Industrial and Corporate Change, Volume 11, Number 5, pp. 955–993

Disruption, disintegration and the dissipation of differentiability

Clayton M. Christensen, Matt Verlinden and George Westerman

This paper proposes a deductively derived model to help managers who preside over decisions to integrate or outsource to assess *ex ante* whether, when and why it might be strategically and competitively important to develop internal capabilities to perform certain activities in-house, and when it would be sensible and safe to outsource elements of value-added. Among the paper's conclusions are that the competitive advantage from vertical integration is strongest in tiers of the market where customers are under-served by the functionality or performance available from products in the market. Vertical integration tends to be a disadvantage when customers are over-served by the functionality available from products in the market. Vertically integrated firms will therefore often dominate in the most demanding tiers of markets that have grown to substantial size, while a horizontally stratified, or disintegrated, industry structure will often be the dominant business model in the tiers of the market that are less demanding of functionality.

1. Introduction

Whether to become or remain vertically integrated is a question of vast strategic importance in many industries. In recent years, firms such as Alcoa, Lucent and General Motors, for whom vertical control over most steps in their value chains had historically constituted an important basis of competitive advantage, have sold upstream businesses that produced components or intermediate materials, in order to focus on the portions of their value chains that they consider to be core to their business. Others, like IBM, continue to own but are de-coupling upstream from downstream operations, tasking the former to sell components openly in the market, and the latter to procure components from external suppliers when necessary to maintain competitiveness. In contrast, Microsoft is aggressively integrating downstream from its initial operating system products into a variety of applications software markets; Intel has integrated into chipsets and motherboards using its microprocessors; and telecommunications and entertainment companies have integrated together in bewildering ways.

Some business experts have praised these actions, while other reputable observers have reacted with skepticism. For example, IBM's management have been criticized for having outsourced the microprocessor and operating system of their personal computer from Intel and Microsoft, choosing to participate primarily in the design and assembly stages of value-added in their product. While history has proven the decision to have been unfortunate for IBM, at the time the decision was made it was judged by many as the right thing to do.[1] It has indeed been difficult to predict, a priori, which of these moves toward or away from vertical integration would be judged in retrospect as having been managerially astute, and which would be viewed as strategically flawed. Too often for decisions as important as these, their wisdom can only be judged with the benefit of history.

This paper proposes a deductively derived model to help managers who preside over decisions to integrate or outsource to assess *ex ante* whether, when and why it might be strategically and competitively important to develop internal capabilities to perform certain activities in-house, and when it will be sensible and safe to outsource elements of value-added. Our conclusions are that:

1. The competitive advantage from vertical integration is strongest in tiers of the market where customers are under-served by the functionality or performance available from products in the market. Vertical integration tends to be a disadvantage when customers are over-served by the functionality available from products in the market.
2. As a result of (1), vertically integrated firms will often dominate in the most demanding tiers of markets that have grown to substantial size, while a horizontally stratified, or disintegrated, industry structure will often be the dominant business model in the tiers of the market that are less demanding of functionality.
3. The tendencies listed in (1) and (2) occur in end-use markets for complete product systems, such as automobiles and computers. But they also can occur in the markets for subsystems and components, which themselves comprise multiple constituent parts and materials.
4. Most often, vertically integrated firms tend to dominate many markets at the outset. Because of the patterns observed in our earlier studies, however—in which the pace of technological progress proceeds at a faster rate than customers in any given tier of the market can utilize that progress—the dominant business model in any given tier of the market will tend to shift over time from vertically integrated firms to a horizontally stratified population of specialized firms.
5. The generalization in (4) can be reversed, however, when performance gaps emerge in markets due to discontinuous shifts in the functionality demanded by customers. When this occurs, the pendulum of competitive advantage is likely to swing back toward vertically integrated firms, as companies seek to compete with each other on the basis of superior product functionality again.
6. When the dominant business model in a tier of the market shifts from vertical integration to horizontal stratification, the ability to achieve above-average profitability tends to transfer from the firms that design and assemble end-use products that historically had not been good enough, to those that build those subsystems

---

[1] See, for example, the discussion of IBM's outsourcing decisions in *Fortune*, 14 April 1997.

which limit performance of the end-use system, and which therefore are not good enough.

These conclusions began to take their initial shape in studies of the patterns of vertical integration and disintegration in the disk drive industry, in which the pendulum of competitive advantage swung repeatedly between integrated and non-integrated firms, in various tiers of the market (Christensen, 1993, 1994; Chesbrough and Kusunoki, 2001). This paper's conclusions have not been built inductively from empirical analysis, however. They have been derived deductively by combining the results of the disk drive studies with other scholars' examinations of technological modularity (Ulrich, 1995; Sanchez and Mahoney, 1996; Baldwin and Clark, 1997), and with concepts of the drivers of change in the basis of competition (Christensen, 1996; Adner and Levinthal, 2001). In this paper we provide some preliminary but promising empirical evidence supporting the model, and use the model to examine briefly the histories of the computer, automobile, software, photonics, financial services and microprocessor industries, to suggest that the model might be more broadly useful. Our primary purpose in offering this paper is to invite other scholars to test empirically these hypotheses, and thereby continue to build deeper understanding of the circumstances under which we might expect integration and non-integration to confer competitive advantage or disadvantage.

2. Relationship to prior studies of vertical integration

The model presented in this paper does not address every rationale for vertical integration and disintegration. We believe, however, that it builds upon and extends the foundations laid by several important scholars who have studied the causal drivers behind integration. Stigler's (1951) causal model, echoing Adam Smith's (1776) original analysis of the specialization of labor, asserted that a driver of specialization was market size. He observed that many industries begin as vertically integrated ones due to their small size. They then increasingly become populated by specialist firms as they grow. Stigler posited that later in the life cycle, when demand begins to contract, industries consequently tend to reintegrate. Although we agree that scale is a factor, our model views scale often as an outcome of other factors that drive specialization, rather than as a fundamental causal driver of it.

