the market (cf. Teece, 1980; Jacobides and Hitt, 2004)—to understand when firms are integrated and when they are not.

Simply put, for market transactions to occur, *both* parties must find specialization advantageous, sufficiently so to outweigh any costs of trading; and this implies particular properties in the distribution of productive capabilities in the industry. To explicate this further, consider an industry with two segments: upstream and downstream. For intermediate market transactions to make sense, some firms must be good upstream *and* not have good downstream capabilities, or at least be capacity-constrained downstream, so that trade becomes beneficial for them. If the efficient firms upstream were as good downstream as the other downstream firms, they would not need or want to specialize, regardless of the potential transaction costs (TC).[3] Thus, when the level of analysis is shifted from the individual firm to the population of firms in an industry, it becomes clear that vertical specialization must be in part a function of heterogeneity in productive capabilities along the value chain.

In particular, if productive capabilities, at the industry level, are symmetrically distributed in the upstream and downstream segment, and if there are no capacity constraints or bottlenecks, then there will be no benefit from intermediate trade (Jacobides, 2004). If, however, capabilities are but weakly correlated across stages—the scatter plot of upstream efficiency vs. downstream efficiency is loose—specialization will occur if transaction costs permit. It seems reasonable to suppose that such weak (or even negative) correlations will be the rule when productive capabilities in different parts of the value chain build on different knowledge bases. In other words, the more dissimilar the segments, and the less strength from one segment can translate into strength in another, the smaller the likelihood that a single firm will be equally good upstream and downstream.

Transaction costs do play a role—but their role is akin to that of a sales tax or a tariff levied in international trade.[4] If productive capability

[3] Only if we assume away capability/productivity differences between firms can we directly compare the workings of 'the firm' and 'the market' as coordination/governance mechanisms. If capability heterogeneity does exist, any comparison between 'the firm' and 'the market' is simultaneously a comparison of the properties of the governance mechanism and of the particular productive capabilities of the potential transacting partners vs. those of the integrated firm. Issues of incentive intensity are also involved here, but their importance may be exaggerated if capability heterogeneity is ignored.

[4] There are multiple sources and aspects of transaction costs. Coase (1937, 1988) emphasized the 'frictional' costs, such as those of identifying a potential supplier, negotiating, drafting a contract and monitoring it, etc. Williamson (1975, 1985) transformed the subject by shifting attention to the costs of transactional hazards and of governance arrangements to limit such hazards. His focus is on the tendency of transactions to run into difficulty for reasons associated with bounded rationality and opportunism, whereas the frictional costs are present even when things go well—a feature of economic reality more like physical friction, being independent of human calculation and motivation. Alternatively, transaction costs may arise from difficulties in measuring and monitoring performance (Alchian and Demsetz, 1972; Barzel 1982), or the inability to specify the goods and services needed (Jacobides and Croson, 2001). While the distinctions among these types of costs are clearly important for some purposes, such as the micro-analysis of governance arrangements, they are all quite similar when viewed in a systemic perspective. They all represent burdens or obstacles to market transactions, and they are all potentially subject to reduction, at least in the long run, through some combination of managerial

asymmetries are high, such 'taxes' might not be sufficient to curb vertical specialization. A firm that contemplates using the market may be willing (even if grudgingly) to pay these high 'taxes,' be they frictional costs or costs of the expected losses from hold-up, renegotiation, and such, if the gains from trade are high enough to compensate for these losses. The relevant calculus is the comparison of the potential gains from trade (due to heterogeneous productive capabilities) as measured against the level of TC. Conversely, even if the level of the TC 'tax' is low, if the gains from trade are *even smaller* (because capabilities are symmetrically distributed) then there will be no reason for vertical specialization to emerge. Similarly, the increase of TC may or may not lead to greater integration: if the increase is not sufficient to outweigh the gains from specialization, no further integration will take place.[5]

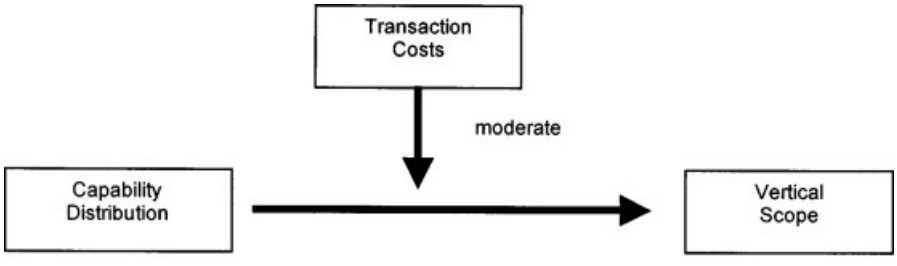
A transaction cost decrease lubricates the workings of the market, in that it allows firms to capitalize on their capabilities and relative strengths. If firms are different, a transaction cost reduction will allow for significant specialization, as each firm will focus on its area of strength. If all firms are alike, or differ in equal proportion across stages, the same reduction will not do much (if anything) to promote specialization and disintegration. The most efficient firm cannot find a superior source to trade with and organizes internally, and then this logic applies to the second best firm, and so on down the line; no latent gains from trade exist to change vertical scope. A way to visualize this is to consider a lake, and think of the water of the lake as being the transaction costs. As transaction costs recede, *if* the lake bottom is rugged—that is, *if* the capabilities are asymmetrically distributed—islands of specialization will increasingly appear. Yet if the lake bottom is smooth, if there is no heterogeneity to uncover underneath the surface, the lake surface also remains smooth—integration will remain (for a formal analysis, see Jacobides, 2004). Figure 1 illustrates.

ingenuity and appropriate investments. In their systemic consequences, these costs are all akin to a 'tax'.

[5] Could a merger of two firms, each efficient at one stage of the production process, be a way to ensure that gains from specialization can still be had within the boundaries of one integrated entity? Perhaps under some circumstances it could. But it is likely that such a vertical merger, presumably motivated by the need to reduce transaction costs, will involve interventions that might erode the superiority in each of the merged divisions, because of the often unwitting and inappropriate need for intra-firm homogeneity in cultures and management styles, pay-equity norms, etc.

HOW DOES THE ISP EVOLVE? UNDERSTANDING THE DYNAMICS OF SCOPE AND CAPABILITY

The previous discussion argues that at the industry level of analysis it is quite clear that the vertical scope of individual firms depends on capability differences. This, however, is a relatively static picture. By considering vertical scope in the

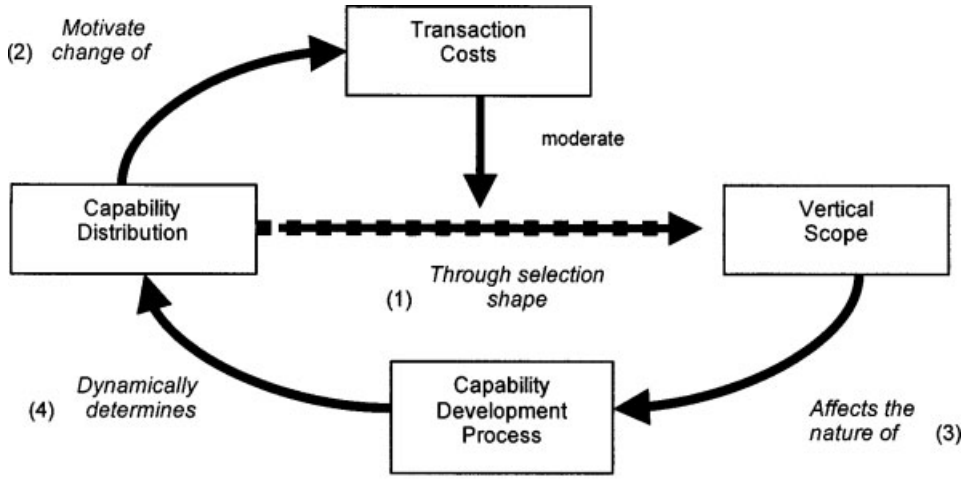


Short-term determination of vertical scope:
Transaction Costs *operate* on the underlying capability distribution

