RIPPLES THROUGH THE VALUE CHAIN:
HOW AN UPSTREAM INNOVATION SHAPES PROFIT AND SCOPE IN A
SECTOR

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
and
Faculdade de Ciências Económicas e Empresariais
Universidade Católica Portuguesa
Palma de Cima 1639-023 Lisbon, Portugal
e-mail: fveloso@fcee.ucp.pt
Phone: +351 21 721-4241
Fax: +351 21 727-0252

CLAUDIO WOLTER
Department of Engineering and Public Policy
Carnegie Mellon University
5000, Forbes Avenue, Pittsburgh, PA 15213-3890, USA
e-mail: cwolter@alumni.cmu.edu
Phone: +1 412 480-8774
Fax: +1 412 268-3757

November 20, 2012 –Version 6.5 / S

Submission for the 2013 Wharton Technology Conference
RIPPLES THROUGH THE VALUE CHAIN:
HOW AN UPSTREAM INNOVATION SHAPES PROFIT
AND SCOPE IN A SECTOR

Abstract

This paper considers how an innovation ‘ripples through’ a sector consisting of a value chain with upstream suppliers, downstream assemblers and integrated firms. We show how profit migrates along the chain as a result of an upstream innovation, and suggest that downstream assemblers, even without any strategic foresight or manipulation, can become a ‘bottleneck.’ How much they profit from their ‘dumb luck’ depends on whether the innovation is a ‘zero-level’ productive innovation or a dynamic change (in learning rates), and to whom the innovation is available. We also show that the level of Transaction Costs (TCs) shapes profitability dynamics along a sector. By becoming more than just a cost in the competitive context, TCs act as an ‘isolating mechanism’ and allow firms to profit.

Keywords: Value chain, industry evolution, profit migration, innovation, bottleneck.
Our world is increasingly one of sectors comprising of interconnected segments. In industries such as telecommunications, financial services, and pharmaceuticals, vertical specialists coexist with partially integrated competitors. Firms are becoming increasingly aware that their fortunes depend on the dynamics of the broader ecosystem and their value-chain partners (Adner, 2012), and there is growing interest in how changes in one part of the sector can affect another. The focus, though, has been at the level of the firm. From Teece’s (1986) seminal analysis to research on how vertical links to complementors shape performance (Adner and Kapoor, 2010; Afuah, 2000; Kapoor, 2012), the focus has been on the plight of one given firm. Yet what happens at the level of the segment as a whole, given the sector dynamics around it? Jacobides, Knudsen, and Augier (2006) speculate that profitability may depend on the comparative dynamics between adjoining stages of the value-adding process (which we term ‘segments’), and on ‘industry architecture,’ i.e. the way these different stages are connected, but do not provide concrete detail on how these sector-level connections operate.

As practitioners and strategy researchers examine these issues, it becomes apparent that we still lack answers to some simple questions about the dynamics of a competitive setting. In particular, how does an upstream innovation affect profitability dynamics downstream at the level of the entire sector? For example, how does a change in laser component technology affect firms making laser systems – or integrated firms? Studying the laser sector, Bhaskarabhatla and Klepper (2012) show how an upstream innovation in semiconductor (diode) pumps for solid-state (DPSS) lasers changed both the makeup of the industry and its survival dynamics. The laser sector, like many others, consisted of upstream suppliers, downstream assemblers / Original Equipment Manufacturers (OEMs) and vertically integrated OEMs doing their own supply. The upstream innovation changed the sector drastically, leading not only to a displacement of the old upstream technology (and the
demise of the firms that produced it), but also to a concurrent shakeout of downstream OEMs and integrated firms. Or consider Kapoor’s (2012) analysis of the semiconductor sector, where some semiconductor firms engaged in fabless design, others focused on the operations of fabs (i.e., factories of microchips), and others remained integrated into both activities. Innovations would happen in all segments, and Kapoor clearly demonstrates that changes in one part of the value chain affect the entire sector – specialists and integrated firms alike. Yet the question remains: How, precisely, do such innovations ‘reverberate’ along a value chain? What are the aggregate patterns that emerge? What drives them?

Although such structural relationships are common, they have rarely been studied from this perspective. Sectors such as disk drives (Christensen, 1997), biotechnology (Murray, 2004), photolithography (Henderson, 1993), and chemical products (Lieberman, 1989) all feature interrelated segments and have been intensively studied. Yet nearly all research has focused on a single segment, leaving the interdependencies unexamined.

The few empirical studies have focused on the plight of an individual firm, and its relative competitive advantage as a function of its upstream and downstream links. Adner and Kapoor (2010), for instance, focus primarily on the competitive advantage of a focal firm as a function of the challenges facing its upstream and downstream complementors. Afuah (2000) considers how links between workstation makers and their microprocessor suppliers affected their ability to innovate. Research drawing on Teece (1986) has looked at how contractual issues affect a firm’s ability to appropriate innovative rents, even though the fallacy of aggregation suggests that what might work for one firm might not work sector-wide. Finally, in terms of theory, some formal work focuses on explaining the strategic game between interrelated monopolies – such as Intel with its links to Microsoft and the PC sector (Casadesus-Masanell and Yoffie, 2007). Such cases, with their unusual network externalities
and their focus on players that ‘own’ a sector, do little to help our baseline intuition of the competitive case.

This is our focus: how upstream changes affect downstream dynamics and the balance between firms in the sector. We provide a computational model that explores how changes in an upstream segment can affect not only that segment itself, but also its downstream complementors and any vertically integrated firms. By looking at these competitive dynamics, we hope to understand how profit shifts along the value chain following a significant upstream innovation.

Our starting point is competitive heterogeneity. We argue that firms differ in their abilities, and competition reduces this heterogeneity over time through selection and differential growth. So we allow heterogeneity to play itself out over time, and consider its impacts on profitability – and scope. However, we also posit that relatively ineffective firms may still stay in the market, since Decreasing Returns to Scale (DRS) limit the growth of superior firms. Building on these assumptions, we create a baseline model of industry evolution and shifting profits, before moving on to the more interesting question: How does an upstream innovation affect these dynamics?

Focusing on an upstream innovation (or ‘technological shock’), we look at whether a one-off increase in firms’ productivity shapes profit dynamics, not only in the innovator’s own segment (the ‘focal segment’), but also downstream. We also look at whether a different type of innovation – a change in firms’ ability to improve, which is a type of dynamic capability – produces different results.

We find that when there is an innovation upstream, the downstream sector is able to capture a good deal of value-add, in spite of its own inaction. Indeed, under some constraints, it can capture much more value than the innovating segment itself. This suggests that simply
“being there” when there is a flurry of innovation in the vertical complements can yield rich pickings – in some cases, even transforming firms or entire segments into ‘bottlenecks,’ i.e. the areas that are most profitable and attractive (Jacobides et al., 2006; Rosenberg, 1969). We also find that the type of innovation matters a great deal: A ‘one-off’ discovery that improves effectiveness is more helpful to the downstream segment than an upstream innovation in terms of dynamic capability.

We present a detailed set of findings showing how the upstream innovation changes both profitability and vertical scope, and how changes in profitability and scope interact. In particular, we provide entire sets of results for different Transaction Cost (TC) levels, showing how TCs mediate value-chain and profit dynamics. These findings contribute to the recent literature on scope and capability co-evolution (see Argyres et al., 2012; Jacobides and Winter, 2005; 2012). We argue that TCs mediate profitability because they act as an ‘isolating mechanism.’ Turning to the way TCs affect ‘ripples’ along the value chain, we show that, under some conditions, high TCs will lead to relatively more profits upstream, but even greater value capture downstream.

Our findings facilitate a structured discussion around the managerial concept of ‘value migration’ (Christensen and Raynor, 2003; Slywotzky, 1996), and show how ‘architectural’ conditions in one segment shape profit patterns in the entire sector. We also apply our insights to the ‘open’ innovation strategies followed by firms in telecoms and software.

This paper is structured as follows. The following section presents the theoretical background, reviews directly relevant research, and motivates our study. Next, we provide a section on the basics of the research design and computational model, describing how simulations were used to uncover the evolutionary pattern of interactions between the variables of interest. This is followed by the presentation and discussion of our model’s results. The final section places our findings in the context of current research and managerial
THEORETICAL MOTIVATION

While the question of how innovation ripples through a value chain has not been directly addressed in the literature, some work has shed light on related topics. Teece (1986) paved the way for subsequent research that looked at how transactional structures enabled a focal firm, or the owners of complementary assets, to benefit from innovation (e.g. Tripsas, 1997). His focus, though, was on the extent to which one firm is able to appropriate the fruits of its innovative efforts as a result of contractual issues linking it with downstream operations; and his emphasis was on contractual form, not on the distribution and study of profit change. In terms of performance (albeit, focusing on innovation as opposed to profits), Afuah (2000) analyzes the changes in the workstation sector as a result of the shift between the CISC and RISC microprocessor architectures and finds that being integrated into the old technology, as well as using a supply base steeped in it, did indeed hamper performance. Adner and Kapoor (2010) look at the photolithography sector and add a further nuance. They show that upstream constraints or difficulties of a firm’s complementors correlate with superior performance of a leading firm, since they can both act as entry barriers and protect the focal firm’s competitive advantage from being eroded by competition. They also find that downstream constraints reduce the effectiveness of a leading focal firm, because they limit its downstream market access. Both these empirical studies underline the importance of

---

1 This stands in contrast to the analysis of the ‘ecological’ implications of an innovation in one segment, and the way it affects another segment. It is, in essence, a fundamentally different research design and question we are interested in, which complements the Teecean analysis. Teece himself, in later work, recognized the need to move beyond the focus on one given firm and its strategies, and shift to the analysis of the overall ‘industry architectures’ (Pisano and Teece, 2007); yet he has not focused on how changes in one segment affect the dynamics on another.
connections between vertically related segments. Yet their focus is largely on incumbents in a focal segment and their plight, on the basis of their links up- and downstream. Yet they do not consider how a segment as a whole fares as its upstream or downstream complement changes, neither do they track profits.