Coase (1937) and Williamson (1985) introduced the role of transaction costs as the causal driver of the optimal boundaries of the organization. There are many types of transaction costs, including threat of intellectual property appropriation (Teece, 1986), lock-in (Williamson 1979), asset specificity (Williamson, 1979; Klein *et al.*, 1978) and the challenges of coordinating interdependent investments (Chandler, 1977). Demsetz (1988) characterized transactions costs as the costs of search and maintenance, showing how these vary across the industrial life cycle. A stream of subsequent scholars within the transactions cost paradigm, including Teece (1986), Langlois (1994), Becker and Murphy (1992), Sanchez and Mahoney (1996), and Chesbrough and Teece (1996), have identified a specific type of transactions cost—the challenge of coordination amongst diverse specialists—as a driver of managerial integration across such interfaces. Monteverde's (1995) construct of 'unstructured technological dialogue' describes the management challenge when an interface between stages of value-added is interdependent and not well specified. Our model builds most directly upon Monteverde's concept.

Scholars working in a parallel stream have studied in engineering terms the concepts of architectural modularity, in order to define more precisely the conditions under which suppliers and customers of products and services might need to engage in structured versus unstructured technological dialogue (e.g. Henderson and Clark, 1990; Clark and Fujimoto, 1991; Christensen, 1994; Ulrich, 1995; Ulrich and Eppinger, 1995; Chesbrough and Kusunoki, 2001; and Baldwin and Clark, 2000).

The contribution we hope to make to the work of these scholars is to define the underlying factors that cause dialogue between customers and suppliers to be unstructured (which can entail high transactions costs if the dialogue transcends the boundaries of firms) or structured (which lowers transactions costs between firms). We also describe underlying mechanisms that may cause structured dialogue to become unstructured, and vice versa. In addition, our model helps explain why the power to earn attractive profits resides at specific locations in a value-added chain, but not at others (Porter, 1985). It also specifies the factors that can cause the power to earn attractive profit to shift to other stages of value-added.

3. Definitions

The key unit of analysis in our model is the *interface* at which a supplier of value-added and a customer of that value-added interact—whether that interface is within or between organizations. It is at this interface that structured or unstructured dialogue occurs. The specific terms that scholars, such as those noted above, use to describe this dialogue vary (Billington and Fleming, 1998; Fixson, 2000). For our purposes, we assert that for structured dialog to occur across an interface between stages or elements of value-added, three conditions must be met.

1. The customer that procures or uses a piece of value-added must understand and be able to specify to its supplier which attributes or parameters of the product or service must be provided, and to what tolerances.
2. Metrics for those attributes must exist, and the technology to measure those attributes must be available, reliable and unambiguous; and
3. The procuring company must understand the interactions or interdependencies between the attributes of what is provided and the performance of the system in which the procurer will use it. If there is any variation in what is provided, the procurer needs to understand how, when and why it will affect the performance of the system (Taguchi and Clausing, 1990).

If these three conditions are met, then the interface between a provider of an element of value-added and its user can be termed a *modular* interface, across which structured technical dialogue can occur. At modular interfaces, the necessary information exists for a market to function efficiently. Modular interfaces can occur across the boundaries of companies and across boundaries of functional groups within a company (such as between product design and manufacturing). They can also occur between groups within a project team; and they can exist between individuals. Such interfaces occur in products, services and systems of use. Henceforth in this paper, when we use the term *product*, we intend for it to apply to a service as well.

When these three conditions are not met at an interface, then we term it an *interdependent* interface,[2] across which *unstructured* technical dialogue must occur. At interdependent interfaces, the necessary information required for an efficiently functioning market does not exist. Management and integration, rather than markets, constitutes the most efficient coordinating mechanism across interdependent interfaces.[3]

Few products, services or systems would be composed exclusively of modular or interdependent interfaces—suggesting that architectures that are entirely modular or entirely interdependent would be rare extremes at opposite ends of a spectrum. This also suggests that we could rarely characterize an entire industry as being dominated by integrated or specialized firms—because this is likely to vary at the interfaces of different pieces of value-added. It will also vary, as shown below, by tier of the market.[4]

The use of these definitions in the model presented below yields results that are consistent with the findings of scholars such as Sanchez and Mahoney (1996) and Chesbrough and Teece (1996). We assert that if these three conditions of modularity—specifiability, measurability and predictability—exist at any interface, it improves the potential for an efficiently functioning market to emerge at that interface. Markets are more effective coordinating mechanisms across modular interfaces than is managerial

---

[2] Ulrich (1995) and others use the term 'integral' to refer to interfaces where these conditions are not met, and Chesbrough and Teece (1996) use the term 'systemic'. We have chosen the term 'interdependent' because it seems more descriptive of the situation. The other terms connote enough other meanings that we have chosen to employ this new term.

[3] This assertion mirrors Monteverde's conclusion that 'Roughly speaking (since other things also matter), firm boundaries . . . should congeal around transactions rich in such technically necessary, unstructured dialog' (Monteverde, 1995: 1629).

[4] In earlier papers about these phenomena (Christensen and Rosenbloom, 1995; Christensen, 1997: ch. 2), we describe the existence of a 'value network'—a nested ecosystem of suppliers and customers whose constituent companies share similar business models and process rhythms, which tend to move up-market and get disrupted as a group. The evidence in Section 5 of this paper suggests that all companies within a particular value network are not likely to uniformly employ modular or interdependent architectures. Elsewhere, Christensen suggests the existence of a 'Law of Conservation of Modularity'—a generalization asserting that interfaces across sequential elements in a value-added chain are likely to be alternately interdependent and modular (Christensen, 2001).

coordination. On the other hand, management will trump market coordination in cases where an interface is interdependent.

4. The causes of swings between vertical integration and stratification

Many companies today are striving to outsource those elements of value-added that do not build upon their strengths and can therefore be procured more cost-effectively from suppliers. History has shown, however, that many industries pass through repeated cycles in which competitive advantage rests alternately with integrated and non-integrated business models (Christensen, 1994; Fine, 1998)—suggesting that decisions to integrate or disintegrate that make sense in one context can create disadvantages when things change.