If capabilities are dissimilar (weakly correlated at most) along the value chain
Then latent gains from trade across firm's boundaries exist
Then a reduction of Transaction Costs ***will*** lead to substantial disintegration

If capabilities are similar (highly correlated) along the value chain
Then there are no latent gains from trade across firm boundaries
Then a reduction of Transaction Costs ***will not*** lead to substantial disintegration

Figure 1. The short-term framework for explaining the institutional structure of production



Dynamics: Evolutionary mechanisms , explained

Mechanism (1)	**Selection amplifies the impact of capabilities on scope** (through competition and imitation)
Mechanism (2)	**Latent gains motivate endogenous reductions of TC**
Mechanism (3)	**Changes in scope affect the capability development process**
Mechanism (4)	**Capability development affects the roster of participants.** (including entrants from new sectors)

Figure 2. How capabilities, transaction costs and scope co-evolve in an industry

historical context of industry evolution, we gain further additional insights by identifying the *evolutionary mechanisms* that shape it. In particular, we argue that (1) selection (through competition and imitation) affects vertical scope at the firm and industry level, reinforcing the static results shown above; (2) transaction costs themselves are partly endogenous, and in particular they respond to productive capability differences: firms that stand to win from, e.g., lower transaction costs make the investments, including lobbying and institutional investments, to shape the transactional environment to their advantage. Also, (3) the capability development process itself changes as a consequence of changing scope. Vertical disintegration in particular often has dramatic implications for the nature of industry participants, the identity of firms within it, as well as the capability development process. Completing the loop, (4) the capability pool in the industry changes as new participants appear and new knowledge bases become relevant. This section, then, adds the dynamic relationships among the key variables, identifies the role of the capability development process on the nature of the industry and explains the evolutionary logic of industry's changes in scope.

Figure 2 provides a visual illustration of our argument, and expands the static picture in Figure 1. Table 1 also summarizes these four mechanisms, which we now consider in turn.

Selection amplifies the impact of capabilities on scope

The first evolutionary mechanism that shapes the dynamics of vertical scope is the selection process, be it through competition (that changes the shares of different firms in overall production), or imitation which determines the vertical scope in an industry.

The literature on capabilities, and recent empirical work (Lieberman and Dhawan, 2001; Collis and Noda, 2001), indicates that the typical situation in an industry, especially early in a technology or industry's life cycle, is that firms have heterogeneous productive capabilities. There are two aspects of this heterogeneity we focus on: diversity

Table 1. The evolutionary mechanisms of the institutional structure of production: how capabilities and transaction costs co-evolve

Mechanism	Causal process	Result
1. Selection amplifies the impact of capabilities on scope	Differential profitability arising from heterogeneous capabilities shapes firm growth and thus promotes particular vertical structures; imitation amplifies these effects	Selection increases scales of effective firms/units, prompts imitation, and reduces heterogeneity in all things, including vertical structure; efficient firms 'impose' their desired scope
2. Latent gains motivate endogenous reductions in transaction costs	Capability differences between firms up- vs. downstream provide incentives to change TC (generally reduce TC, turning latent into realized gains from trade)	Endogenous changes of TC; usually, reduction of transaction costs and initiation of contracting, *provided* capability differences along the value chain provide incentives to do so
3. Changes in scope affect the capability development process	Changes in scope affect the capability development process	Changes in the capability development process
4. Capability development affects the roster of participants	As a result of new vertical scope, capability pool in the industry (which comes from existing or new types of players) changes, leading to new capability distribution	New mix of participants and hence new capability pool in the industry Changes in industry definitions and competitor/capability profile; industry convergence

of unit production costs and diversity in the nature of the vertical division of labor. The mechanism is then as follows:

First, the firms with the lower overall costs (or better product quality) are more profitable and are likely to grow—the more so because of attractive opportunities in the marketplace, the threat from ambitious rivals and the desire to seize advantageous positions before others do. Specifically, in every period, the economic selection process of profit-driven growth implies that the scale of the most successful firms, or the successful vertical units of firms, will increase—in many cases dramatically. To the extent that specialized firms have superior productive capabilities, selection will soon push for greater specialization; conversely, if firms that are superior are also integrated (by happenstance or for a good reason), selection will beget more integration.

A firm that finds itself with sharply asymmetric capabilities is likely to commit itself to the segment in which it is stronger. This is particularly obvious when the weaker segment is actually unprofitable. But it is also common for firms to withdraw from segments in which they have managed to be at least marginally profitable. Why does this happen? Why would a firm want to abandon what might be a profitable operation, sourcing it outside? The answer is that expansion and profitable growth will have to focus on the areas where a firm has a comparative advantage when compared to the rest of the industry. Even if a firm can make a profit in both segments, it may choose to drop one of them to focus managerial attention and other resources on building its core strength. Concern with the stock market's emphasis on rates of return, as opposed to total return, also leads firms to avoid 'diluting' the profitable performance of one segment with marginally satisfactory results from the other.[6] Yet these developments can only transpire when the industry context supports it: the balance of competencies in the rest of the industry must imply a match between what the focal firm wants to outsource and what the rest of the industry wants, on net balance, to supply.[7]

Provided that transaction costs permit an intermediate market to appear, selection also affects

[6] From a strategic point of view, a firm needs also to consider its future vulnerability to market power in the abandoned segment. The possibility of countering such exploitation via re-integration depends on the ease of acquiring the capabilities (see Karim and Mitchell, 2000)—which depends, in turn, on the underlying knowledge conditions and the evolution of technology (cf. Prahalad and Hamel, 1990).

[7] For such specialization to occur, however, one more condition must be met: the selection environment must be stringent enough to prompt the abandonment of the competitively weaker market/value chain position. This proviso may explain why vertical reorganization is often prompted by an economic crisis (North, 1986; Silver, 1984). Indeed, a period when demand is sufficiently strong to motivate full utilization of all capacity, upstream and down, is a period when the competitive impact of asymmetric capability is muffled.

the overall degree of vertical integration. To the extent that successful firms tend to replicate those arrangements as they extend their capacity, selection forces produce a decrease in the effective variety of vertical arrangements across the industry. As a consequence, success and failure become more apparent as time passes. This increases the power of a related mechanism that is also operating to decrease heterogeneity—imitation of the more successful firms by others (Peteraf and Shanley, 1997), including the particularly important case of imitation in new firms founded by managers who began their careers in those more successful firms and know a good deal about what is done there (Klepper, 2002). As Shanley and Peteraf (2004: 12, 15) note:

> Through trial and error interactions, firms learn about the resources and capabilities of other actors ... about the degree to which other firms complement their own activities and whether there are synergistic effects from joint activity. Firms also learn about the behavior of other firms as exchange partners and as competitors for jointly created value ... [Also,] through their contacts in professional organizations, managers may identify the types of vertical linkages and groupings that are commonly employed and how they appear to work for the firms involved. Managers can also go beyond their immediate set of industry contacts to focus on larger and highly visible firms as exemplars, whose activities are publicized through the business media. As members of other firms' boards, managers can examine potential linkages ... Informational intermediaries, such as consulting firms, [further] spread information.

Furthermore, even in the absence of hard evidence on the efficiency of conducting business in any particular way, pressures of institutional isomorphism tend to develop and soon restrict the menus of intra-firm organization of activity (Meyer and Rowan, 1977; DiMaggio and Powell, 1983). For all these reasons, vertical structures within the industry become increasingly homogeneous, as selection reinforces the impacts of productive capability differences on scope.