Jacobides et al. (2006) work on ‘industry architectures’ makes more directly related conjectures, which we formalize and test. They make the point that firms may want to manipulate the conditions of entry and competition in the segments where their complementors operate, and benefit as a result. They suggest that segments can thus become ‘bottlenecks’ and that those bottlenecks can capture a disproportionate share of the industry-wide profits. Thus, they hypothesize that changes (including innovation) affect relative mobility both within and into a segment, a thesis that is analyzed in the entertainment sector by Ferraro and Gurses (2009) and in the computer sector by Jacobides and Tae (2012) – also, see Pisano and Teece (2007). This emerging body of work suggests two observations. First, firms try to shape the architecture of their sector to their advantage – or to the advantage of their vertical segment. Second, changes in conditions in one part of the value chain shape those in another. It is the latter insight that we try to advance.

Our paper also extends related computational models. For instance, Malerba et al. (2008) provide a model of the ‘history-friendly’ evolution of the computer ecosystem, providing a rationale for the observed patterns of disintegration and concentration. Their model explains how heterogeneous firms compete, sometimes as integrated entities and sometimes as specialized ones, and suggests that conditions along the value chain affect the collective outcome. Their focus, as with much of the preceding literature, is on scope and concentration

---

2 One way forward is to consider the ability of firms to impose the terms by which they trade; i.e., to say that some firms along the value chain play complex strategic games, both in deciding their scope, and setting their terms of interaction (see Fixson and Park (2008); Gereffi, Humphrey and Sturgeon (2005)). Such models will always depend on the distribution of power in the sector, and cannot yield generalizable predictions. In particular, they are of no help to understand cases such as the laser sector, where market power is clearly not a substantial issue. Another possibility is to focus on firms who are simply trying to ‘do the best they can’, and profit along the way – provided there is some heterogeneity to account for profits. This is the approach we will pursue.
levels, profits being incidental to their analysis – whereas our objective is to study profit
dynamics explicitly, leaving scope changes in the background. Arora and Bokhari (2007) also
provide a computational model of vertical structure that deals with scope rather than profits.\(^3\)

Moving to more tangentially related research, Industrial Organization (I/O) has
considered interdependencies along a value chain, albeit in a very narrow and delimited way.
One major stream of literature has focused on the role of successive oligopolies, and on the
potential distortions they can bring to pricing and welfare (see Perry (1989) for a review). A
separate but related line of research has focused on the particular interactions of two
vertically related stages in the presence of both complementarity and/or network externalities,
motivated in particular by the Wintel case (i.e., the link between Microsoft Windows and
Casadesus-Masanel and Yoffie (2007) provide a formal model showing how two
complementors (Microsoft and Intel) are yoked together through their joint ownership of
Wintel, the quasi-monopoly platform. They also show how changes in upstream and
downstream R&D investments are connected to these two firms’ payoffs, leading to complex
and not always harmonious competitive dynamics. Although the focus is on R&D
investments and their competitive outcome, the results are highly contingent on the
specificities of this oligopolistic model, which also shows unusual network externalities and
lock-ins. Nevertheless, this approach provides yet another example of value-chain
interdependencies.

A more distantly related part of the literature looks at the role of Transaction Costs (TCs)
and capabilities in industry evolution (with or without innovation). TCs have had a long and

\(^3\) A rich discussion on the same empirical basis, which further motivates our model, can be found in Bresnahan and
Greenstein (1999). They provide a detailed verbal analysis of the evolution of the computer sector, with keen interest in the
entry process, changes in scope, and the results in terms of the overall level of integration, concentration, and (less
systematically) profitability levels in the sector. Langlois and Robertson’s (1995) book provides further empirical examples
of vertically interrelated sectors and histories of how certain parts of the value chain have co-evolved with others. We
consider such systemic analyses of sectors as motivation for our work.
distinguished pedigree since Coase’s (1937) seminal contribution. Interestingly, though, the
dependent variable is almost always some version of vertical scope, whether one takes
Transaction Cost Economics (TCE; Williamson, 1985), the Resource-Based View (RBV;
Leiblein and Miller, 2003), or the more recent synthesis between institutional and
evolutionary views (Jacobides and Winter, 2005, 2012). Profit evolution, our main interest, is
treated as an incidental variable that affects firms’ decisions about whether or not to integrate.
While our model will look at scope and vertical integration, our main focus is profits. Having
said that, our analysis complements the discussion on scope and capability co-evolution,
since it does consider profitability and scope as a function of TCs and heterogeneity – a
theme we will return to in our discussion.

In terms of modeling, our paper draws on the platform developed by Jacobides (2008),
who examines the co-evolution of scope and capabilities under different TC conditions. His
setup comprises two vertically linked segments, where firms can choose whether to use their
internal upstream segment or the market at large (or both) for the production of a component
needed in the assembly of a final product. He shows that heterogeneity of capabilities, and in
particular the correlation patterns between upstream and downstream segments, plays a
determining role, and that TCs act as an ‘operator’ on the underlying ‘bed’ of capability
heterogeneity. We retain the modeling platform he built, but we focus on profitability, not
only scope, and, most importantly, we change the ‘experimental design’ in terms of the
simulation questions we ask. Rather than allowing firms to choose their scope endogenously,
we posit that there are upstream suppliers, downstream assemblers, and integrated firms, and
explore the impact of innovations on all types of participants, under varying TC conditions.

Our final observation on modeling relates to our choice of a competitive analysis as
opposed to an oligopolistic setup (e.g., à la Casadesus-Masanel and Yoffie (2007); or Perry
(1989)). This is motivated by two factors. First, taking the competitive case allows us to
provide some fairly general results, dependent only on production assumptions. As noted before, the competitive case is an important baseline, and one that has hardly been addressed. It is also possible to superimpose a strategic layer on these results, once we have them. Second, evidence has been amassing on the ‘persistent performance differences between seemingly similar enterprises’ (Gibbons et al., 2010), and large-scale studies have shown that differentially capable firms are in fact the rule (Bloom and Van Reenen, 2007; Syverson, 2011). Therefore, profits may not be driven (or even primarily driven) by oligopolistic restraints. Rather, they will be sustained by efficiency differentials (Lippman and Rumelt, 1982, 2003; Jacobides, Winter, and Kassberger, 2012; Winter, 1995). The novelty of our approach is not just to look at such efficiency-differential-based profits, but also to consider them in the context of an interdependent value chain and see how innovation (under varying TCs) shapes the outcome.

In conclusion, ‘ripples along the value chain’ have not received direct attention in the academic literature, although they have been discussed in the popular business press, offering some evidence that the phenomenon exists. Slywotzky’s (1996) discussion of ‘value migration’ considers some value-chain dynamics, and Christensen and Raynor (2003) go as far as to posit a ‘law of conservation of attractive profits’ whereby profits don’t change, but simply jump around between vertically related segments. But these ideas have not received any analytical scrutiny – hence our model, which focuses on a specific and tractable piece of the puzzle.

**RESEARCH DESIGN AND MODEL STRUCTURE**

4 For completeness’ sake, it might be worth noting that a related analysis exists in terms of whether profits accrue to firms, or to some resources that firms might own or employ. Rather than using a ‘ripples along the value chain’ model, this approach asks ‘where do profits accrue (firms or resources)?’ However, the thorny question arises of whether these are profits, or rents to some underlying scarce resources not properly accounted for (Lippman and Rumelt, 2003; Jacobides et al., 2012; Winter, 1985). The intuition here dates back to Ricardo (1817) and his incisive analysis of why taxes on agricultural products would only benefit landowners who had more fertile land in their possession, given that land was in scarce supply: The idea is that fixed resources of some sort needed for production can become a capacity bottleneck, and thus accrue profits through rents they will obtain (if markets function properly). The problem we want to tackle, though, is different: We want to understand how vertically related segments, each of which has its own resources (which we will assume are priced in the market) make profits. We allow for firms to expand, and new ones to enter, so it is not ‘input scarcity’ that we will consider here.
We build a computational model with two vertically connected segments and firms of heterogeneous capabilities, whose formal details and parameterization are discussed in the Appendix. Since we want to consider the dynamics between different types of firms, we create three groups: integrated firms, specialized suppliers (or upstream specialists), and specialized assemblers (or downstream specialists). Integrated firms don’t pay TCs.