Our studies of how disruptive innovations can cause well-managed companies to fail can shed some light on the drivers behind this cyclical pattern. Because this work has been reported elsewhere, it will only be briefly summarized here.[5] There are two elements to this model, as depicted in Figure 1. The first asserts that in most markets there is a trajectory of performance improvement that customers can actually absorb or utilize over time, represented by the gently sloped lines. Secondly, as depicted by the steeply sloped lines, there is a distinctly different trajectory of performance improvement that the innovators in an industry provide to their market, as they introduce new

---

[5] The initial findings that the pace of technological progress can outstrip the abilities of customers to utilize that progress were detailed in Christensen (1992a,b). Having assembled a complete census of data on every disk drive model introduced by each company in the world disk drive industry between 1970 and 1990, Christensen measured through regression analysis the trajectory of improvement in the storage capacity of each form factor of disk drives during this period. Then, using data on disk drive capacity actually used in various classes of computers, he measured through regression analysis the trajectory of improvement utilized by customers in various tiers of the market. These results were described in Christensen (1993), Christensen and Rosenbloom (1995), Bower and Christensen (1995), Rosenbloom and Christensen (1995) and Christensen and Bower (1996). Similar econometric analysis was used in Christensen (1997) to measure the trajectories of improvement in functionality that manufacturers of excavating equipment provided, in contrast to the trajectory of performance improvement that various types of contractors were able to utilize. The 'disruptive technologies model' was inductively derived from these empirical analyses. Dan Monroe of Bell Laboratories (Monroe, 1999) and Mick Bass of Hewlett Packard (Bass, 2000) subsequently have empirically measured the same phenomena in semiconductor products. Christensen (1997) also uses the model in a deductive mode, comparing the predictions of the model to qualitative data about the histories of established and entrant companies in the computer, steel, retailing, motor controls, motorcycle and accounting software industries. Subsequent studies have found the same phenomenon in medical education (Christensen and Armstrong, 1998), retailing (Christensen and Tedlow, 2000), healthcare (Christensen *et al.*, 2000), macroeconomic growth (Christensen *et al.*, 2001) and semiconductor products (Bass and Christensen, 2002). Professor Ron Adner and his colleagues (Adner, 1999; Adner and Levinthal, 2001; Adner and Zemsky, 2001) have recently examined the same phenomenon using deductive, modeling methods.

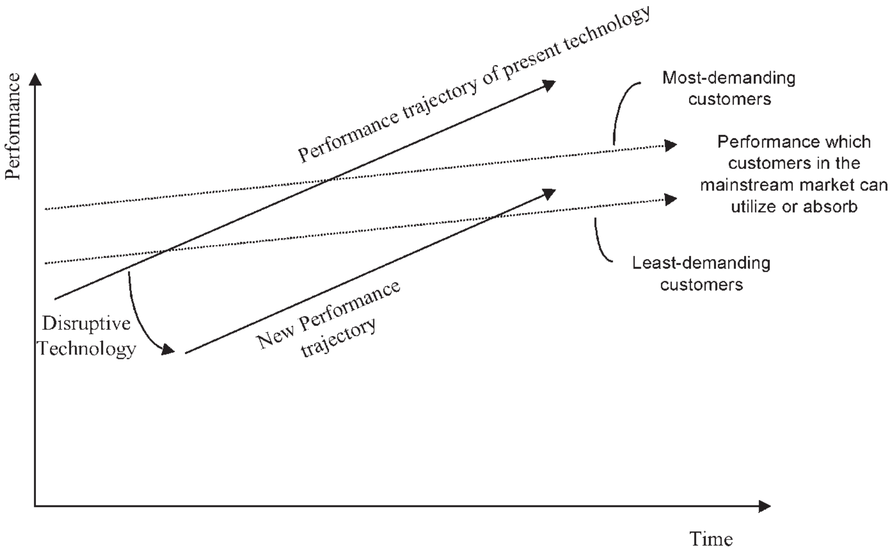

Figure 1 The intersecting trajectories of improvements that customers can utilize versus those that innovators provide.

and improved products. Our studies have shown that the trajectory of technological progress almost always outstrips the abilities of customers to utilize the improvement.

This means that companies whose product functionality is closely tuned to what customers in a tier of a market may need at one point in time typically improve products at such a rate that they overshoot what those same customers actually can utilize in later years. In other words, the functionality of a product can over-satisfy what less-demanding customers in lower tiers of the market need, even while customers in more demanding tiers of the market continue to need more functionality than even the best available products offer. It also means that 'disruptive technologies'—simpler, more convenient products that initially do not perform well enough to be used in mainstream markets—can take root in undemanding tiers of the market and then improve at such a rapid rate that they can squarely address mainstream market needs in the future.

This model has been used to describe how minicomputers displaced mainframes, and how personal computers displaced minicomputers. It illustrates how hydraulic excavator manufacturers overthrew makers of cable shovel makers; and how the Japanese automakers assaulted western car markets. It describes the mechanism through which steel minimills have been displacing integrated mills; by which the packet-switched telecommunications infrastructure is disrupting the circuit-switched network; and many others. The model has been expanded and refined, using very different research methods, by Adner and his colleagues (Adner, 1999; Adner and Levinthal, 2001; Adner and Zemsky, 2001), among others.

During the early years of many industries, in the left-most regions of Figure 1, product functionality is not good enough to satisfy the needs of customers in most tiers of the market. Competition during this era therefore focuses predominantly on product functionality: designing and producing higher-performing products is a fundamental mechanism by which companies strive to get ahead of each other (Christensen, 1996, 1997; Adner and Levinthal, 2001; Adner and Zemsky, 2001). These competitive pressures compel engineers to fit the pieces of their product together in new and untested ways in each successive product generation, as they work to wring as much performance as possible from the technology that is available. As a result, product designs tend to be interdependent, rather than modular, during this era: the design of each part tends to be contingent upon the design of other parts, and upon the way they interact within the overall system architecture. There are often powerful interdependencies between design and manufacturing during this era that are similarly based in the competitive need to stretch functionality to the frontiers of what is possible.[6]

There are two reasons why interdependent architectures predominate during eras when product functionality is not yet good enough for what customers need. The first was articulated by Ulrich (1995), who showed that creating a modular architecture—especially one that is defined by industry standards—forces designers to compromise or back away from the frontier of what is technologically possible. At the left side of Figure 1, backing off is not competitively feasible. The second reason is that new technologies are often employed in the stages and tiers of an industry where competitors are stretching toward the frontier of functionality. It is when new technologies are used to do things that have never been done before that engineers most often encounter interdependent interfaces: they do not know what to specify, cannot accurately measure important attributes and do not yet understand how variation in one subsystem will impact overall system performance. Unstructured technical dialogue is therefore the language required to compete successfully when a product's functionality is not good enough to address targeted customers' needs.