Latent gains motivate endogenous reductions in transaction costs

Perhaps the most important insight one gets by observing the actions of firms in an industry over time is their conscious and considerable effort to reshape their institutional environment. In other words, rather than considering transaction costs as a given, firms actively try to manipulate and shape the transactional environment to their advantage. Although the transaction cost context is fixed in the short run, it is subject to change in the medium term as firms make incremental adaptations and explore new directions in the quest for profit.

This observation sets our framework apart from the firm-level (and almost always short-run) analyses of transaction cost economics or incomplete contract theory (cf. Hart, 1995). In these research streams, the problem of vertical scope is basically one of choice from a menu of governance alternatives. We argue that the choice menu from which a firm picks is determined by the conditions of the industry as a whole, at each point of time, so that even an optimal choice is constrained by an institutional environment that is *fixed* in the short run. That environment is determined by such slow-evolving things as prevailing contracting norms, firm reputations and transactions technology, by the ongoing process of learning to contract (Mayer and Argyres, 2004), as well as the existence of technological or transactional 'interfaces' (Baldwin and Clark, 2003) that can support market exchange to a greater or lesser extent.

Yet in the intermediate and longer term there is significant firm agency in shaping the TC context of the entire industry (Jacobides, 2005). For instance, an individual firm may come up with a particular way to organize its production by, say, creating a new way to measure and assess an intermediate good, or a new way to coordinate the upstream and downstream operations. To the extent that this reduces TC, it enables specialization, thus reshaping its institutional context for the next period (Argyres and Liebeskind, 1999; Madhok, 2002; Argyres and Mayer, 2004; Mayer and Argyres, 2004). We further argue that, by and large, the benefits of that action also affect the economy as a whole—at least as far as such transactional solutions proliferate through imitation. This dynamic element, whenever it is at play, generally tends to push for a decrease of TC. Such investments in the creation of market interfaces are often imperfectly appropriable. Firms make them when they anticipate that their private benefits will exceed their costs, but even firms that had not been concerned with funding the creation of a new intermediate market might, once the fixed costs are sunk by their competitors, adopt their ways of transacting and thus fuel further specialization.

Such effects arise, for example, in the context of a 'standards war,' like IBM vs. Apple in personal computers, or VHS vs. Betamax in VCRs (Cusumano, Mylonadis, and Rosenbloom, 1992).

Changes in scope affect the capability development process

So far, we have argued that vertical scope changes as a result of the forces of selection, and also as a result of the conscious efforts to change the transactional environment. We now discuss how changes in scope produce changes in the process of capability development and knowledge accumulation.

Our first observation is that the scope of a firm shapes its incentives to invest in capability improvement. Larger scale, conferred by previous success in the market place, motivates investments to further enhance productivity—the more so when the costs of such investments are relatively independent of the scale of application. For a firm with asymmetric capabilities, these incentives operate more strongly in its areas of existing strength, creating a self-reinforcing dynamic of specialization.

There are subtler effects that may be at least equally important. An integrated bank, for instance, views itself as a 'bank;' its points of reference, its competition, and its processes draw on the banking world. Issues of organizational identity and framing (see Albert and Whetten, 1985; Whetten and Mackey, 2002), which significantly affect the way in which capabilities are formed and managerial actions are taken (Tripsas and Gavetti, 2000), are related to the scope of the firm, and to its self-perception. On the other hand, a focused financial intermediary draws from a more specialized knowledge base; it has different points of reference, and different means of searching its competitive environment as it improves its business practices. For instance, a unit that focuses exclusively on acquiring customers may improve customer acquisition practices more effectively, not only because of greater accountability and clarity of objectives of a focused unit (Lawrence and Lorsch, 1967), but also because independent divisions can consider solutions that depart more radically from existing practices. The scope of references of a specialized firm is deeper in its own domain than even a specialized department for the firm as a whole. A customer acquisition vertical specialist, for instance, may consider techniques not only in its own field, but also in other sectors, and it will try to integrate the experience from a more narrowly, yet also more deeply, calibrated set of related projects and similar cases. Overall, the capacity to absorb new productive knowledge arising externally (Cohen and Levinthal, 1990) depends on scope.

While the process of capability development depends on how integrated a firm is, it depends even more on how integrated an industry is. That is, the organizational identity of firms that compete in a vertically disintegrated industry, and the concomitant processes of capability development and knowledge accumulation, are drastically different from that of an integrated industry. In a wide range of sectors that have recently disintegrated, ranging from financial services, to health care, to utilities, to automobile production, vertical specialization has transformed both the identity of firms and the capability development process.

In financial services, for instance, previously integrated sectors have unbundled, partly as a result of deregulation, partly as a result of changes in information technology (Evans and Wurster, 1999), and partly as the result of selection and endogenous TC reduction described above. In these sectors, the change from an integrated to a disintegrated structure drastically changed the nature of the industry, and the structure of the capabilities that firms needed to compete; the 'special' status of being 'a bank' or 'an insurance company' gave way to the realization that banks and insurers are collections of information processors, data handlers, risk pricers, etc, and as such that specialization could help them by drawing from different sectors (Crane and Bodie, 1996; Hagel and Singer, 1999). This shift in the self-perception of banks has led to changes in the capability development processes both for the integrated and the disintegrated entities in the industry. While integrated banks still do cling more to the mindset that 'banks are different,' the emergence of vertical specialization has changed the way they look at themselves and the related capability development process (Albert and Whetten, 1985; Whetten and Mackey, 2002). Thus, changes in vertical scope at the firm and especially at the industry level can and do affect the nature of the knowledge accumulation and capability development process; indeed, such changes may be among the most pervasive and least studied drivers of capabilities over time.

To the extent that the specialized production leads to faster knowledge accumulation, vertically specialized firms may be able to improve more quickly than the integrated ones. So even if the initial capability endowment favors some larger, integrated entities, the knowledge accumulation process may ultimately lead to gains from specialization, and thence to gains from trade. This means that the emergence of gains from specialization may depend (a) on the extent to which the knowledge bases of the different segment are divergent; and (b) on the extent to which focusing on one of the two improves the ability of any given firm to enhance its productive capability (Argyres, 1996; Christensen, Verlinden, and Westerman, 2002).

Yet while vertical specialization significantly affects the knowledge development process, and often allows specialized firms to develop superior capabilities as they can draw on a more extensive knowledge base, it can also be detrimental when it inhibits systemic innovations (Chesbrough and Teece, 1996). That is, excessive and sustained specialization may create 'silos' that inhibit systemic business improvement. Some new capabilities, especially those resting in knowledge bases or technologies new to an industry, are beyond the reach of existing specialized firms, and are necessarily introduced by new, integrated players (Raynor and Christensen, 2002). More generally, whether integration or disintegration provides the stronger basis for capability improvements is an empirical issue; the answer tends to vary with industry life-stage. It depends on the nature of new capabilities, and whether they span multiple stages of the production process or not. As a general rule, new capabilities that draw on new, integrated knowledge bases, or which depend on a systemic reorganization of production, shift the industry towards reintegration. We discuss this further in the following section, and illustrate it with the example of Swiss watch manufacturing (for a similar view see Fine (1998).