Specialized suppliers and specialized assemblers, who need each other’s services to produce, pay per valorem TCs as they procure an intermediate product or component on the market, and they do so through a ‘pool,’ i.e. without bilateral links between firms. All three types of firm are of differential capability – some good, others less so. The industry, like most new industries, starts with excess demand and profits that are gradually, yet never entirely, competed away.

Our model sets up incumbents who compete, consider costs and market prices, and maximize their profits, without exercising market power or engaging in output restraint. We then allow for entry (of more efficient firms) and exit (of loss-making firms). As more efficient firms displace less efficient firms, we see prices go down – but we also see scope and, crucially, profits changing as a result. We also allow firms to learn by doing, in order to enhance the verisimilitude and relevance of the model further.

As we are explicitly abstracting away from oligopolistic profits to focus on efficiency (competitive) profits, we have not provided for oligopolistic or strategizing behavior in our setting. In keeping with Williamson’s (1985) approach, our emphasis is on ‘economizing’ as opposed to ‘strategizing,’ partly because we find it useful and an important benchmark, and partly because the economizing element has not been adequately studied. So firms compete, but do not manipulate the market: They are price-takers. Our empirical observations of fast-moving sectors such as telecommunications and media also lead us to believe that firms are often focused on competing effectively in their own markets, as opposed to engaging in
cartel-type behavior to shape their downstream dynamics. We do, however, consider heterogeneous firms, with differing capabilities. Furthermore, we present firms that experience Decreasing Returns to Scale (DRS), so that the better firm cannot expand without limit (or cost) and capture the entire market effortlessly. The assumption of DRS, which we share with Lippman and Rumelt (1982) as well as Arora and Bokhari (2007), requires some explanation, as it suggests that a firm with a superior ‘recipe’ cannot replicate it and capture the market. This may be due to tacit, non-replicable knowledge; the existence of costs that make such expansion undesirable; or any fixed factor that cannot be indefinitely leveraged (such as customer base, customer loyalty, reputation, entrepreneurial / managerial attention or focused expertise in a region). While this choice has its limits (see Jacobides et al. (2012) for a critique and an alternative suggestion), it still provides us with a realistic structure for u-shaped average cost curves that are consistent with empirical observations. However, we have run an entirely different specification under Constant Returns to Scale (CRS), using Jacobides et al.’s (2012) approach, and the main results on profit ‘ripples’ are robust.  

Entry and exit play an important role in our model – as in reality. Entry happens at every period: We allow a new firm to enter, and its upstream and downstream capabilities are randomly drawn from independent distributions. On the basis of these capabilities, firms decide whether they will compete as upstream suppliers, downstream assemblers, or integrated producers, choosing the form that will yield the largest profits. Alternatively, if their draw is not good enough for them to produce above fixed costs, they do not enter at all. Firms exit as soon as they become loss-making.

The model yields a platform that allows us to pose a series of meaningful ‘if-then’ experiments, where for a given level of initial capability heterogeneity, and a given/fixed

---

5 Details on the CRS available upon request
initial distribution of specialized and integrated players, we observe the adjustment paths for profits (and scope) that emerge. The inclusion of TCs allows us to vary them as a parameter to (a) establish their role in shaping patterns of scope and profitability evolution; and (b) use them as a potential moderator in terms of how an innovation in one segment might affect others. To accomplish the latter, we introduce a technological ‘shock to the system’ to see how it ‘ripples through’ the entire sector in term of profit migration and its interaction with TCs. When the innovation takes place, the capability distribution upstream gets ‘re-based’ so that the average (post-innovation) firm is as good as the best (pre-innovation) firm still in the industry.

Since innovation is our focus, we pay particular attention to operationalization. We consider the case of innovation in the upstream segment alone and then how TCs mediate aggregate scope and profits along the value chain. We explore two different interpretations of upstream innovation. First, we model it as a change in the productive capabilities available to firms – i.e. a change in the ‘zero-level capabilities’ (Winter, 2003), or simply the efficiency with which firms turn inputs into outputs. Second, we consider innovation as a change in productive capabilities coupled with a change in rates of improvement. This is consistent with the ‘dynamic capabilities’ approach (Teece, Pisano, and Shuen, 1997), where innovation can be thought of as a change in improvement (or learning) rates.

The Appendix provides the formal details of our model, including parameterization and calibration, as well as robustness checks, with relevant methodological discussion and references.

RESULTS

Baseline results

The top panel in Figure 1 represents the natural evolution of the sector’s vertical scope for
three TC levels – low, medium and high – without the innovation being introduced. Scope is measured as the share of final product sold by integrated firms, as compared with the total production, which includes specialized assemblers as well. We start with the same uniform distributions of upstream and downstream capabilities, and then run the full stochastic ‘history’ of the sector. We repeat the exercise 30 times, so that we obtain results that show us both the mean history of the sector and the variance that can occur, lest we overestimate the impact of our results. Thus, the top of Figure 1 shows the evolution of the mean aggregate scope for 30 runs, and of the standard deviation for these same runs.

**INSERT FIGURE 1 HERE**

Long-run equilibrium tends to partial integration; the sector does not become totally integrated or disintegrated. While this result is consistent with Jacobides (2008), the present model differs because it includes entry and exit. At the very beginning of the industry, entry of integrated firms is more common due to the opportunity of capturing profits in both upstream and downstream segments. Hence we see in Figure 1 that integration initially increases for all TC levels. With low (or no) TCs, however, the probability of integrated entry decreases more quickly than the probability of specialized entry (either upstream or downstream), as noted by Arora and Bokhari (2007) – a result we replicate here. With Darwinian selection (and learning) driving the average capabilities of firms in the industry upwards via entry and exit, the probability that any capability draw will be large enough to allow for profitable entry decreases. But because the draws upstream and downstream are independent, potential entrants need two good draws to enter as an integrated firm, which becomes less probable much more quickly. The result is that scope reverts its path and moves towards disintegration at this TC level, before entry ceases and the sector stabilizes. In contrast, higher TCs help balance the probabilistic advantage of specialized entry by making

---

6 An analytical treatment of the entry probabilities is available from the authors upon request.
it less profitable in general, therefore increasing the relative frequency of integrated entry. Hence, we end up seeing more integration at these TC levels.

The bottom panel of Figure 1 depicts the profitability of the entire sector – that is, the combined profits of both upstream and downstream segments. It shows that there are no significant differences in total profits as a function of TCs and that for all three levels overall profits decline over time, as more efficient firms enter the industry and natural selection weeds out inefficient ones. Because of decreasing returns to scale and the fact that firms are price takers, they will produce until their (increasing) marginal cost reaches the industry-wide equilibrium sales price. Thus, the point where supply meets demand will be set on the marginal cost curve of the least efficient firm, which will be producing at a level that allows it to barely recoup its fixed costs. Profits are then made inframarginally – i.e., firms always produce the last unit at a break-even (as costs rise with DRS), and may make profits in those units that cost them less to make. Therefore, as the industry evolves and more efficient firms enter, forcing the exit of less efficient ones, the price buffer provided by the least efficient firm shrinks, thus leading to lower profits overall.

However, our interest here is not total sector profits, but the distribution of profitability across segments. This can be seen in Figure 2, which depicts the shares of total profits commanded by specialized suppliers, specialized assemblers, and integrated firms. As one would expect, profits in the sector accrue mostly to specialists when TCs are low and suppliers and assemblers can use the intermediate market to find trading partners efficiently. In contrast, integrated firms dominate the sector with high TCs, as these shield them somewhat from competition by making it more difficult for specialized firms to enter and by burdening specialized incumbents with the TC ‘levy.’

**INSERT FIGURE 2 HERE**
With this background set, we now move on to the ‘ripple’ effects of an upstream innovation. This is modeled initially as a reduction in production cost for upstream entrants, whether specialized firms or parts of potentially integrated entrants. Figure 3 shows how aggregate industry scope and profits change as a result. Figure 3 shows the ‘treatment effect’ – i.e. the (percentage) difference between the sector dynamics with and without the innovation, which happens in period 250. To make these effects comparable, and meaningful, we have provided graphs that show the percentage change from the baseline runs after the treatment, for scope and total profits.

Figure 3 shows that TCs do play a role in shaping aggregate scope and (temporarily) profits after the innovation is introduced. First, we note in the top panel of Figure 3 that the lower the TC, the greater the disintegration of the sector after the innovation. This is consistent with our expectations. Perhaps more interestingly, even though new integrated firms do have access to the innovation and have a cost advantage by virtue of not ‘paying’ TC, they do not seem to be able to fully capitalize on that, and vertical integration decreases even at the highest TC level. This reflects the fact that it is hard for a new, innovative, integrated firm to also be very good downstream, while it is relatively easy for innovative suppliers to take advantage of efficient downstream complementors by contracting out the assembly.\(^7\)

**INSERT FIGURE 3 HERE**

In the bottom panel of Figure 3, we can also see the change in total industry profits. There is an initial spike, as the innovation allows some new firms to benefit from their capabilities (upstream), while still being buffered by (in relative terms) not-so-capable firms. If TCs are low, then the initial spike in profits is high. The main reason for this is that specialists

---

\(^7\) This arises in part because as the industry evolves, incumbent assemblers have been able to increase their average efficiency via selection and learning, whereas integrated entrants draw new downstream capabilities from the original distribution. This result remains even in the absence of learning.
dominate the industry with low TCs, while integrated firms are more numerous when TCs are high. So there are greater profit opportunities when specialized assemblers hook up with suppliers than there are for new integrated firms, who are always tied to their own internal supplier segment.