In early mainframe computers, for example, the logic circuitry could not be designed until the operating system was designed; the operating system could not be designed until the core memory was designed; and the core memory could not be designed until the logic circuitry had been designed. Manufacturing methods powerfully affected whether the system performed as it was designed to do. Everything depended upon everything else. A company could not have existed in that industry as an independent supplier of logic circuitry or operating systems, or as a contract manufacturer, because clear, modular interfaces had not yet been established to define how the parts would fit together. This implies that integrated companies can be

---

[6] Stuckey and White (1993) assert that industries will remain vertically integrated when there is *asset specificity* (a fixed asset is geographically so restricted that it is *de facto* tied to another asset), *technical specificity* (two pieces of equipment can work only with each other, and will not easily work with others) and *human capital specificity* (people whose skills are of value only within a particular working relationship). In the parlance of this paper, each of these situations is architecturally *interdependent*.

expected to dominate at the interfaces between pieces of value-added where functionality is not good enough. The dominance of IBM in mainframe computers, Digital Equipment in minicomputers, General Motors and Ford in the automobile market, Alcoa in aluminum, Standard Oil in oil, and Xerox in photocopying are all examples of firms whose vertical integration conferred competitive advantage during an era when performance was not good enough.

Because these conditions often typify an industry during its early years, scholars such as Stigler (1951) and Chandler (1977) have observed that integrated firms generally comprise the predominant business model as most industries grow towards substantive mass. Certainly industries must achieve a certain critical mass in order to support specialized competitors. We assert, however, that the fundamental causality of integrated firms being dominant at the outset and then displaced by specialized ones is not the passage of time or a general evolution towards 'maturity' or large scale *per se*. Rather, it is this causal sequence:

1. When functionality is not good enough to address what customers in a given tier of the market can utilize, firms compete by making better products.
2. In order to make the best possible products with the technology that is available, product architects tend to employ interdependent, proprietary architectures, because building a modular system around industry standards forces them to back away from the frontier of what is technologically possible. In tiers of the market where product functionality is not good enough, competitive conditions penalize companies that attempt to do this. New technologies are often employed in these conditions.
3. Because this entails unstructured technical dialogue, transactions costs are minimized through integration. Integration constitutes an important competitive advantage in managing the interdependencies in design, manufacturing, sales, service and procurement during this period.[7]

When the functionality of available products surpasses what customers in a tier of a market can utilize, however, competition changes. Customers experience diminishing marginal utility from further improvements, and consequently are less willing to

---

[7] Stigler's (1951) observation that industries tend to reintegrate and consolidate as they became mature in their later stages also may not result from shrinking scale *per se*. We have written elsewhere that after the dimensions of innovation in functionality, reliability and convenience are exhausted, price-based competition becomes predominant. It is possible that costs can be minimized most effectively from within an interdependent product architecture and integrated business model. For example, our conversations with some of Dell Computer's competitors have surfaced the possibility that Dell is over-serving the market in terms of convenience and customization—and that there are real overhead costs associated with its business model and product architecture. If an integrated supplier like IBM now offered to the market a single one-size-fits-all personal computer with more-than-enough microprocessor speed, display pixels and memory capacity, it might possible be able to steal substantial share at the low end from Dell. It is possible, therefore, that the causality of what Stigler observed is the mechanism that we discuss here.

reward further improvements with higher prices. Innovators therefore need to find other ways to compete profitably for the business of customers in tiers of the market who are over-served by functionality. Our research suggests that very often, speed to market becomes a critical dimension of competition in the lower-right regions of Figure 1. Similarly, the ability to conveniently customize the features and functions of products to the specific needs of customers in ever-smaller market niches becomes a critical trajectory of innovation that enables firms to get ahead of their competition and maintain profit margins (Pine, 1992; Christensen, 1996, 1997; Adner and Levinthal, 2001).

The efforts of disruptive competitors to be fast and flexible in this era of overshoot at the right side of Figure 1 forces them to create modular product designs in order to be competitive—because modularity creates many more options for speed, cost reduction and customization (Baldwin and Clark, 2000). When available functionality more than satisfies what customers can utilize, designers have the slack to back away from the frontier of what is technologically possible, in order to define modular architectures (Ulrich, 1995). Modularity often begins to take form in companies' proprietary interface specifications, which enable them to outsource components and subsystems at arm's length from other organizations (Sanchez and Mahoney, 1996). When one company's modular interface specifications become accepted by multiple competitors, they can become industry standards. Industry standard modularity enables firms that design and assemble products to introduce new and customized products even more rapidly than they could when interfaces were modular but proprietary, as designers and assemblers can mix and match the most effective components from the best suppliers.

Over time, the lower overheads and scale economics that focused component suppliers enjoy, coupled with the speed-to-market and flexibility advantages enjoyed by non-integrated assemblers, enables a population of horizontally stratified firms to displace vertically integrated firms (Grove, 1996; Fine, 1998; Baldwin and Clark, 2000). Modular specifications constitute sufficient information for an efficient market to work; and market-based coordination (structured technical dialogue) trumps managerial coordination across modular interfaces (Sanchez and Mahoney, 1996).

In summary, the chain of causality that shifts competitive advantage in a given tier of a market from integrated firms is this:

1. When the functionality of available products outstrips the ability of customers in a tier of the market to utilize further improvements, companies must compete differently to win the business of customers who are over-served by functionality. Innovations that facilitate speed to market, and the ability to customize features and functions in response to the needs of customers in ever-smaller market niches, become the trajectories of improvement that customers reward with premium prices.
2. Efforts to compete along these dimensions of speed, flexibility and customization cause product architectures to evolve toward modularity. This facilitates speed and flexibility.