Capability development affects the roster of participants

In addition to changes in the balance among existing industry participants, the changes in the knowledge development process have yet another, and potentially important impact: changes in vertical scope not only shape the capability distribution of existing players, but also determine the roster of potential entrants/industry participants, potentially altering (perhaps dramatically) the pool of players and the relevant capabilities in the industry. For instance, when specialization breaks an industry into pieces, some pieces may correspond closely to activities in other sectors. Vertical specialization in banking and the creation of specialist data-handling units, for example, highlighted the promise of drawing from data management practices outside of banking. It also allowed companies previously not associated with banking, such as EDS and IBM to export their capabilities to the newly specialized sector. Thus the nature and boundaries of the industry, as viewed at a particular time, may be called into question by the adjustment of intra-firm knowledge accumulation processes. The more a specialized unit draws on knowledge sources outside the industry, the more it draws attention to its particular function as one that is performed well outside the industry. The result may be, for example, that a 'model' functional specialist from outside the industry sees an attractive niche within the industry and tries to get the business—and if successful, this may ultimately become a 'toehold' entry.

This change mechanism is highly visible in the increasingly ubiquitous (and controversial) practice of outsourcing and offshoring vertical segments (Doig *et al*., 2001; Adler, 2003). The growth of this practice has been predicated on the idea that information processing, data handling, call center and customer relation management are generic between different industries. While each sector (e.g., airlines, PC manufacturers, banks, insurances, travel agents) formerly viewed their information processing as a very 'special' part of their operation, companies such as EDS and Infosys saw that these functions could be broken off, disintegrated from the rest of their respective value chains. Furthermore, once these service vendors saw data handling, IT or human resource management as relatively generic applications of their own capabilities (Adler, 2003), they proceeded to develop superior capabilities in these focused functions, and leveraged the knowledge and capabilities gained from one sector (e.g., call centers for banks) to another (e.g., call centers for airlines). As a result, they started participating in several industries, and helped reduce the transaction costs through interface standardization and other innovations; that was the price they needed to pay to become industry participants in new segments.

This changed the nature of the industries they colonized, amounting in fact to the creation of a new 'industry' that cut across the vertical boundaries of a number of existing sectors in the economy.

The impulse to reintegrate

The mechanisms described above, and the interactions among them, tend to produce a secular drift in the direction of a less integrated industry structure. This tendency is illustrated in the histories of many sectors, as we and others before us have remarked.

Yet this is not the whole story. There is a contrary push from disintegration back into integration, which also comes about as the result of a self-reinforcing evolutionary mechanism. Specifically, we argue that the cycle pushing toward specialization gets reversed when new and superior capabilities arise from knowledge bases that are misaligned with the existing vertical structure of the industry. This sets in motion a process that may eventually make vertical integration typical, endogenously increasing TC along the way.

This process has been alluded to by Silver (1984), Langlois and Robertson (1995), Fine (1998), and even Williamson (1983). In particular, Williamson (1983: 107ff.) argues that integration comes in whenever asset-specific investments are called for, which tend to make integration an economizing solution. This, we suggest, happens precisely when firms confront new choice menus that involve the appearance of technologies relying on new knowledge bases—often different from any of the constituent parts of the previous ones (e.g., micro-electronics as opposed to mechanical technologies for cash registers—see Rosenbloom, 2000). The new and superior knowledge base may initially favor integrated players; and this means that the capabilities of such players are superior to those of the existing industry participants. The ensuing selection process reduces overall specialization, as integrated players out-compete the existing co-specialized ecosystem.

Furthermore, the organizational requirements that emerge as new technology is brought into practice tend to be reflected in an increase in the relevant transaction costs (which thus stand in a bidirectional causal relation to integration itself). In the absence of an accepted interface with stable capabilities on both sides of it, the incentive to tame such transaction costs is limited because there are initially no gains from trade to be realized by doing so. This is true in the first place because the new downstream producers may have no effective and profitable way to trigger the appearance of a supplying industry to meet their needs. The question of whether a required intermediate product or service is available on the market may have a sharp 'no' answer, or it may have a more or less ambiguous 'yes' answer. The suitability of existing products for the new role is often a matter of degree—some adaptation or improvement may be required. This is the second reason for integration to surge—the difficulty of getting appropriate responsiveness from existing external suppliers, and also the need to identify and learn to manage an efficient interface between the stages. In the wake of a radical technological discontinuity, as in the early stages of an industry, the management task is itself changing along with design details upstream and downstream, making the coordination task particularly challenging. Langlois (1992: 116) summarizes: 'Ultimately, the costs that lead to vertical integration are the (dynamic) transaction costs of persuading, negotiating with and coordinating among, and teaching outside suppliers in the face of economic change or innovation.'[8]

It may also happen that new technologies and related management capabilities for a focal sector are imported from an existing integrated sector. The industry will be transformed on the basis of new productive structures—new architectures (Henderson and Clark, 1990) which require new ways of organizing (Chesbrough and Teece, 1996). This argument is also consistent with the implicit assumption by Christensen, Raynor, and Verlinden (2001) and Christensen *et al.* (2002), and the

[8] An example of this concerns the meat packing innovations of Gustavus Swift, a tale recounted by Chandler (1977), Porter and Livesay (1971), and Langlois (2003), among others. Chandler's seminal analysis of the emergence of integrated corporations (Chandler, 1962, 1977) could be readily recast in this same light. The new ways to organize production, spurred by changes in transportation and communication technologies, originally favored integrated firms because they led to a new set of capabilities, cutting across different stages of the production process, which were superior to the pre-existing specialized capabilities. The new integrated corporations did not primarily respond to an increase in transaction costs; rather, they involved a fundamental transformation of both the process of organizing production, and the types of knowledge or expertise such production relied upon. The capability itself comprised superior and firm-specific ways of managing interdependencies throughout the value chain, while seizing the benefits of newly found economies of scale (cf. Lamoreaux and Raff, 1995). This reading of the history is also consistent with Lamoreaux, Raff, and Temin (2003).

more explicit argument by Raynor and Christensen (2002) that vertical integration may lead to superior innovative abilities (read: superior ability to improve capabilities). As Raynor and Christensen (2002: 4) argue for the case of telecommunications, 'In particular, the ability to develop and deploy new technology services depends on ... the control a firm exercises over the value chain ... interdependent technological architectures are best developed by firms with integrated value chains.' The shift of several manufacturing or component sectors towards 'total solution provision' (Foote *et al*., 2001), by redefining the scope and the way the firm and its identity (and capabilities) are defined, is another case in point; Cacciatori and Jacobides (2004) provide a detailed illustration of this tendency in their study of how the U.K. construction sector shifted from co-specialized vertical specialists to a set of solution providers. We also describe one such example, the organization of production in the watch industry, which shifted from disintegration to integration, later in the paper.

Summary and implications

To summarize, we propose that the evolution of the institutional structure of production in general, and vertical scope in particular, is driven by the selection mechanisms that operate within an industry; that these selection mechanisms will tend to reduce variety in the ways labor is organized within firms (though not necessarily yielding maximally efficient ways of organizing production). Once a clear division of labor emerges within firms, across-firm comparisons of segment performance may begin, and these may even lead to the creation of a new intermediate market if integration was the only choice earlier on. The intensity of the TC reduction effort depends on the latent gains from so doing; and the latent gains from trade depend, in turn, on the distribution of capabilities along the value chain. Finally, the TC context further feeds back into the process of capability development: specialization opens up the industry to a new knowledge base, and possibly new industry participants. Yet when this cycle of specialization exhausts the benefits it can offer, or whenever a new knowledge base comes about that relies on a more integrated structure, integrated firms may displace the specialized ones, and the inverse process takes place: TC endogenously increase within the newly integrated structures; the more efficient firms displace the older, disintegrated ones, and this process is reinforced as the capabilities of integrated firms continue to increase at a quicker pace; this further changes the capability development process. So industries may be characterized by phases of disintegration, driven by the four mechanisms delineated above, succeeded by reintegration, which is driven by product or process innovations that pose novel problems of vertical coordination or rely on new, often integrated capabilities.