With time, however, we see that selection and increased competition from new innovative entrants tend to return profits to the level that would have prevailed had the innovation never been introduced. But even when this happens, we will see that there are winners and losers, if we once again disaggregate profits by firm type.

**Profit changes rippling through the value chain after the upstream innovation:**

**Bottlenecks, dumb luck, and industry dynamics**

To consider what underlies the results above, we need to delve some more into the micro-mechanics of industry evolution. To do so, we disaggregate the changes in total profits across each of the constituent segments. The top panel of Figure 4 shows the treatment effect on specialized assemblers. These non-innovating assemblers display ‘dumb luck’ and increase their profits at all TC levels. More interestingly, the relative profit increase that they derive from the upstream innovation is strongly correlated with TCs. These two important effects arise because the downstream assemblers, post-innovation, increasingly become ‘bottlenecks’ – sought after by the ever-improving upstream innovators – and it takes a while for this impact to subside. In steady state, they see an increase of 50–180 percent in profit.

**INSERT FIGURE 4 HERE**

Why should this be? The answer lies largely in the dynamics. When TCs are the highest, aggressive industry selection up to the period where the innovation is introduced weeds out most specialized assemblers, leaving only a handful of highly capable ones. When the shock comes, entry of new specialized assemblers will be limited, not only because high TCs would
favor the entry of an integrated firm in the case of a good downstream draw combined with an average (from the new distribution) upstream one, but also because the capability threshold for entry as a specialized assembler has been significantly raised by natural selection and learning. In other words, the relative entry dynamics put the skills of the downstream assemblers in ‘short supply.’ They become, to use Rosenberg’s (1969) analogy, the ‘bottleneck,’ and since they are a needed complement they have the ability to keep the profits.

Figure 5, which plots the treatment effect on the cumulative entry of firms, clearly shows what happens with entry to the sector after the innovation. With very low TCs, the top panel shows that there are only a few more specialized assemblers entering the industry after the technological shock, and the difference is small when compared to high TCs. This again reflects the fact that potential entrants find it hard to compete with those assemblers who survived until the introduction of the innovation, but might still find it profitable to enter as a supplier or integrated firm depending on TC.

Therefore, higher TCs allow the highly efficient surviving specialized assemblers to benefit from the relatively few innovative specialized suppliers who manage to enter the industry despite the fact that the TC ‘tax’ is high. TCs are not just a cost, but also a barrier that means that the few select specialized firms that survive downstream are efficient – and they can leverage their efficiency by drawing on the equally efficient new supply of upstream partners who must jump this high TC barrier in order to enter. Hence, entry of specialists is asymmetrical: efficient downstream assemblers, having improved still further through learning and production, need not fear the threat from less effective assemblers who do come in to take the extra demand. If anything, these entrants come in and become the ‘umbrella’ – the least effective firms that set the price for that segment high.
In contrast, at very low TC levels, surviving specialized assemblers up to the onset of innovation will be on average less efficient than at high TC levels. And while they will benefit from the significant entry of innovative upstream suppliers (although these still have, on average, lower capabilities than new suppliers at high TC levels), the opportunities offered by the superior upstream technology will not be as good as when TCs are high, given the industry evolution. In other words, the ‘bottleneck’ that they can command is much weaker.

Let us next consider the profits of specialized suppliers. First, we see that the shift in their overall level of profitability over time is lower than that of their downstream co-specialized brethren. Second, we see that they cannot sustain profits, as the innovation forces more competition for entry – the more so with lower TCs (see the middle panel of Figure 5). At the lowest TC level, profits become even lower than they would have been without the innovation, as can be seen by the negative values of the treatment. The innovation ends up intensifying competition, shifting profits away.

The bottom panel in Figure 4 completes our value-chain-level exposition by showing the fate of integrated firms. We can see that integrated firms suffer from this innovation, and all the more so when TCs are low – something we expected when we saw the drop in vertical integration in the sector. Despite the significant amount of entry by new integrated firms shown in Figure 5 when TCs are high, their overall profits slump by an average of 20 percent soon after the shock, before recovering to 10 percent below the long-run no-innovation level. The recovery relates to the stochastic properties of entry – it just takes longer to get two good draws from one firm, as opposed to two firms that get good draws up- or downstream using the market. And the lower TCs are, the bigger this problem is. Furthermore, low TCs mean

---

8 Again, part of the stark reduction in profitability of integrated entrants is due to the dynamics that have preceded the innovation. Recall that the surviving integrated incumbents are relatively efficient up to the time of innovation at this low TC level. Yet the incumbents cannot innovate. So, incumbents face an internal inefficient upstream segment, while the potential integrated entrants can draw on the upstream innovation, yet do not have the benefit of the historical learning curves for their new downstream segment. This places integrated firms at a competitive disadvantage.
that the innovation (which happens upstream) makes the specialized up- and downstream specialist combination a more formidable opponent for the integrated incumbents; so, all in all, they will fare worse.

To synthesize, innovation creates profitability ripples along the value chain. In addition, TCs work as an isolating mechanism, by favoring entry of integrated firms and thus limiting the entry of innovative specialized suppliers who then go on to pair up with highly efficient surviving specialized assemblers. They also impact value in subtler ways in the other segments. When TCs are low, there are a significant number of new entrants, especially upstream (Figure 5). Since price formation in the upstream segment is determined among suppliers themselves, the increased competition brought about by entry ends up homogenizing marginal costs (or capabilities), thus squishing margins and competing the segment’s profits away. So, in the case of low TCs, suppliers’ advantage is more ephemeral, and the ‘bottleneck’ moves downstream, with assemblers increasing their market share at the expense of integrated firms and suppliers.

**Firms in competition: From segment- to firm-level dynamics**

We can supplement this analysis of sector-wide evolution by looking at the underlying drama that happens as firms compete with each other. To see this, we contrast the profits of incumbents (firms in the sector *ad initio*, plus those that enter up to the time of innovation) with those of new entrants, who draw on the superior capability distribution that the innovation made possible. The left-hand panels in Figure 6 below show the total amount of profits made by incumbents – all three types – while the right-hand ones show the profits of firms that enter from the time of the innovation onwards. Note that, unlike in the case of the ‘treatment’ graphs at the sector level, these graphs consider how profit levels change in *absolute* as opposed to relative terms for the different categories of firms.
The top two panels clearly illustrate the ‘bottleneck’ effect on specialized assemblers. With selection driving away inefficient firms until the upstream innovation becomes available, the surviving incumbent assemblers greatly benefit from it, leaving very little opportunity for further entry of similar firms. The middle panels, in contrast, demonstrate how new entrants almost immediately replace incumbent specialized suppliers. They also show how continuous entry at the lowest level of TCs ends up competing away the profitability of the whole segment. Finally, the bottom panels show how the innovation also leads to new integrated firms replacing incumbent ones in the sector. However, this substitution process is much slower than that of specialized suppliers, due to the ability of incumbent integrated firms to cross-subsidize their segments for some time, particularly when TCs are high.\(^9\)

Finally, we should note that we have run another setup, not reported here, where the innovation also originates from within the sector. The main results do remain, and some interesting, and intuitive, differences emerge.\(^10\)

**What if the innovation is dynamic?**

A more radical conception of innovation revolves around its nature, as opposed to its origin.

---

\(^9\) Recall that Figure 6 provides the evidence on the rates of total profitability in the sector in terms of different types of firms. Comparing and contrasting this with the rates of change of profits following the innovation presented in Figure 4 provides us with an additional intuition. While TCs are positively correlated with relative profitability increases in the case of specialists, they are negatively correlated with absolute profit levels. The reason is that when innovation occurs, specialists are fewer, and more capable on average: so they can be more profitable, but the total amount of profits is not large. It’s the percentage that is large! TCs restrict the amount of production that co-specialized pairs of superior assemblers and new suppliers can undertake, especially when competing with integrated firms; and this reduces the total amount of profit that can be achieved. Yet this also protects, by virtue of this restriction, the margins of these suppliers, so that TCs help boost profitability after an innovation shock.

\(^10\) This scenario is closer to the status quo ante than the previous case; the shock is smaller. The reason is that integrated incumbents now have a fair chance to become innovators and compete with disintegrated players on a fairer basis. As for total sector profits, they jump soon after the innovation is introduced. However, with both entrants and incumbents able to access the innovation, competition intensifies more swiftly, and profit streams are shorter-lived. Specialized assemblers still get their “dumb luck” advantage, with profits increasing soon after the innovation (though less so than before) and the more so with higher TCs. Suppliers’ profits behave similarly to those presented in Figure 4, and the isolating effect of TCs is visible, though more limited due to the ability of integrated incumbents to partake in the innovation process. TCs cannot ‘isolate’ as much, and as such matter less for profitability. As incumbents can also innovate, they do so, and profit; the lower TCs are, the better off they become.
In both of setups so far, we modeled innovation as a one-time increase in capabilities (or a decrease in costs). This is akin to a change in what Winter (2003) calls a ‘zero-level capability.’ Yet we know that much of the action takes place around higher-order capabilities. We consider the simplest, which is the effectiveness of the learning process – technically speaking, the gradient of the learning curve (Argote & Epple, 1990), or the efficacy with which cumulative production experience translates into a reduction in costs. We can then explore the impacts of a change in this dynamic capability upstream, and see how this affects the dynamics downstream, and how TCs moderate this relationship. To motivate this further, we consider the case of ‘competence-enhancing’ innovations, positing that those who innovate will have both superior upstream capability and a better gradient to their learning curve. So the question here is how a change to these dynamic abilities would change the results of the model. To explore this, we allowed innovators’ upstream learning rates (for both new specialized suppliers and integrated firms) to become three times their pre-shock values. To keep the comparison with our initial setup and Figure 4 simple, we have provided the results when innovation is available only to entrants. The results are summarized in Figure 7.