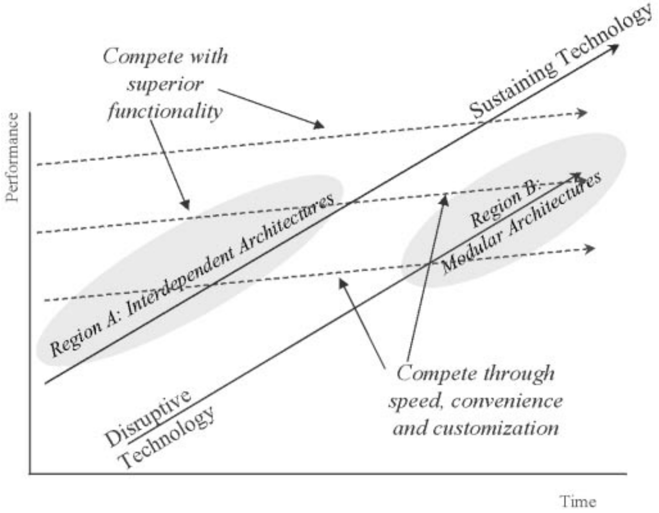

Figure 2 Overshooting the functionality that customers can utilize triggers change in the way companies must compete.

3. Modularity then enables independent, focused providers of individual pieces of value-added to thrive, because transactions cost-minimizing structured technical dialogue can occur. As a result, an industry which at one point was dominated by integrated firms becomes dominated by a population of specialized, non-integrated firms.

Figure 2 summarizes the conditions in which we would expect an industry to be characterized by functionality-based competition amongst integrated firms employing interdependent architectures (Region A) versus those in which the industry would be characterized by speed- and convenience-based competition within a population of specialized competitors who interact within modular architectures (Region B).[8]

A significant body of scholarship (e.g. Teece, 1986) has focused on the appropriability of knowledge as a critical factor affecting decisions to integrate or disintegrate. We hope that our model casts additional insight on this phenomenon as well. It implies that when the functionality of a product is not good enough to address customers' needs, the language of successful competition must be unstructured technical dialogue. The interactions through which this dialogue occurs are the 'locations' where the organization's capabilities to design and manufacture better products reside. This tacit knowledge or capability cannot be appropriated by competitors. When overshooting has occurred and competitive forces drive architectures toward modularity, however,

---

[8] The advent of a modular architecture in many cases seems associated with the emergence of a dominant design (Abernathy and Utterback, 1978; Christensen *et al.*, 1998). This association is not yet clear enough in our minds to say more than this. The possibility of this linkage, however, is something that we invite other scholars to study with us.

then the capability for fitting the pieces of the product together which had resided in unstructured technical dialogue becomes embodied in the interface standards—structured technical dialogue—that define how the modules fit and work together. This enables competitors to appropriate what had been proprietary capability and know-how.

4.1 *Case evidence supporting the disintegration model*

The model presented above was deductively derived through a synthesis of various scholars' work. In this section we offer preliminary empirical evidence—some of it in the form of numerical analysis, some in the form of narrative history—that is consistent with the chain of causality in this model. In case studies of industries as diverse as disk drives, computers, financial services and microprocessors, we observe a process similar to the one outlined above that transferred competitive advantage from integration towards non-integration. We summarize these observations in the following.

4.2 *Evidence from the disk drive industry*

Our earlier research described how the performance of disk drives improved at a more rapid pace than the ability of customers in any given tier of the market could absorb those improvements. Over and over, this enabled disruptive innovators piercing into the market's underbelly to displace the industry's leaders (Christensen and Rosenbloom, 1995; Christensen and Bower, 1996).

This continuous process of up-market migration implies (in the language of this paper) that architectural modularity is likely to occur in the least-demanding tiers of the market first; and that at any point in time we should expect the most demanding tiers of the market, which are the most under-served by the functionality of available products, to be populated by more technologically interdependent products. Consequently, we would expect integrated firms' market positions to be strongest in the most demanding tiers of the market, and the market shares of non-integrated firms to be strongest in the least-demanding, most over-served tiers of the market.

In our study of the disk drive industry, we devised a method to measure the degree to which the architecture of a drive was modular or interdependent. It is an indirect measure, but seems to support the notion that modularity appears first in the least-demanding tiers of the market, where the phenomenon of overshooting first occurs. The analysis suggests that modular architectures then migrate toward more demanding tiers of the market, as this on-going process of overshooting successively more demanding tiers of the market continues.

To do this analysis, we built a database of every model of disk drive introduced by any company in the world between 1975 and 1998—4334 models in all. This constitutes a complete product census for the industry in these years.[9] For each of these models, we

---

[9] The data were obtained as a generous gift from Mr James Porter, Editor of *Disk/Trend Report*. We have

had data on the types of components that were used in the drive—including hardware components, and the types of firmware and software coding that were employed. We then estimated regression equations, in which the dependent variable was the recording density of the drive.[10] The independent variables were the year in which the drive was introduced, the size of the drive[11] and the components that were used in the drive—represented by dummy variables for each type or generation of component technology. Where interviews with engineers suggested that interactions amongst components might affect the recording density achieved in a product, interaction terms were included in the analysis. The equation was estimated in the following form:

$$\ln(\text{Recording Density}) = B_1 + B_2(\text{Year}) + B_3(\ln \text{Disk Diameter}) + \\ B_4(\text{Component Dummy } 1) + \ldots + B_n(\text{Component Dummy } n)$$

The coefficients that were estimated for each of the component dummy variables measured the extent to which the use of various component technologies added to or detracted from the recording density of the product. The coefficient of the year variable measured the annual improvement in recording density that resulted from general, incremental advances that could not be linked to the use of particular new architectural, component, software or firmware technology. Detailed results from this analysis are reported in Appendix 1. The adjusted R^2 was 0.95, indicating that the variables accounted for most of the variation in density across products in the sample.

This equation allowed us to estimate the expected recording density of each drive, given its size, the components that were used and the year in which it was designed. We could then compare the expected density with the density that its engineers actually achieved. We called the ratio of the actual recording density to the expected density the *architectural efficiency* of the drive, and calculated this ratio for every disk drive model in the database.[12] An architectural efficiency ratio of 1.0 indicates that the engineers

---

entered the data in a huge Excel spreadsheet, and would be happy to share the data with colleagues who wish to analyze it further (available on request from C.M.C.). Although *Disk/Trend Report* recently ceased publication, the San Jose Public Library holds past copies of the reports.

[10] Recording density is measured as the number of megabits of information that can be stored on a square inch of disk area.

[11] The diameter of the disk actually has a strong effect on the recording density that is feasible, because the inertial problems of precisely positioning larger components over a particular track of data are much greater in large drives than small drives.