Thus, the four mechanisms we identify and illustrate in Table 1 and Figure 2 explain the mechanics and causal drivers of this co-evolutionary process. These mechanisms shape the menu of transactional choices and the relative capability positions each firm faces in an industry.

CO-EVOLUTION, ILLUSTRATED

The process of transaction cost and capability co-evolution, then, can have two distinct stages, or aspects: one which moves from integration to disintegration, and one which moves from disintegration back to integration. This section provides a historically grounded illustration of these two parts of the co-evolutionary cycle, using mortgage banking in the United States and watch manufacturing in Switzerland in the 1980s as the focal examples.

From integration to disintegration: mortgage banking in the United States, 1981–1989[9]

In the early years following World War II, mortgage financing was a very localized and vertically integrated business. Integrated firms, especially savings and loan associations, used to produce (i.e., originate), hold, and service mortgage loans. In the late 1970s, this model of integrated housing finance started giving way to a vertically specialized one. The process as a whole included a number of discrete episodes that illustrate the mechanisms and causal paths discussed above. We focus on the role of mortgage banks, which are non-depository financial institutions that originate loans and sell them to the secondary market. More specifically, the focus is on the development of vertical specialization between these first two parts of the mortgage loan origination process—retail

[9] This discussion is based principally on field and archival research by Jacobides. See Jacobides (2001, 2005).

production (identifying the lender, selecting the right customer, assessing the risk, approving the customer), and wholesale loan management (closing and funding the loan, and then keeping the loan in the bank's portfolio, for subsequent sale of the capital claim to the secondary market for loans). Note that, unlike many other developments in the financial services sector, this particular process of vertical specialization was not influenced, facilitated, or prompted by regulatory changes.

In the early 1980s, several mortgage banks shifted from localized, small players to larger, national entities. As a result, the organization of work within them became increasingly modular. In the early days, mortgage banking executives were involved in all aspects of loan origination and in selling the capital claim to the secondary market. As time progressed, however, firms found the need to create separate functions. The success of the firms that adopted this structure (as well as the imitation of it) led to its proliferation in the industry. Furthermore, the divisions that overperformed the market ensured that their contribution got recognized; and they tried to find a way to break free of the restrictions imposed by what they saw as their relatively inadequate colleagues in the bank. So in the short run, a firm would identify that it was more profitable in, and could productively expand in, only *some* parts of the value chain, such as origination *or* warehousing. Thus, the basic prerequisite of our 'static' model, capability heterogeneity which can, with due TC reduction, lead to specialization, was in place (Jacobides, 2005).

These industry dynamics also illustrate the first evolutionary mechanism. The increasing separation of upstream and downstream, driven by the heterogeneity of capabilities of firms in the industry, led to a very drastic reduction of the variety in cost structures and to an intensification of competitive pressures, a pattern visible in industry statistics (Duncan, 1998). After the emergence of the 'market for closed loans,' necessary to support vertical specialization, the inefficient origination divisions of large integrated firms were simply axed, and in-house loan production was substituted with purchased loan production, done by more efficient firms. In the relatively scalable warehousing and servicing segments, significant concentration came about, as the efficient firms were no longer held hostage to their slow origination growth, and as they squeezed out their less productive competitors.

The next evolutionary mechanism, endogenous reduction of transaction costs, was also in operation in parallel to these competitive developments—indeed, it precipitated them, facilitated by historical happenstance. Given the short-term conditions of relative efficiency in the production process, the quest for market-mediated solutions had begun, with those firms favored by the highest productivity differentials pushing the hardest. Some firms found themselves to be better than average in origination and worse than average in warehousing, and other firms realized the inverse; and this imbalance of capabilities led to the gradual increase of across-firm-boundaries transactions. They thus actively tried to find ways to make such trade possible, through efforts to standardize the requisite information (Jacobides, 2001, 2005). The early (and initially sporadic and exploratory) use of the market produced initial solutions to some of the problems of using the markets, including recognition of the desirability of some standards both for products and ways of contracting. Such adjustments by pioneering mortgage banks were largely one-time costs, which irreversibly disposed the industry toward greater use of the market, thus facilitating further vertical disintegration.

Also, consistent with our third mechanism, the superiority of the specialized entities was a function of their ability to better conduct and improve their operations when focused on only one segment. Differences in culture, incentives, and knowledge bases along the value chain made the combination of the activities in a single firm problematic. In mortgage origination, integration into the local community and links with the builders (who help individuals find a package that matches their finances) are the key success drivers; the mentality in running the business revolves around sales, and around client relationship management. The appropriate incentives are high powered; agents are paid in commissions more than in salaries. The structure should be flexible and responsive, with a knowledge of the local markets, customers and products, and, most important, of how to convince. Wholesale, on the other hand, is an entirely different business. It consists of managing the pipeline of loans; of being able to guess how many loans that are made in a falling interest rate environment will fall off and never close; knowing when to sell these loans to the secondary market; and ensuring there is a good

management of the different types of risks, including underwriting/default risk early on. Compensation is more stable, based on salaries; skills revolve around finance and risk management, and the types of successful executives are vastly different. It is hardly surprising, then, that institutional separation enabled the specialized firms to improve and expand *faster* than integrated production, leading to the benefits that pushed toward further specialization in the sector.

Finally, consistent with the fourth mechanism, the pool of players in the industry changed. New entrants, such as real estate agencies, came to mortgage brokerage on the basis of their capabilities. Likewise, retail banks migrated into mortgage loan warehousing, which they could both understand and perform profitably on the basis of their experience, something that was not true of the previously integrated mortgage banking. So vertical specialization changed the 'points of reference,' the relevant knowledge base, as well as the types of industry participants who could, on the basis of their capabilities, now compete.

On the basis of this new capability development process and the changing roster of participants, the vertically co-specialized warehousers and brokers populated an increasing share of the industry as it evolved.

From disintegration to integration: Swiss watch manufacturing, 1980–1992

A brief illustration of the shift from disintegration to integration, driven by capability differences, comes from the Swiss watch-making industry, discussed in Enright (1995), Maillat *et al*. (1996), and Radov and Tushman (2000). The watch-making industry that had developed in the Jura mountains, close to Geneva, dominated the watch-making business until the 1970s. The industry had slowly evolved into a set of independent manufacturers, each occupying a part of the production process. Some watchmakers specialized in making hands for the watch, some finishing the cases, some working on the straps, etc. The limits of what each did were clear; and the intermediate markets had converged on a simple way to decompose the watch-making activity, with few interdependencies between the parts of the production process, and with well-defined dimensions on which the intermediate markets traded. As long as improvements were needed within each of these modules/predefined parts of production, and as long as these related to the attributes that the intermediate market was based on, the industry prospered—and did so for a long time (Bumbacher, 1992).

This long-successful pattern was disrupted by the appearance of the quartz movement technology. Interestingly, this development arose from one of the vertically specialized firms. However, the Swiss companies failed to grasp its importance or appreciate its commercial potential, and as a result did not pursue this option. The reason was that it was not quartz technology alone that enabled the reorganization of the watch-manufacturing process; while such watches might have been feasible in the context of a vertically co-specialized chain, their efficient production required a major overhaul of the production process (Taylor, 1993; Radov and Tushman, 2000: 2). This was not easy to envisage, let alone engineer, in such a co-specialized environment caught in the set ways of a disintegrated equilibrium (Silver, 1984; Langlois, 1992; Barrett, 1999). Thus, the vertically disintegrated, market-based organization of the watchmakers in the Jura mountains was unable to respond effectively. Their ability to react was circumscribed by the limits implicitly imposed by the existing division of labor in the value chain.