A comparison with Figure 4 suggests that these results do not differ qualitatively in terms of total scope, yet there is an important change in terms of total profits. Initially, profits increase in a familiar fashion. However, instead of returning to the pre-shock levels in the long run through competition, profits remain higher than in the absence of the shock. This happens because increased learning accentuates capability heterogeneity, leading to a greater concentration of production in the hands of leaders. In other words, an innovation that also includes a higher-order capability change makes leaders even stronger, even more dominant, and this means that entrants do not have the ability to catch up, induce competition, and
reduce the competitive profits. So the type of innovation has real implications in terms of profitability.

But what of the ripples down the value chain? What happens when we disaggregate and consider the profits of different segments? Assemblers still get their fair share of dumb luck, with profits increasing for all TC levels, and the more so the higher the TCs, which still act as isolating mechanisms. As for suppliers, we can see that learning brings a particularly important effect to specialized suppliers. First, their overall profits increase on average more than those of their co-specialized downstream counterparts — the opposite of what happened when the innovation consisted only of a zero-order capability improvement. When an innovation is in terms of dynamic capabilities, the innovating segment has more of a chance to protect its own advantage and not pass it through the value chain so that the ‘dumb luck’ complementors end up benefiting more than the innovators themselves. Comparing Figure 7 to Figure 4 shows that higher learning leads to a positive treatment of suppliers’ profits even for the lowest TC level. This is consistent with the previous analysis, which suggests that higher-level capability innovations enhance and sustain efficiency profits, as they reduce the threat of effective competition for the best firms: In the baseline model, too much entry of innovative specialized suppliers ended up competing away their profits and transferring the profitability of the innovation to the ‘bottleneck’ represented by assemblers. Plotting the treatment of supplier entry in the last panel of Figure 8 and comparing it to Figure 5, we see that higher learning significantly restricts entry by making the threshold that must be overcome by the upstream capability draw much higher in a shorter time span. This in turn helps to isolate the new suppliers from further competitive pressures that would have resulted from further entry upstream. Combined with the higher profit opportunities now available, the ultimate result is that suppliers’ profits increase by approximately 100 percent, instead of decreasing when TCs are very low.
Finally, results also change in an interesting and subtle way for integrated firms. As in Figure 4, we have an initial drop in profitability, the more so with lower TCs. Yet as new integrated firms enter, some (the best of them) are also ‘turbo-boosted’ by their upstream capability enhancement – just as the specialized upstream suppliers are. And, given sufficiently high TCs to act as a protection, some of the integrated firms are able to recover lost ground and have their joint relative profitability move into positive territory.\textsuperscript{11}

**DISCUSSION**

Beginning with the aim of exploring how an innovation ripples through the value chain, we restrict our attention to the case of an upstream innovation, and a competitive setup. We create a model with upstream suppliers, downstream assemblers, and integrated firms, and consider different TC levels to discover their impact in the profit evolution process. This well-circumscribed exercise yields both some general findings, and some subtler conditional results.

In terms of general findings, we show how the downstream assemblers who did not engage in any innovation ended up capturing a substantial part of the total sector profitability (at least in relative terms), as a result of ‘dumb luck’ – simply being in the right place at the right time. In some of our illustrations, these assemblers even had the majority of the additional profits. In other words, we see how, in the absence of any strategic interaction, an upstream innovation caused more entry and mobility in that segment, shifting profits downstream and turning assemblers into a ‘bottleneck’ (Jacobides et al., 2006; Rosenberg, 1969).

In addition, we show how different types of innovation (zero-level vs. higher-level

\textsuperscript{11} Interestingly, the existence of a dynamic ‘boost’ to innovators means that profits can go up even if the market share does not increase. The initial drop in profits at the highest TC level is the result of competition against efficient surviving assemblers. Given greater opportunity for entry of integrated firms at this level, and the higher learning that accompanies them, new profit opportunities are created. And these are not created at the expense of specialists. If we look at market shares for this TC level (results omitted), we will see that integrated firms actually lose market share, despite increasing their profits like specialized firms.
capabilities such as learning) affect the competitive outcome. First, we demonstrate that higher-level capability-based innovations dynamically increase heterogeneity and also help sustain higher total efficiency (competitive) profits. Second, we show that innovations on the basis of dynamic capabilities favor the innovating segment in relative terms, and reduce the extent of entry (which is precisely why the segment can hang on to its profits).

We have also illustrated the importance of TCs in the process of profit migration. We find that TCs can affect not only total profit, but also who profits, and when. TCs can act as an ‘isolating mechanism,’ which, regardless of the exact setup of innovation dynamics, represents not only a cost but also a shield against competition, which can lead to profits and protect them from rapid erosion. This happens not because of uncertain imitability (as in Lippman and Rumelt (1982)), but rather because of their role in competitive dynamics. Furthermore, TCs also affect the inter-temporal profit pattern, softening the onset of the profit ‘spike’ and delaying its decay. This is particularly important when we look at the NPV of profits, and questions of discounting through time and capitalizing at different points in time emerge as considerations.

Our paper thus extends work on scope, the co-evolution of (heterogeneous) capabilities, and TCs to encompass profitability dynamics. While industry-friendly simulations such as Malerba et al. (2008) have considered how scope and concentration, or scope and technologies, co-evolve, the analysis of profitability has not been explicitly considered. We think that the explicit focus on profit evolution in a sector is a promising theme, and one that complements recent empirical work, such as Jacobides and Tae’s (2012) analysis of profitability dynamics along the computing value chain.

That said, the core contribution of our paper is to explain the mechanisms through which changes in one part of the value chain can ‘ripple through.’ We explain why it is downstream assemblers who can benefit most from a cost-saving innovation in the upstream sector. The
intuition is that competition can make them the ‘bottleneck’ – which leads to higher profits, as hypothesized by Jacobides et al. (2006). In our model, innovation allows new firms to enter a particular part of the value chain with superior capabilities and displace existing players. This result is consistent with several hypotheses in the emerging ‘industry architecture’ literature (Brusoni, Jacobides and Prencipe, 2009; Ferraro and Gurses, 2009; Pisano and Teece, 2007; Wolter and Veloso, 2008), which looks at conditions of replaceability and relative entry; some connections might also be made to Porter (1980), although only superficially, since his line of research focuses almost exclusively on oligopolistic profits.

Our analysis sheds new light on the mainstream discourse on profit migration, a familiar feature of the popular business press (Slywotzky, 1996). Christensen, Raynor, and Verlinden (2001), for instance, exhort firms to ‘skate to where the money will be,’ suggesting that profits will shift from integrated to disintegrated firms as a result of the dissipation of integrated firms’ differentiability (see also Christensen, Verlinden, and Westerman (2002)). In a similar vein, Christensen and Raynor (2003: 168) propose a ‘law of conservation of attractive profits,’ suggesting that when commoditization and modularization drive profits from one part of the value chain, they will re-emerge in another. Such efforts arguably lack truly robust theoretical or empirical support – but, if nothing else, they testify to the importance of predicting or explaining how conditions in one part of the value chain affect profits. Our discussion of how, in a competitive setup with DRS, profits shift (and why) offers a concrete and analytically consistent discussion of these interesting but elusive phenomena.

This approach could be useful in a variety of contexts. From workstations and microprocessors to innovations in loan origination and impacts on loan processing firms, examples abound of upstream innovations potentially bestowing ‘dumb luck’ profits on
downstream firms. Of course, the beneficiaries of such ‘dumb luck’ are not really ‘dumb’; they are (as in our model) highly efficient firms. Yet they tend to ascribe their success solely to their own merits, rather than sector-wide dynamics. As an example, consider the case of the offshore drilling value chain (Ernst & Young, 2012). Years of innovations in drilling technologies have allowed oil to be extracted, albeit riskily, from great depths. The exploration- and production-drilling segment, where much of this activity was located, showed some increase in profitability, with EBIDTA margins in the neighborhood of 11–12 percent. However, the main beneficiaries were at the very end of the value chain: the seismic analysis and reservoir prospecting firms who identify deep-drilling areas. Their margin went as high as 24 percent (before falling, probably temporarily, as a result of the Macondo/BP spill). So, while the seismic/reservoir firms continued to improve and innovate, it was drilling technology advances that arguably ‘lifted all boats’ and added a ‘dumb luck’ component to their profitability.