[12] Professor Marco Iansiti (Iansiti, 1997) used different methods to develop an analogous measure in his study of product design processes in the computer workstation industry. He labeled his measure 'technological yield'. We prefer to use the term 'architectural efficiency', to be consistent with earlier publications that employed this measure (Christensen, 1992a,b) and because it is more descriptive of the phenomenon we are trying to measure. Whereas Iansiti compared what was theoretically achievable versus what was actually achieved, we have measured the average of all engineers' work versus the work of the individual product design teams that developed each of the products.

achieved exactly the expected density. Ratios above 1.0 indicate that, through clever product design, the engineers were able to wring more recording density out of the same set of components than the average engineer would have done. A ratio of less than 1.0 suggests that the drive's engineers got less-than-expected performance, given the components that they used.

If the interface standards amongst the components were so completely defined that engineers had no degrees of freedom in designing how they would fit the components together in the drive's architecture—i.e. if the design were completely modular—then we would expect the architectural efficiency of the population of drives to be 1.0 and the standard deviation of architectural efficiency to be 0. The larger the standard deviation in the architectural efficiency of the product models in this population, the greater the scope for differentiated techniques for integrating components. In other words, the greater the standard deviation, the greater the degree of architectural interdependency in product designs. And the lower the standard deviation, the greater the degree of architectural modularity in the drive.

Figure 3 maps on its vertical axis the standard deviation of the architectural efficiency of drives sold into the desktop computer market between 1980 and 1995, by tier of the market—ranging from the drives in the lowest-capacity quartile at the front to those of the highest-capacity quartile in the back.[13] Fleming and Sorenson (2001) used very different methods to arrive at a similar conclusion.

Note that the standard deviation increases from front to back—from the lowest quartile to the highest quartile in each of the periods—suggesting that the degree of interdependency was always greater in the most-demanding tiers of the market. While the degree of interdependence/modularity seems to have been stable over time in the lowest quartile of products, in the second, third and fourth product quartiles the degree of interdependency decreased monotonically over time, as the ongoing process of overshooting and increased modularity progressed upward through tiers of the market. This suggests that the scope for product differentiation was always most limited in the lowest tiers of the market in which customers were most over-served and greatest in the most demanding tiers, where customers' thirst for improved performance still required more interdependent architectures. The stable standard deviation in the lowest tier of

---

[13] There is a common problem in analyses of this sort, as one of the reviewers of this paper pointed out. Quoting from the review letter, '[If] you plot the size of the residual at different levels of the dependent variable, [you often] find that the residual is larger on the larger end of the distribution of the dependent variable, and smaller at the smaller end of the distribution of the dependent variable. It would be very surprising to have found any other pattern, [because] is common for errors to be heteroscedastic in proportion of the to the dependent variable.' He or she is correct. It is for this reason that our measure of architectural efficiency we have used in these studies is the ratio of actual to expected recording density, rather than the absolute magnitude of the residual. Using the ratio normalizes for the effect of the absolute magnitude. Indeed, had we not normalized in this way, the plot would have been extraordinarily misleading because recording densities have increased dramatically over the period.

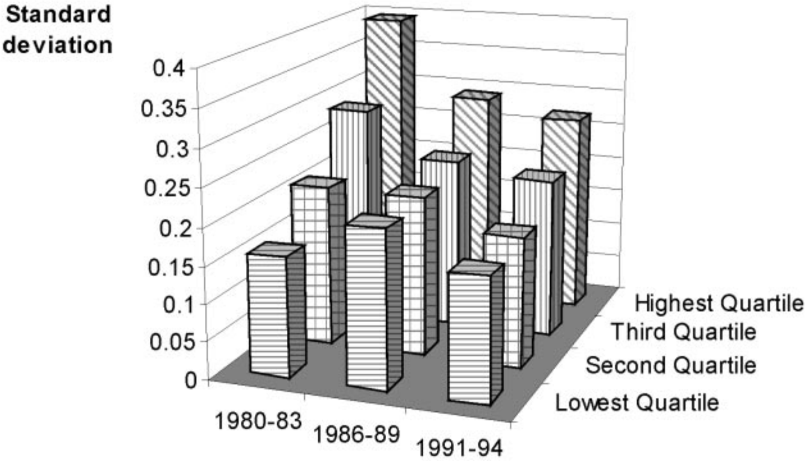

Figure 3 The progress of modular architectures through progressively demanding tiers of the disk drive market.

the market supports the notion that few products are entirely modular at every interface.[14]

As shown below, the industry's vertically integrated firms—particularly IBM—have dominated the most-demanding tiers of the market, while non-integrated manufacturers such as Quantum and Western Digital held the largest shares in the least-demanding end of the market.[15]

4.3 The computer industry

We do not have a similar set of detailed, component-level data for the computer industry as for disk drives, but it appears that a similar pattern holds in this industry as well. Products in the most performance-demanding tiers of the market are architecturally interdependent and proprietary, and are supplied by integrated companies. The architectures of products targeted at progressively less performance-intensive tiers

---

[14] The only deviation from the trend towards increased modularity seems to have occurred in the late 1980s, when the drive makers in all tiers of the market began to use thin film heads. As Waid (1989) notes, thin film heads constituted a fundamentally interdependent technological challenge during the earliest years of their use, because many elements of the drive's design were interdependent with elements of the thin film head design. This supports the point suggested above that when new technologies are used their interactions with other elements in the system design are not well understood.

[15] Some readers of earlier drafts of this paper have wondered whether, in the lower tiers of truly commoditized product markets, architectures might become interdependent again. If everybody in a significant portion of a market wanted exactly the same features and functions, and their desires were stable over time, the flexibility and options value of modularity might have little value (Baldwin and Clark, 2000). It then might be possible that a single interdependent product design might indeed be a lowest-cost solution.

of the market are progressively more modular, and are supplied by progressively less-integrated companies.

In the early years of digital computing, when the functionality of available products fell short of what the mainstream markets needed, the computer industry was dominated by integrated players such as IBM. Even today, the most demanding tiers of the mission-critical enterprise server business continue to be dominated by integrated companies such as Hewlett Packard, IBM and Silicon Graphics. Their products are technologically interdependent, built around proprietary reduced instruction set computer (RISC) microprocessors and proprietary UNIX operating systems whose key properties are interdependently designed and manufactured, largely in-house. The performance of their products has overshot what is utilized in all but the most demanding tiers of the market—where unit volumes are so small, in fact, that Silicon Graphics' once-spectacular growth trajectory has sputtered.