It is telling that the new competitors drew from an entirely different knowledge base—one based on miniaturization and micro-electronics, spearheaded by Japan's Casio and Seiko companies. These companies were initially more integrated, and, more to the point, they were drawing from a more integrated and *new* knowledge base which enables the production of miniaturized and reliable products based on electronic displays.[10] Thus, this

[10] Explaining the lack of response of the Swiss to these new technologies, Maillat *et al*. (1996: 17) suggests 'The development of new skills, especially in micro-electronics, new materials, etc. that were indispensable to watch-making in the 1970s, enabled those bodies [i.e., micro-electronic firms and integrated Japanese producers] to reach other industrial sectors which, like watch-making, needed miniaturized devices, integrated systems, sensors, special materials, etc. Subsequently ... restructuring operations [in the watch manufacturing companies] ... became necessary because the integration of new, constantly-evolving technologies in microtechnology requires extensive scientific and technical knowledge.'

case example illustrates the limits of the old, dis-integrated capability development processes, the emergence of new capabilities which drew from a new, and more integrated knowledge base, and also the resulting competitive selection process which pushed out the co-specialized method of the Swiss, to the advantage of the integrated mode of the Japanese (Radov and Tushman, 2000: 5–6).

Further confirmation of this account comes from a closer look at Swatch—the venture that arguably put Swiss watch manufacturing back on track (Barrett, 1999; Falletti, 2004). The company that manufactures Swatch, SMH (now accounting for 15% of the global watch market), became unusually integrated in the 1980s. It was this very integration that permitted it to develop its innovative Swatch project (Taylor, 1993; Radov and Tushman, 2000). The new production mode was the result of integrated production, not only because it drew on an integrated capability, but also because it required firm-specific investments that were not standardized—so that the TC in the new mode of production were higher.[11] So better capabilities were linked with higher TC, and this led to an increase of integration, which was then followed in the industry (Maillat *et al*., 1996: 17).

DISCUSSION AND CONCLUSION

We have set forth a dynamic view of vertical scope, in which the choices that firms make about their boundaries are conditioned by an evolving industry context. This dynamic analysis reveals important links between transaction costs and capabilities that go well beyond simply using transaction costs and capabilities in an additive function when assessing an individual firm's decision to make or buy (Williamson, 1999). First, we argued that to understand the choice of vertical scope it is necessary to complement the micro-analytic view at the firm level with analysis at the level of the industry and its diverse population of firms. Such analysis reveals transaction cost changes as catalysts of changing vertical scope, yet also identifies capability differences as a necessary driver: absent gains from trade, arising from heterogeneous capabilities, any degree of TC reduction will not lead to specialization.

We further argue that scope is not only statically determined by capabilities as they interact with TC, but also that there are some specific evolutionary mechanisms that explain how scope, TC, and capabilities *co-evolve*. We identified the four key evolutionary mechanisms, which explain both how capabilities affect scope, and, crucially, how scope affects capability. We then illustrate these mechanisms by examining how they work in mutually reinforcing ways to shift an industry's dominant way of organizing from integration to disintegration and back into integration. In particular, we argue that along a particular technological trajectory, while innovations remain incremental on the technology side, induced innovation in transacting practices provide access to specialization gains and produces a secular drift toward disintegration. However, when more radical change occurs in technology, new, integrated capabilities become valuable, and the older, prevailing transacting practices are often rendered obsolete as well, leading to a phase of vertical reintegration.[12] Reintegration may also be further sped up by the effort of specialized firms to avoid the commoditization of their core (specialized) business, or to migrate to more profitable parts of the value chain—themes more fully explored by Cacciatori and Jacobides (2004).

Implications for strategy research

While we claim to have demonstrated the value and importance of the systemic view, it must be acknowledged that this paper is in some respects a preliminary probe of new terrain. We hope it will help stimulate further research, leading to a better understanding of what drives industry structure,

[11] As Falletti (2004: 1) notes, SMH even chose to produce the 1.5-volt chip which was needed for the Swatch, despite the reaction of outside parties which thought such an integration was excessive. Yet this proved useful not only for the integrity of the design of the watch and the reliability, but also as a distinct product stream. SMH sold this chip to other segments, outside the watch-making business, allowing it to remain on the forefront of technology and also develop requisite scale in this critical input. Also, see Nicholas Hayek's interview in Taylor (1993).

[12] See Fine (1998) for a similar view. He, however, attributes the integration phase more to the efforts of dominant specialist firms to leverage their positions into other vertical stages. Our analysis also challenges the argument by Christensen *et al.* (2001, 2002) that integration is always 'better' in terms of satisfying customers, and that dis-integration occurs whenever the inferior but steadily improving dis-integrated solutions 'catch up' with customer requirements. Integration may or may not be superior, depending on the capabilities of integrated firms.)

vertical scope, and capability development. We conclude with a few suggestions along this line.

First, this perspective opens up some unexplored venues for research into the capability- and resource-based view of the firm. In this account, transaction costs shape the trajectories of capability development; they determine the nature of the knowledge acquisition process, and quite possibly the type of competitors that can emerge in an industry, and thus the nature of the underlying resources and capabilities that can be leveraged for competitive advantage (Dierickx and Cool, 1989). Thus vertical specialization enables new types of references, and new competitors, to invade previously integrated sectors; it also calls upon incumbent firms to reconsider the appropriateness of their knowledge bases. The process of capability development and the identity of industry participants change, often irrevocably, as a consequence of changes in vertical scope. This dimension, which has not yet received its due attention (cf. Helfat, 2003; Dosi, Nelson, and Winter, 2001), may well be one of the next challenges for capabilities research. To understand capabilities, we need to follow the shaping roles of transaction costs and scope.

By highlighting the connection between disintegration and industry participants, this perspective illuminates the ongoing, and controversial economic transformations of outsourcing and offshoring. These involve the vertical break-up of previously integrated sectors, on the basis of superior capabilities, leveraged across sectors in newly emerging 'industries' that cut across traditional classifications. Such transformations also suggest that we may need to reconsider the boundaries of existing industries, which are actively redrawn as a result of the process of vertical disintegration and reintegration. This further suggests we may be in need of a new empiricism, which defines industries in a more dynamic way, transcending the traditional SIC definitions, and focuses on the comparative analysis of value chains instead.

Opportunities also open for considering new levels and types of analysis, shifting away from the focus on either individual transactions, or on the particular, unique capabilities and resources of specific firms. The *distribution* of capabilities in an industry, and in particular the distribution of productive capabilities along different parts of the value chain, becomes a key issue, which affects vertical scope and competitive dynamics alike (cf. Jacobides and Hitt, 2004). Questions then arise as to what determines such distributions, i.e., what makes some industries have firms that are capable in all parts of the value chain, and some for which this is not the case. We have suggested that considerations such as similarity of knowledge bases and managerial styles will be important drivers of capability distributions, and, ultimately, of vertical scope. Research on the nature and determinants of capability distributions seems to hold much promise.

We also argue that changes in integration may dramatically affect the patterns of competition if they open up a sector to entirely new capability bases which have been developed in different contexts (Slywotzky, 1996; Jacobides, 2005). Disintegration opens up industries to new breeds of vertically specialized competitors, as we have seen in sectors like financial services; and reintegration, seen in the trend toward 'total solutions' in industrial and customer markets, also changes the nature of capabilities that are brought to bear, the basis of competition, as well as the identity of the participants in the industry. The interplay of changes in TC/scope and competitive dynamics/competitor identification is a promising area for future research.