Another, more complicated sector where we think our model offers fresh, if partial insights is the telecommunications equipment market. For example, contrast Apple, which has been vertically focused on the device, leaving the development of applications to outside firms, with Nokia, which has kept the system much more integrated. Apple benefits from innovations in applications: Over and above the hefty 30 percent commission it obtains from app sales, it also benefits from the excitement created by both the device and the apps. While this case is far more complicated and strategic than our model, the intuition remains: Excess entry of innovators (in apps), while beneficial to a few of these innovators, tends to benefit downstream complementors (Apple) when they become the bottlenecks. In contrast, Nokia, whose technology has benefited much less from inputs by outside firms, has suffered from an inability to innovate effectively, and as a result has come off worst in the battle with Apple.

Furthermore, the same TCs that protected Nokia before innovation in apps transformed
the use of the phone have now been the source of its downfall. What used to be isolating mechanisms became the barriers of adaptation in an evolving world of technological discontinuity.

More broadly, we can consider how the insights from this model on value migration might help us understand when and why firms use ‘open’ innovation strategies, since they expect benefits to accrue to them inasmuch as they are the bottlenecks in the vertically related segment. From Google to P&G, the use of such selective processes along the value chain seems to relate to the desire to capture value, not just an enthusiasm for openness.\(^\text{12}\)

**Methodological choices and limitations of our approach**

Our paper provides a simple, stylized rendition of the conditions that lead to profit migration along the value chain. Like any model, ours is open to the charge of being overly stylized. We do not purport to offer a comprehensive map of an evolving value chain, nor a ‘history-friendly’ depiction of the main forces behind a sector’s evolution (per Malerba *et al.*, 2008). Rather, we aim at retaining the smallest number of analytical foundations required to illustrate a particular set of phenomena. As a result, we consider the competitive setup, heterogeneous firms, varying TCs, and well-specified dynamics of entry and exit. We abstract away from issues of product differentiability, monopolistic control, network effects, platform dynamics, vertical restraints, contractual strategies, appropriability regime differences, and strategizing to shape sectors as considered in the industry architecture literature. We aimed for a model that combined verisimilitude (i.e. not too distant from

\(^\text{12}\) As a final and highly speculative and macro application of our model, consider the explosive growth and innovation (in terms of cost containment) in developing countries such as the BRICs. It may be that the massive improvements in the upstream segment, done in China and India, both led to significant vertical disintegration, and also created the illusion of success (temporarily) to the first-world downstream partners. That is, in the initial period, those firms that were vertically specialized downstream benefitted from the upstream innovation and cost reduction. They had a form of ‘dumb luck’ as a result of the BRIC innovators efforts. Yet as the competitive process unfolded, those ‘dumb luck’ assemblers from the first world found themselves with profits diminishing, and BRICS started entering downstream as well. Making matters even worse, it seems that the innovation in the BRICs has shifted from zero-order to dynamic capabilities, since their abilities are now in learning and innovating as much as production efficiency (Kumar & Puranam, 2012). This means that these ‘upstream’ BRIC firms will benefit more, and the hollow downstream complementors from the first world might not have the same luck in the next few years – an ominous sign for a declining Western world.
reality) and parsimony (not too complex or ad hoc).

In terms of verisimilitude, the key point to note is the emphasis on non-oligopolistic structures. We feel that this is both empirically important and intellectually defensible. In terms of the empirics, the growing literature on firm performance differentials strongly suggests that there are substantial, persistent efficiency differences between firms (see Bloom and Van Reenen, 2007; Gibbons et al, 2010; Jacobides and Winter, 2012). This, in turn, suggests that a good part of profitability may be due to efficiency differences rather than oligopolistic restraints or other forms of strategic behavior. Intellectually, the underlying analytics are justified because we know far less about the evolution of competitive profits than we do about multilateral strategizing. In terms of parsimony, it is noteworthy that we offer a ‘quasi-experimental’ computational model as opposed to an analytical, closed-form calculation. This was done in order to represent a set of heterogeneous competitors and their evolution through time, and to model innovation as a change in the underlying process of knowledge generation – none of which can be effectively encapsulated in a closed-end formulation.

In terms of limitations, the results of our dynamics depend on our assumptions on entry and production, even if our approach was to show exactly how these choices shape collective outcomes. While our structure clearly has limits (including the fact that it does not allow for oligopolistic price-setting and strategic interaction or shaping of the value chain), we hope that our simple case can add value by dint of its detailed reporting of causal forces. Another limitation is the use of a DRS formulation, which means that the heterogeneity of firms’ capabilities is associated with a lack of replicability. This is a fairly frequent case, but an alternative approach would be to model a production function that exhibits short-term Constant Returns to Scale, but where DRS can emerge over the long term – provided firms invest in customizing their resources (as in Jacobides et al. (2012)) or training and
development (Rubin, 1973). To ensure that DRS were not absolutely crucial in driving the results, we did try CRS with an ‘investment module’ that limited the growth of the best firms, as in Jacobides et al. (2012).

Many extensions can be presented. Several have been explored but not yet reported, to conserve space. For instance, rather than taking heterogeneity distribution as a given (and varying it as a robustness check), we can alter it as an independent variable to see the profit consequences of initial heterogeneity on evolutionary dynamics.\textsuperscript{13} We also explored the effect of (positively or negatively) correlated capabilities coming into existence only after the technological shock. On the basis of the results we have already seen, we confirm that our findings are robust.

Summing up, we hope that the paper has provided an interesting glimpse of the systemic impacts of co-evolving changes in profitability and scope across a sector. We have allowed for differentially capable firms, an empirical regularity often absent in formal modeling, and focused on the competitive case. Based on these primers, we tried to push our intuition on how an industrial system evolves (in the spirit of Malerba et al. (2008)), focusing on differentially capable firms, transaction costs, and the impacts on profits and scope from an upstream innovation. We hope that this provides a concrete sense of how innovation can shift a ‘bottleneck’ from one part of the value chain to the next, and how TCs combine with the relative entry/production conditions to shape not only vertical scope but also profitability and segment profits. With the ever-increasing number of closely connected systems of vertically co-specialized firms (in industries ranging from telecoms to financial services to healthcare), a better understanding of the way profit ‘ripples’ along a value chain is essential.

\textsuperscript{13} In particular, we already considered a setup where the capabilities of the upstream and downstream segments draw from a positively correlated set of capability distributions, where success in one part drives success in another. As expected, positive correlated capabilities make it easier for integrated firms to enter at the expense of specialized ones. Integrated firms can profit more, and especially so at high TC levels. But the central claim of the paper, i.e. that the profit will move downstream to the assemblers after an upstream innovation, is preserved in the correlated capability context.
REFERENCES


Lieberman, MB. 1989. The Learning curve, technology barriers to entry, and competitive
survival in the chemical processing industries. Strategic Management Journal 10(5):
431–447.

Lippman, SA, Rumelt, RP. 1982. Uncertain imitatibility: An analysis of interfim differences

Lippman, SA, Rumelt, RP. 2003. A bargaining perspective on resource advantage. Strategic

of computer firms: A history-friendly model of the coevolution of the computer and

Murray, F. 2004. The role of inventors in knowledge transfer: Sharing in the laboratory life.

Perry, MK. 1989. Vertical integration: Determinants and effects. In Handbook of Industrial

Pisano, G, Teece, DJ. 2007. How to capture value from innovation: Shaping intellectual


London, UK.

Rosenberg, N. 1969. The direction of technological change: Inducement mechanisms and


Slywotzky, A. 1996. Value Migration: How To Think Several Moves Ahead Of The

326–365.

Teece, DJ. 1986. Profiting from technological innovation: Implications for integration,


Tripsas, M. 1997. Unraveling the process of creative destruction: Complementary assets and
incumbent survival in the typesetter industry’. Strategic Management Journal, 18(S1): 119–
142.


Winter, SG. 1995. Four Rs of profitability: Rents, resources, routines and replication. In
Resource-Based and Evolutionary Theories of the Firm: Towards a Synthesis. Kluwer
Academic Press: Boston, MA.


Wolter, C, Veloso, FM. 2008. The effects of innovation on vertical structure: Perspectives on
Figure 1: Evolution of industry scope and total profits.

Figure 2: Evolution of profit shares by firm type.
Figure 3: Treatment of industry scope and total profits.

Figure 4: Treatment of profits by firm type.
Figure 5: Treatment of the cumulative number of entrants.

Figure 6: Profits by cohort (incumbents and new entrants) for each firm type.
Figure 7: Results for scenario where innovation is introduced by new entrants and is accompanied by higher learning rates.
APPENDIX: Model Formulation, Parameterization, and Robustness Checks

Model structure: Overview

Our model has two parts: A static phase and a dynamic phase. In the static part, all companies compete with their capacities in place. They decide whether they will produce, and, if so, how much; they buy resources to produce; and, if they are specialized firms, they trade through the intermediate market. The entire static/short-term market equilibrates, and then firms count their profits (or losses) and make decisions for the next period. As the period ends, firms consider whether they can afford to continue operating in the next period; if they can’t, they exit. Then follows an ‘evolutionary update,’ in which potential entrants ‘try their luck’ by drawing on a capability distribution and consider whether, with the capability level they have drawn, it makes sense for them to compete – and, if so, whether they should compete as integrated or specialized firms, endogenously choosing their preferred part of the value chain. With both phases complete, the process begins afresh with a new static period.