Products in the next-lower tiers of the server business are more modular in character. Sun Microsystems' Solaris operating system, for example, is rapidly becoming a standard. Predictably, this market tier is dominated by less-integrated manufacturers. Sun, for example, continues to design its own microprocessor and operating system, but licenses them to competitors and outsources fabrication. Sun is aggressively pushing up-market to disrupt Hewlett Packard, IBM and Silicon Graphics, carrying its more modular architecture with it in the process.[16]

The less-demanding tiers comprising the business computing market are dominated by suppliers such as Compaq, Dell and Gateway, whose products are consummately modular. These firms are not integrated; most components in their products are supplied by specialist companies. Manufacturing and the in-bound and outbound logistics are often managed by contractors such as Solectron, and even the design of some products is being out-sourced. Dell, in particular, leverages its status as a non-integrated assembler of modular products to conveniently customize its computers to the specifications of individual customers, and deliver the machines to their doorsteps within 48 hours. These firms began their histories squarely in the personal computer space, and have aggressively carried modularity and disintegration up-market, disruptively stealing market share in the workstation and server space from Sun. As components get more capable, the non-integrated companies carry their modular architecture up-market—hence, disintegration is occuring in progressively more demanding tiers of the market.

4.4 Mortgage banking

The mortgage banking industry has historically been dominated by integrated institutions such as savings banks and savings and loans institutions, which collected and serviced deposits, originated loans, evaluated borrowers' credit worthiness, assessed property values, closed loans and serviced them. In terms of the definition of

---

[16] The pattern in which these waves of disruptive technologies are sweeping through the tiers of the computing market is described in greater detail in Christensen and Verlinden (1999).

modularity noted above, there were no standard ways to measure the riskiness of a loan made to any borrower, and as a result, markets could not emerge at the interface of these stages of value-added.

Asset securitization and credit scoring systems that originated in the credit card industry essentially replaced bank officers' judgement with simple metrics—they brought modularity. With credit scoring came knowledge of which attributes of the borrower needed to be specified, and technology for measuring those attributes became known. Likewise, asset securitization transformed loans from non-standard assets with uncertain risks and returns, to standardized units with easily measured risk and return. Credit scoring and securitization took root in the 1960s in the lowest tier of the lending market—credit cards of retailers such as Sears. These then migrated up-market, usurping open credit cards, auto loans, mortgage loans and, most recently, small business loans. In each of these market tiers, integrated commercial and savings banks have been replaced by a horizontally stratified population of specialist firms such as MBNA, GMAC, GE Capital, Countrywide and FNMA. Integrated banks' share of the mortgage market, for example, has eroded from over 90% in the 1960s to 39% by 1999. A population of specialized firms now originate most mortgages, perform credit checks, value collateral, close loans and service them (Hodes and Hall, 1999).

4.5 Disintegration of the microprocessor industry

Our final case is the microprocessor industry. Although the microprocessor is a component within a modular personal computer, the microprocessor itself is a complex, technologically interdependent system. Projects to develop next-generation microprocessor platforms consume time and resources of a magnitude similar to those that were required to design new mainframe computers. Intel is an integrated company, designing for itself each element of the microprocessor in an interdependent, iterative process. Its 'copy exactly' method of transferring designs into volume production is a testament to the complex and poorly understood interdependencies between design and manufacturing.

While the speed of complex instruction set (CISC) microprocessors has gotten fast enough that Intel and AMD are disrupting RISC microprocessor-based machines in the higher tiers of the market, their products have overshot the speed that typically is utilized in mainstream business applications. In fact, Monroe (1999) has shown that the Moore's Law pace at which transistors are being made available per area of silicon is outstripping the ability of circuit designers to utilize transistors by 40% each year. As a consequence, in the less performance-demanding tiers of the market, the architecture of microprocessors, such as the Intel Celeron processor, is becoming more modular (interview with Mr Randy Steck, Intel Architecture Labs, July 1999). And at the lowest end, for those chips in hand-held wireless digital appliances, companies like Tensilica have begun to offer web-based tools that enable applications developers to assemble from modular components custom-designed microprocessors and systems-on-a-chip whose features and functionality are tuned exactly to the requirements of the application. Design cycles for these modular microprocessors are measured in weeks, rather than years (Bass and Christensen, 2002).

The findings of Macher (2001) support these assertions. He has shown that integrated semiconductor manufacturers perform better than non-integrated ones in the most performance-demanding tiers of the market, whereas the opposite is the case in less-demanding tiers. Furthermore, with clearer design-for-manufacturing rules as an interface, chips positioned away from the leading edge are increasingly being fabricated in independent silicon foundries, which he shows are able to bring products to market much more rapidly than integrated firms.

4.6 Synthesis across these cases

Table 1 summarizes the patterns revealed in this set of cases. It lists down the left-most column the causal chain in the model we are proposing. For each of the industries we have studied, an X in the cells of the table indicates where that phenomenon was observed.

4.7 Case studies in reintegration

The trajectory maps in Figures 1 and 2 suggest that the predominant business model in many industries generally will evolve from integrated firms toward non-integrated, specialized business models. But on occasion the trend has reversed itself, back towards integration. Other financial reasons for reintegration are considered in Section 5. The factor that seems to have driven the re-ascendance of integration as a source of competitive advantage, however, was the occurrence of a ‘performance gap’—an upward shift in the functionality that customers needed. In terms of Figures 1 and 2, this involves an upward shifting of the dotted, gently sloping lines to a new height. The emergence of these performance gaps can throw an industry back into a ‘Region A’ situation, as depicted in Figure 2. When this happens, it demands again true managerial and technological reintegration, as innovators through unstructured technical dialogue are again forced to piece the components of their products together in unconventional and untested ways, in order to push performance as close as possible to what customers have begun to demand. The following sections describe in some detail why and how this happened in disk drives, and then recount a similar pattern in the software industry as well.