On the level of strategy and policy, there is an implied need to shift focus from the individual decisions firms make, to the dynamics of industry evolution and the mechanisms that change the business landscape. Executives and policy-makers alike are increasingly interested in understanding the evolutionary dynamics of their business environment, in the interest of extending their limited foresight in dynamic settings. They do care about the forces that shape the structure of their competitive environment; about the nature of the competitors and capability pools that emerge as a result of changes in vertical scope; and the political game of reshaping the institutional and transactional environment. Therefore, a focus on evolutionary mechanisms can thus be a welcome addition to our predictive and normative toolkit.

Finally, on the methodological level, our analysis suggests that while the factors driving individual decisions to integrate or not are important, an exclusive focus on them may block the way to understanding the evolution of the institutional structure of production. The analysis of the evolution of the ISP should complement understanding of the individual transaction, and of the individual

firms' capability, much like the study of evolution should complement our knowledge of botany and zoology. Although great progress has been made in the micro-analytic understanding of firm choices of scope, Coase has pointed the way not merely in the analysis of that aspect of transaction costs, but in proposing the problem of understanding the systemic implications. We believe that we have made a useful start down the path he suggested, but there is clearly much more to be done.

ACKNOWLEDGEMENTS

At various stages of this work, helpful comments have been received from a large number of individuals. First, we would like to acknowledge the guidance, advice, and inspiration provided by our late friend and colleague, Sumantra Ghoshal. He is sorely missed. We also thank Carliss Baldwin, Julian Birkinshaw, Costas Markides, Peter Moran, Phanish Puranam, Freek Vermeulen and conference or seminar participants at Michigan, Stanford, MIT, USC, INSEAD, ETIC (Strasbourg), the Sant'Anna School (Pisa); and at the Academy of Management (Seattle), the Strategic Management Society (Baltimore), and the Organization Science Winter Conference. We would also like to thank Associate Editor Will Mitchell and two anonymous referees for useful comments. Jacobides acknowledges financial support by the Centre for the Network Economy and the Leverhulme Trust Project on Digital Transformations at the London Business School; Winter acknowledges financial support by the Reginald H. Jones Center at the Wharton School.

REFERENCES

Adler P. 2003. Making the HR outsourcing decision. *Sloan Management Review* **45**(1): 53–60.

Afuah A. 2001. Dynamic boundaries of the firm: are firms better off being vertically integrated in the face of a technological change? *Academy of Management Journal* **44**(6): 1211–1229.

Albert S, Whetten D. 1985. Organizational Identity. In *Research in Organizational Behavior*, Vol. 7, Cummings LL, Staw BM (eds). JAI Press: Greenwich, CT; 263–295.

Alchian AA, Demsetz H. 1972. Production, information costs and economic organization. *American Economic Review* **62**(5): 777–795.

Argyres NS. 1996. Evidence on the role of capabilities in vertical integration decisions. *Strategic Management Journal* **17**(2): 129–150.

Argyres NS, Liebeskind J. 1999. Contractual commitments, bargaining power, and governance inseparability: incorporating history into transaction cost theory. *Academy of Management Review* **24**(1): 49–64.

Argyres NS, Mayer K. 2004. Contract design capability and contract performance in high technology industries: implications for the roles of managers, engineers and lawyers. Manuscript, Boston University School of Management.

Baldwin CY, Clark KB. 2003. Where do transactions come from? Working paper, Harvard Business School, 11 February.

Barney JB. 1991. Firm resources and sustained competitive advantage. *Journal of Management* **17**(1): 99–120.

Barrett ME. 1999. Time marches on: the worldwide watch industry. Thunderbird GSM Case A07-99-0006.

Barzel Y. 1982. Measurement cost and the organization of markets. *Journal of Law and Economics* **25**(1): 27–48.

Bumbacher U. 1992. The Swiss watch industry. Harvard Business School Case 792-046.

Cacciatori E, Jacobides MG. 2004. The limits of the market: vertical re-integration explained. Working paper, SIM Area, London Business School.

Chandler AD. 1962. *Strategy and Structure*. MIT Press: Cambridge, MA.

Chandler AD. 1977. *The Visible Hand*. Harvard University Press: Cambridge, MA.

Chesbrough H, Teece D. 1996. When is virtual virtuous? Organizing for innovation. *Harvard Business Review* **74**(1): 65–72.

Christensen C, Raynor M, Verlinden M. 2001. Skate to where the money will be. *Harvard Business Review* **79**(10): 72–82.

Christensen C, Verlinden M, Westerman G. 2002. Disruption, disintegration and the dissipation of differentiability. *Industrial and Corporate Change* **11**(5): 955–994.

Coase RH. 1937. The nature of the firm. *Economica* **4**: 386–405.

Coase RH. 1988. The nature of the firm: meaning. *Journal of Law, Economics and Organization* **4**(Spring): 19–32.

Coase RH. 1992. The institutional structure of production. *American Economic Review* **82**(4): 713–719.

Cohen W, Levinthal D. 1990. Absorptive capacity: a new perspective on learning and innovation. *Administrative Science Quarterly* **35**(1): 128–152.

Collis DJ, Noda T. 2001. The evolution of intra-industry firm heterogeneity: insights from a process study. *Academy of Management Journal* **44**(4): 897–925.

Crane DB, Bodie Z. 1996. Form follows function: the transformation of banking. *Harvard Business Review* **74**(2): 109–117.

Cusumano M, Mylonadis Y, Rosenbloom R. 1992. Strategic maneuvering and mass—market dynamics:

the triumph of VHS over beta. *Business History Review* **66**(Spring): 51–94.
Demsetz, H. 1988. *The Organization of Economic Activity*. New York: Blackwell.
Dierickx I, Cool K. 1989. Asset stock accumulation and sustainability of competitive advantage. *Management Science* **35**(12): 1504–1514.
DiMaggio P, Powell W. 1983. The iron cage revisited: institutional isomorphism and collective rationality in organizational fields. *American Sociological Review* **48**: 147–160.
Doig SJ, Ritter RC, Speckhals K, Woolson D. 2001. Has outsourcing gone too far? *McKinsey Quarterly* (4): 25–37.
Dosi G, Nelson RR, Winter SG. 2001. *The Nature and Dynamics of Organizational Capabilities*. Oxford University Press: Oxford.
Duncan DG. 1998. Relative profitability of the mortgage industry. *Mortgage Banking*, June: 95.
Enright MJ. 1995. Organization and coordination in geographically concentrated industries. In *Coordination and Information*, Lamoreaux NR, Raff D (eds). MIT Press: Cambridge, MA; 103–142.
Evans P, Wurster T. 1999. *Blown to Bits*. Harvard Business School Press: Boston, MA.
Falletti R. 2004. *Swatch: A Case*, DisegnoIndustriale. http://www.disegnoindustriale.com/progetti/swatch-01eng.html [15 August 2004].
Fine C. 1998. *Clockspeed*. Perseus Books: Reading, MA.
Foote NW, Galbraith J, Hope Q, Miller D. 2001. Making solutions the answer. *McKinsey Quarterly* (3): 84–93.
Hagel J, Singer M. 1999. Unbundling the corporation. *Harvard Business Review* **77**(2): 133–141.
Hart O. 1995. *Firms, Contracts and Financial Structure*. Oxford University Press: New York.
Helfat CE (ed). 2003. *The SMS Blackwell Handbook of Organizational Capabilities: Emergence, Development, and Change*. Blackwell: New York.
Henderson R, Clark KB. 1990. Architectural innovation: the reconfiguration of existing systems and the failure of established firms. *Administrative Science Quarterly* **35**: 9–30.
Hoetker G. 2005. How much you know versus how well I know you: selecting a supplier for a technically innovative component. *Strategic Management Journal* **26**(1): 75–96.
Jacobides MG. 2001. Mortgage banking unbundling. *Mortgage Banking* **61**: 28–39.
Jacobides MG. 2005. Industry change through vertical dis-integration: how and why markets emerged in mortgage banking. *Academy of Management Journal* **48**(3).
Jacobides MG. 2004. How capabilities, transaction costs and scalability interact to drive vertical scope. Working paper, London Business School.
Jacobides MG, Croson DC. 2001. Information policy: shaping the value of agency relations. *Academy of Management Review* **26**(2): 202–224.
Jacobides MG, Hitt LM. 2004. Losing sight of the forest for the trees? productive capability differences as drivers of vertical scope. Leverehulme working paper, London Business School.
Karim S, Mitchell W. 2000. Path dependent and path breaking change: reconfiguring business resources following acquisitions in the U.S. medical sector: 1978–1995. *Strategic Management Journal* **21**(10/11): 1061–1081.
Klepper S. 2002. The capabilities of new firms and the evolution of the U.S. automobile industry. *Industrial and Corporate Change* **11**(4): 645–666.
Lamoreaux N, Raff D. 1995. *Coordination and Information*. Conference report (NBER). University of Chicago Press: Chicago, IL.
Lamoreaux N, Raff D, Temin P. 2003. Beyond markets and hierarchies: towards a new synthesis of American business history. *American Historical Review* **108**: 404–433.
Langlois RN. 1992. Transaction cost economics in real time. *Industrial and Corporate Change* **1**: 99–127.
Langlois RN. 2003. The vanishing hand: the changing dynamics of industrial capitalism. *Industrial and Corporate Change* **12**(2): 351–385.
Langlois RN, Foss NJ. 1999. Capabilities and governance: the rebirth of production in the theory of economic production. *Kyklos* **52**(2): 201–218.
Langlois RN, Robertson PL. 1995. *Firms, Markets, and Economic Change: A Dynamic Theory of Business Institutions*. Routledge: London.
Lawrence P, Lorsch J. 1967. *Organization and Environment*. Harvard University Press: Cambridge, MA.
Levinthal DA. 1997. Adaptation in rugged landscapes. *Management Science* **43**(7): 934–950.
Lieberman MB, Dhawan R. 2001. Assessing the resource base of Japanese and U.S. auto producers: a stochastic frontier production function approach. Working paper, UCLA.
Madhok A. 2002. Reassessing the fundamentals and beyond: Ronald Coase, the transaction cost and resource-based theories of the firm and the institutional structure of production. *Strategic Management Journal* **23**(6): 535–550.
Maillat D, Lechot G, Lecoq B, Pfister M. 1996. Comparative analysis of the structural development of milieux: the example of the watch industry in the Swiss and French Jura Arc. Working paper, University De Neufchatel, Switzerland. http://www.unine.ch/irer/wp9607.doc [15 August 2004].
Mayer K, Argyres NS. 2004. Learning to contract: evidence from the personal computer industry. *Organization Science* **15**(4): 394–410.
Meyer J, Rowan B. 1977. Institutionalized organizations: formal structure as myth and ceremony. *American Journal of Sociology* **83**: 340–363.
Nelson RR, Winter SG. 1982. *An Evolutionary Theory of Economic Change*. Belknap Press of Harvard University Press: Cambridge, MA.
North DC. 1986. The new institutional economics. *Journal of Institutional and Theoretical Economics* **142**: 230–327.
Penrose E. 1959. *The Theory of the Growth of the Firm*. Oxford: Blackwell.

Peteraf M, Barney JB. 2003. Unraveling the resource-based tangle. *Managerial and Decision Economics* **24**(4): 309–323.

Peteraf M, Shanley M. 1997. Social learning and the 'fundamental paradox' of transaction cost economics. In *Advances in Strategic Management*, Vol. 14, Walsh JP, Huff AS (eds). JAI Press: Greenwich, CT; 193–222.

Poppo L, Zenger TR. 1998. Testing alternative theories of the firm: transaction cost, knowledge-based, and measurement explanations for make-or-buy decisions in information services. *Strategic Management Journal* **19**(9): 853–878.

Porter G, Livesay HC. 1971. *Merchants and Manufacturers*. Johns Hopkins: Baltimore, MD.

Porter ME. 1996. What is strategy? *Harvard Business Review* **74**(6): 61–78.

Prahalad CK, Hamel G. 1990. The core competence of the corporation. *Harvard Business Review* **68**(3): 79–91.

Radov DB, Tushman ML. 2000. Rebirth of the Swiss watch industry, 1980–1992 (A). Harvard Business School Case 9-400-087, 12 June.

Raynor M, Christensen C. 2002. *Integrate to Innovate*. Deloitte Research Publication (www.deloitte.com/research).

Richardson G. 1972. The organisation of industry. *Economic Journal* **82**: 883–896.

Riordan MH, Williamson OE. 1985. Asset specificity and economic organization. *International Journal of Industrial Organization* **3**: 365–378.

Rivkin J. 2001. Reproducing knowledge: replication without imitation at moderate complexity. *Organization Science* **12**: 274–293.

Rosenbloom RS. 2000. Leadership, capabilities and technological change: the transformation of NCR in the electronic era. *Strategic Management Journal*, Special Issue **21**(10–11): 1083–1103.

Schilling MA, Steensma H. 2001. The use of modular organizational forms: an industry-level analysis. *Academy of Management Journal* **44**(6): 1149–1169.

Siggelkow N. 2001. Change in the presence of fit: the rise, the fall, and the renaissance of Liz Claiborne. *Academy of Management Journal* **44**: 837–858.

Shanley M, Peteraf M. 2004. Vertical group formation: a social process perspective. *Managerial and Decision Economics* **25**(6–7): 291–298.

Silver M. 1984. *Enterprise and the Scope of the Firm*. Martin Robertson: London.

Slywotzky A. 1996. *Value Migration: How to Think Several Moves Ahead of the Competition*. Harvard Business School Press: Boston, MA.

Taylor W. 1993. Message and muscle: an interview with Swatch titan Nicholas Hayek. *Harvard Business Review* **71**(2): 98–110.

Teece DJ. 1980. Economies of scope and the scope of the enterprise. *Journal of Economic Behavior and Organization* **1**(3): 223–247.

Teece DJ, Rumelt R, Dosi G, Winter S. 1994. Understanding corporate coherence: theory and evidence. *Journal of Economic Behavior and Organization* **23**: 1–30.

Teece DJ, Pisano G, Shuen A. 1997. Dynamic capabilities and strategic management. *Strategic Management Journal* **18**(7): 509–533.

Tripsas M, Gavetti G. 2000. Capabilities, cognition and inertia: evidence from digital imaging. *Strategic Management Journal*, Special Issue **21**(10-11): 1147–1161.

Walker G, Weber D. 1984. A transaction cost approach to make-or-buy decisions. *Administrative Science Quarterly* **29**(3): 373–392.

Wernerfelt B. 1984. A resource-based view of the firm. *Strategic Management Journal* **5**(2): 171–180.

Whetten DA, MacKey A. 2002. A social actor conception of organizational identity. *Business and Society* **41**(4): 393–414.

Williamson OE. 1975. *Markets and Hierarchies: Analysis and Antitrust Implications*. Free Press: New York.

Williamson OE. 1983. Organizational innovation. In *Entrepreneurship*, Ronen J (ed). Lexington Books: Lexington, MA; 101–134.

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

Williamson OE. 1999. Strategy research: governance and competence perspectives. *Strategic Management Journal* **20**(12): 1087–1108.

Winter SG. 1988. On Coase, competence, and the corporation. *Journal of Law Economics and Organization* **4**: 163–180.

Winter SG. 2003. Understanding dynamic capabilities. *Strategic Management Journal*, Special Issue **24**(10): 991–995.