Static structure: How competition happens in each period

For the static model, we build on Jacobides (2008) and provide a computational general equilibrium (CGE) model of the evolution of a sector’s vertical scope, as well as the profitability of each participating firm. The sector is comprised of two vertically linked segments, one upstream and one downstream. Downstream firms compete in a final market to sell a homogenous product to price-sensitive consumers. Firms in this segment can be either specialized assemblers or integrated firms. They have varying degrees of ‘zero-order’ capability (per Winter’s (2003) definition) and extend their distinctive approaches to the competitive struggle.

To manufacture the final consumer product, an intermediate homogenous product (or component) is needed. Specialized assemblers buy this component from specialized suppliers who compete to sell it in an intermediate market. Integrated firms, in contrast, manufacture the intermediate product in their own upstream segment and cannot make use of the intermediate market – i.e. there is no tapered integration. In addition, firms can neither integrate (e.g., via a merger or acquisition), nor disintegrate (e.g., via a divestiture or spin-off) throughout the industry evolution.

The model also allows for two independent markets for resources, an intermediate and a final one, in which firms procure two types of homogenous resources needed for the production of intermediate and final goods, respectively. These resources are in elastic supply to both segments of the industry. Production functions display DRS, a feature empirically observed in many sectors (Basu and Fernald, 1997).

Specialized firms trading in the intermediate market must pay per valorem TCs, while integrated firms avoid paying this cost by procuring the intermediate product in-house. Firms

---

14 Since we use Jacobides’ (2008) model as a starting point, we will retain whenever possible the same nomenclature used. Note that our model differs from Jacobides first, in including entry and exit; and, second, in having a fundamentally different research design. Here, we artificially restrict firms to being assemblers, suppliers, or integrated (so as to be able to track the implications of innovation in these three different ‘firm types’. Our research design is also different. We examine a small number of TC values and consider how the industries evolve with and without an exogenous upstream shock in productivity. We also track not only scope, as a function of the market share of integrated vs. specialized firms, but also profits.

15 Since specialized assemblers can buy the intermediate component from any specialized supplier, these firm pairs represent an open modular system. In contrast, integrated firms, by being tied to their upstream supplier, represent a closed modular system.

16 Although these actions will be considered in future extensions, they are not expected to alter the outcomes of the model significantly, due to the flexibility provided by the entry and exit mechanisms. For example, a divestiture can be modeled as the exit of an integrated firm and the entry of two specialized ones.
also have a fixed production cost per vertical segment.

There are no capacity constraints in the model. As firms respond to price signals by growing or shrinking, there are no limits imposed on how much they can adjust their production from one period to the next. Thus, the extension of capacity by more capable firms is the central mechanism driving the Darwinian selection.

Operationally, firms maximize profits in any given period. The Karush-Kuhn-Tucker conditions from these constrained optimization problems, together with the market-clearing equations, define a mixed complementarity problem that can be solved in GAMS via the PATH algorithm (Dirkse and Ferris 1996; Ferris, Kanzow, and Munson 1999; Ferris and Munson, 2000).\(^\text{17}\)

The simulation begins with an equal number of integrated firms and specialized pairs (10 specialized suppliers, 10 assemblers and 10 integrated), each endowed with upstream and downstream capabilities \(a_{I,i}\) and \(a_{F,i}\), respectively, drawn independently from a uniform distribution defined in the interval \([0.9, 1.1]\). This guarantees that capabilities are uncorrelated between both segments of integrated firms. Given this initial heterogeneity draw, the optimization is solved, so that quantities produced, resources consumed, and industry-wide prices are determined.

Hence, we have firms behaving optimally within this first (and any subsequent) period, but also myopically in relation to any future period (i.e. without any forecasting or strategizing). Although this optimal behavior departs in principle from the simpler routines advanced by evolutionary models, it can also be seen as just another rule based on the signals provided by industry-wide equilibrium prices.

Mathematically, we have a profit maximization problem, where the objective functions for specialized suppliers, specialized assemblers, and integrated firms are given, respectively, by:

\[
\begin{align*}
\max_{RI_i} \Pi_{su,i} &= QIS_i \times PIM((1-TC) - RI_i \times PRI - FC), \\
\max_{RF_i} \Pi_{as,i} &= QF_i \times PF - QIB_i \times PIM(1+TC) - RF_i \times PRF - FC, \\
\max_{RI_i,RF_i} \Pi_{in,i} &= QF_i \times PF - RI_i \times PRI - RF_i \times PRF - 2FC,
\end{align*}
\]

where \(QF_i \geq 0\) is the quantity of the final good produced, \(PF\) is its industry-wide price, \(QIS_i \geq 0\) is the quantity of the intermediate component sold by suppliers, \(QIB_i \geq 0\) is the quantity of components bought by assemblers, \(PIM\) is the price of the component in the intermediate market, \(RI_i > 0\) and \(RF_i > 0\) are, respectively, the quantities of intermediate and final resources used by the firms, and \(PRI\) and \(PRF\) are the industry-wide prices of these resources used in the upstream and downstream manufacturing processes. Finally, \(TC\) represents per valorem transaction costs that subtract from or add to the price of the intermediate product sold by suppliers and bought by assemblers. \(FC\) is the ordinary production fixed cost for each segment, and has the desirable property, combined with the diminishing returns to scale we note below in equations (2a) and (2b), to give us a production function that initially yields economies of scale, and then, beyond a certain level, diseconomies of scale.

The quantities produced are determined by firms’ intermediate and final production functions:

\(^{17}\) The implementation code in GAMS is available from the authors upon request.
\begin{align*}
Q_{IP_i} &= a_I \times R_{I_i}, \\
Q_{F_i} &= a_F \times R_{F_i},
\end{align*}

where \( Q_{IP_i} \) is the quantity of components manufactured by suppliers or by the upstream segments of integrated firms, \( a_I \geq 0 \) and \( a_F \geq 0 \) are the firm-specific zero-order upstream and downstream capabilities, and \( b_I \geq 0 \) and \( b_F \geq 0 \) the respective elasticities of production (or returns to scale).

Moreover, since integrated firms can neither buy components in the intermediate market nor sell any excess capacity, the production of specialized suppliers, assemblers and integrated firms must satisfy, respectively,
\begin{align*}
Q_{IS_i} &= Q_{IP_i}, \\
Q_{F_i} &= Q_{IB_i}, \\
Q_{F_i} &= Q_{IP_i},
\end{align*}

which simply imply that specialized suppliers sell all of their production in the intermediate market, and also that the Leontieff coefficient is normalized to 1, so one intermediate component only is used in the assembly of the final product.

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:
\begin{align*}
S(PRI) &= SI \times PRI^{es} \geq \sum_i R_{I_i}, \\
S(PRF) &= SF \times PRF^{es} \geq \sum_i R_{F_i},
\end{align*}

where \( SI \) and \( SF \) are constants and the exponents represent the elasticities of supply for both resources. Similarly, the prices for the intermediate and final products are obtained from:
\begin{align*}
\sum_i Q_{IS_i} &= \sum_i Q_{IB_i}, \\
D(PF) &= DF \times PF^{es} \geq \sum_i Q_{F_i},
\end{align*}

the latter of which represents the exogenous and elastic consumer demand function where \( \varepsilon_F \) is its elasticity and \( DF \) a simple constant.

**Evolutionary update: Entry, exit, and capabilities**

After the first (and each subsequent) period when the static optimization is solved, a dynamic updating occurs. First, firm-level capabilities in both segments increase due to learning-by-doing.\(^{18}\) Learning as a function of cumulative output takes place according to:
\begin{align*}
a_{I_i}(t) &= a_{I_i}(1) \left[ 1 + c_I \times \ln \left( \sum_i Q_{IP_i}(t) \right) \right],
\end{align*}

---

\(^{18}\) Learning is not necessary in this model, but it helps to make it more stable and to attain convergence with a smaller number of runs. It is also an empirically plausible stipulation; see Argote and Epple (1990).
\[ aF(t) = aF(1) \left[ 1 + cF \times \ln \left( \sum QF(t) \right) \right], \tag{6b} \]

where \( cI \) and \( cF \) define the learning rates for the upstream and downstream segments, respectively.

Then, unprofitable firms exit the industry.\(^{19}\) To consider whether they will make profits or losses in the next period, firms assume that prices will remain the same, and consider how gradual departures might affect their plight. We implement a model that aspires to accommodate firms’ adaptive expectations on the basis of exit. We consider that the firm with the highest losses (if losses are made in the sector) will exit; and then in order to consider the remaining firms’ reactions, we allow them to take into account the fact that exit will occur. To do so, we analytically compute the equilibrium without the firm that exited (taking into account the updated capabilities due to learning as we do so) and then identify the next firm with the highest losses and allow it to exit, until the sector appears to be in the black (in these trial runs). This process is repeated until only firms with positive forecasted profits remain in the industry.\(^{20}\) This process of ‘capacity check’ is, we feel, justified in behavioral terms, as it corresponds with firms’ ability to assess whether their presence in a sector will be justified on the basis of their production costs and price trend.

Following the capability update and firms exiting the industry, entrepreneurial entry (of one firm per period) takes place.\(^{21}\) This is based on the following mechanism: a potential entrant takes two independent capability draws (one for \( aI_i \) and one for \( aF_i \)) from the same uniform distribution that originated the initial industry heterogeneity at \( t = 0 \). This potential entrant then makes three forecasts about its possible profitability if it should decide to enter as a supplier, an assembler, or an integrated firm, given the current state of incumbent competition. In other words, three new equilibriums are calculated, one for a new upstream entrant with \( aI_i \), one for a new downstream entrant with \( aF_i \) and one with a new integrated entrant with both \( aI_i \) and \( aF_i \). If one or more of these new equilibria results in the new entrant being profitable, it will then enter. The entrant will choose the scope that yields the largest profits according to their estimates, and will stick with this ‘form’ for the duration of its life. Finally, if all 3 profit forecasts are all negative, no entry will occur. With the choices of entry made, the evolutionary update is full, and with the modified roster of participants (some having exited, another entered) the next period is run, and the process is repeated.

**Evolutionary update: Innovation**

In our first scenario, we consider the impact of introducing a cost-reducing technology upstream, modeled as an increase in the mean of the distribution from which potential entrants draw their initial upstream capability: at the mid-point in the simulation, potential entrants draw substantially higher values of \( aI_i \), thus representing a significant cost reduction in the upstream segment.

So, from \( t = 1 \) to \( t = 250 \), potential entrants independently draw \( aI_i, aF_i \) from a uniform distribution that is centered at 1 and draws from the interval \([0.9, 1.1]\). When the innovation comes into place at \( t = 251 \), \( aF_i \) is still drawn from a uniform distribution with identical variance, but \( aI_i \) now becomes centered at 100 percent of the highest existing value of \( aI \) in

---

\(^{19}\) Due to their fixed costs, firms will have operational losses even if they produce nothing. This is different from Jacobides’ (2008) model, where the absence of fixed costs allows firms to remain idle without actually exiting the industry.

\(^{20}\) These forecasts do not take into account the possibility of entry, although we have also run the process with entry first and exit next, and the results are robust to whether exit happens before or after entry.

\(^{21}\) Note that we have also run the model with multiple entrants for robustness’s sake and the qualitative results are not dependent on this assumption. Also, note the model is run for 500 periods – which eventually allows for ample entry.
the previous period \(t = 250\), among all firms. That is, the innovation means that the mean of the distribution shifts ‘to the right,’ so that last period’s ‘most effective player’ is now but the mean, and that substantially more effective firms can now emerge. This means that the distribution of the capability function becomes \([\max a_I = 250 - 0.1, \max a_I = 250 + 0.1]\).

In our second innovation setup, which addresses the dynamic capabilities approach, relies on our baseline innovation approach but with an additional feature: a post-shock upstream learning rate for new entrants that is three times the downstream one (and also the pre-shock one for both segments).

Table A1 summarizes all of the model’s variables and parameters.

**Parameterization, robustness checks and experimental design**

Significant effort was made to guarantee that the results presented were not borne of a particular combination of parametric values or modeling choices. So a number of changes were made to verify the model’s robustness, as summarized in Table A2.

In terms of the static model, we also considered TCs as a fixed cost as opposed to per valorem taxes. We also combined the per valorem TC formulation with the ‘fixed cost’ of transacting through the market, and this ‘two-part tariff’ formulation yielded very similar results to the ones we present here.

**Table A1: Model glossary**

<table>
<thead>
<tr>
<th>Single-period variables and parameters</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Pi_{su}, \Pi_{as}, \Pi_{in})</td>
<td>Firm Profits</td>
</tr>
<tr>
<td>(QF_i, QIP_i)</td>
<td>Output of upstream and downstream segments</td>
</tr>
<tr>
<td>(QIS_i, QIB_i)</td>
<td>Quantities of intermediate product sold by suppliers and bought by assemblers</td>
</tr>
<tr>
<td>(RI_i, RF_i)</td>
<td>Quantities of intermediate and final resources consumed</td>
</tr>
<tr>
<td>(PIM, PF)</td>
<td>Prices of intermediate and final goods</td>
</tr>
<tr>
<td>(PRI, PRF)</td>
<td>Prices of intermediate and final resources</td>
</tr>
<tr>
<td>(TC)</td>
<td>Per valorem TCs paid by specialized firms</td>
</tr>
<tr>
<td>(FC)</td>
<td>Fixed production costs</td>
</tr>
<tr>
<td>(aI_i, aF_i)</td>
<td>Upstream and downstream productive capabilities</td>
</tr>
<tr>
<td>(bI, bF)</td>
<td>Elasticities of production (returns to scale)</td>
</tr>
<tr>
<td>(SI, SF)</td>
<td>Resource supply constants</td>
</tr>
<tr>
<td>(\varepsilon SI, \varepsilon SF)</td>
<td>Elasticities of supply for resources</td>
</tr>
<tr>
<td>(DF)</td>
<td>Demand constant</td>
</tr>
<tr>
<td>(\varepsilon F)</td>
<td>Elasticity of demand for final product</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dynamic update variables and parameters</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(cI, cF)</td>
<td>Upstream and downstream learning rates</td>
</tr>
</tbody>
</table>

In terms of the dynamics, we also ran the model without a learning component. In terms of the change in the roster of industry participants, instead of a sequential exit mechanism, a simpler one where all firms with operational losses exit at once was implemented. We switched the order in which entry and exit occur in the model, allowing the latter to take place after the former. We also changed the implementation of the modular innovation by limiting access to the innovation only to specialized suppliers, as opposed to suppliers and
integrated firms. And multiple entries were also allowed to take place, as opposed to one firm per period. In all cases, the results remained qualitatively the same as the ones to be presented. No major changes in the overall dynamic evolution of the sector were observed.

All simulations start with the same initial conditions in terms of capability heterogeneity among firms. But as the model has a stochastic component in the capability draws, each run generates a unique evolutionary path. Hence, for each of the three relative transaction cost levels, we run 30 simulations with different random seeds. The resulting output is then averaged and its standard deviation determined before being presented in graphical form. Since there are 500 periods with three additional optimizations solved per period to account for entry, each full model simulation results in $3 \times 30 \times 500 \times 3 = 135,000$ MCP problems being solved. This number was expanded substantially, as calculations that firms make with regards to their exit also require additional MCP solutions within the simulation to take place.

In the model, production functions are defined by the following elasticities of production: $bI = bF = 0.9$. The supply of resources and final product demand are characterized, respectively, by $\varepsilon_{SI} = \varepsilon_{SF} = -\varepsilon_F = 1$. The supply and demand constants take the values: $SI = SF = 200$ and $DF = 400$. Learning coefficients are set at 0.005. Fixed costs are set at $FC = 0.5$, while transaction costs have three levels $TC = 0.001, 0.015$ and 0.030.

The full parametric robustness checks performed are described in Table A2.

**Table A2: Parametric choices and robustness checks.**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
<th>Range</th>
<th>Result</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elasticity of demand</td>
<td>-1</td>
<td>-0.8 to -1.2</td>
<td>No change</td>
<td></td>
</tr>
<tr>
<td>Elasticities of supply</td>
<td>1</td>
<td>0.8 to 1.2</td>
<td>No change</td>
<td></td>
</tr>
<tr>
<td>Elasticities of production (returns to scale)</td>
<td>0.9</td>
<td>0.7 to 1</td>
<td>No change</td>
<td>CRS implementation follows Jacobides et al. (2012)</td>
</tr>
<tr>
<td>Capability distribution</td>
<td>$U \sim [0.9,1.1]$</td>
<td>$U \sim [0.7,1.3]$ to $U \sim [0.95,1.05]$</td>
<td>No change</td>
<td>Normal distributions with same variances also tried</td>
</tr>
<tr>
<td>Learning coefficients</td>
<td>0.005</td>
<td>0 to 0.020</td>
<td>No change</td>
<td></td>
</tr>
<tr>
<td>Exit</td>
<td>Sequential</td>
<td>Firms with $\Pi &lt; 0$ exit at once</td>
<td>No change</td>
<td></td>
</tr>
<tr>
<td>Entry/exit order</td>
<td>Exit then entry</td>
<td>Entry then exit</td>
<td>No change</td>
<td></td>
</tr>
<tr>
<td>Transaction costs</td>
<td><em>Per valorem</em></td>
<td>Fixed and ‘two-part tariff’</td>
<td>No change</td>
<td></td>
</tr>
<tr>
<td>Innovation</td>
<td>Available to integrated firms</td>
<td>Available to suppliers only</td>
<td>No change</td>
<td></td>
</tr>
<tr>
<td>Entry</td>
<td>1 firm per period</td>
<td>Multiple firms per period</td>
<td>No change</td>
<td>Scope moves faster to extremes</td>
</tr>
</tbody>
</table>

22 Setting $\varepsilon_F = -1$ implies that total revenues are constant (and equal to $DF$), thus playing no part in driving the dynamics. However, we have varied this parameter, and considered both elastic and inelastic demand conditions, to ensure our results are robust.

23 Most of these values follow from Jacobides (2008). These are chosen so as to calibrate the model to start at a production level that is about a third from its long-run equilibrium – a fairly realistic structure for an early part of an industry’s lifecycle.