4.8 Reintegration in disk drives

Through most of the 1990s the 3.5-inch drive market was largely in Region B of Figure 2. These drives were used primarily in desktop personal computers, and their capacity—as big as 60 GB—had substantially overshot what customers actually were able to utilize in the mainstream tiers of that market. As shown above, the architecture of drives sold into this market was increasingly modular, especially in its less-demanding tiers. This meant that components from a variety of suppliers could be mixed and matched with predictable results in new product designs. This market, consequently, was dominated by less integrated companies—Seagate, Quantum,

Table 1 Supporting evidence in case studies for key elements of the model

	Disk drives	Computers	Financial Services	Micro-processors
1. At the outset, when available functionality is insufficient to meet customer needs in mainstream tiers of the market, product architectures are interdependent.	X	X	X	X
2. The industry is dominated at this time by vertically integrated firms.	X	X	X	X
3. The functionality provided by the leading integrated firms overshoots what customers in lower tiers of the market can utilize and are willing to pay for.	X	X		X
4. The basis of competition in those tiers of the market that are over-served in functionality changes. Speed to market, and the ability to conveniently customize features and functions become competitively important.	X	X		X
5. Product architectures become modular to facilitate competition on new dimensions.	X	X	X	X
6. Modularity enables non-integrated firms to compete. In those tiers of the market in which overshooting and modularity have occurred, the industry tends to disintegrate; a horizontally stratified population of specialist firms displaces integrated ones.	X	X	X	X
7. Because the pace of technological progress proceeds faster than the ability of customers in given tiers of the market to absorb it, the sequence of events in steps 1–6 above recurs, in each progressively more demanding tier of the market.	X	X	X	X

Western Digital and Maxtor. IBM, the most extensively integrated competitor, was barely been able to sustain a foothold in that market.[17]

The 2.5-inch disk drive market, in contrast, was in Region A. Even though they

---

[17] We have deliberately used the term 'less integrated' rather than 'non-integrated' here because, by this point in the industry's history, all of these firms were integrated to some extent, positioned at various points along the spectrum. Seagate (especially after having acquired Conner Peripherals) had a thriving disk-making operation, and had a magneto-resistive (MR) head operation that was beginning to bear fruit. It had deep expertise in thin film head-making. Quantum designed its controller circuitry but outsourced everything else, including manufacturing. It had attempted to begin making MR heads by purchasing Digital Equipment's disk drive business, but stumbled badly. By 1998 it had essentially

emerged chronologically after the 3.5-inch drive, the functionality of 2.5-inch drives used in notebook computers was not yet good enough. The reason? Computer users attempted to use notebook computers for essentially the same applications as they used desktop computers. Because the 2.5-inch drives in notebooks have one-sixth the surface area for recording than their 3.5-inch desktop siblings,[18] for most of the 1990s notebook computer users were largely dissatisfied with the capacity, weight and power consumption of 2.5-inch drives. As a result, 2.5-inch drives were built around MR heads and PRML error detection codes—complex, non-standard technologies that required interdependent, iterative design processes in order to wring as much performance as possible out of these new technologies.[19]

The 2.5-inch drive market was dominated by the industry's most technologically

---

passed the MR head hot potato to Matsushita Kotobuki Electric (MKE), its manufacturing partner, and MKE was laboring to learn how to make and integrate the heads. Western Digital and Maxtor were the least integrated. IBM was by far the most extensively integrated—especially in the linkages between its components, its research activities that supported advanced-technology components and its read–write channel design activities. The main differences in integration between IBM and Seagate are in IBM's extensive research activities, and in the manner in which IBM's engineers seem to be able to integrate their pieces of value-added—especially at the partial response, maximum likelihood (PRML)—MR head interface.

[18] Each 2.5-inch disk has half the recording area of a 3.5-inch disk, but because the 2.5-inch form factor is used in notebook computers, it must be much thinner, allowing fewer disks to be stacked on a spindle than in the 3.5-inch architecture.

[19] Evidence that the architecture of 2.5-inch disk drives is interdependent rather than modular comes from many sources. This first was a set of twenty-four interviews conducted with engineering managers at IBM, the industry's most integrated company; Seagate, a partially integrated firm; Quantum, a non-integrated assembler of drives; and Read-Rite and Komag, which were non-integrated suppliers of heads and disks, respectively. In every case, they noted that in using PRML codes and MR heads (defined below) to maximize the density of 2.5-inch drives, they could not work with suppliers because they could not specify what suppliers had to deliver, and could not measure whatever attributes of the heads were most critical for maximizing performance. They all attributed IBM's success in this market to its ability to conduct all of the required design and manufacturing in-house, in integrated teams. We also conducted statistical analyses of the phenomena, showing how the statistical significance of interaction terms between components in the regression equations described in the appendix varied across 3.5-inch and 2.5-inch drives. We have not included those results in this paper because of length constraints, but they point to the same conclusion. Drives whose functionality is nearer to the frontier of feasibility have more interdependence in their architectures. The interdependency between these technologies occurs in what engineers in the industry call 'the channel'. MR heads offer a completely different, and much more sensitive, method for detecting changes in the flux field on a disk than prior inductive head technology—enabling much smaller magnetic domains to be created on disks. PRML software algorithms detect when errors in reading data might have occurred and, based upon patterns in other data, estimate what the missing or erroneous data are. The ability to maximize recording density by using the most advanced MR heads depends upon the ability of PRML coders to identify and correct error patterns, which arise because of the way the heads are designed. Both pieces of technology must be done interdependently. This is not the case with the technologies that are used away from the frontier of possibility, such as inductive thin-film heads and run-length-limited (RLL) error-correction codes. Both can be procured and used off the shelf from third parties.

integrated companies—IBM, Toshiba, Hitachi and Fujitsu. Although this market had been served for a few years at its outset by non-integrated firms, as the character of customers' needs became clear, the non-integrated players with their modular product architectures were completely driven from that market. Their share fell from 96% in 1990 to 13% in 1996 and 3% in 1998, as the integrated firms learned to focus their diverse technological capabilities on the customers' needs for maximum recording density. Evidence of the re-ascendance of the integrated business model is summarized in Table 2.

There is now some evidence that the capacity trajectory of 2.5-inch drives has begun to intersect with the capacity demanded in the notebook computer marketplace. This portends another pendulum swing towards modular architectures, shifting competitive advantage back toward non-integrated competitors in this particular market.

Chesbrough and Kusunoki (2001) describe the difficulties that non-integrated firms have escaping the 'modularity trap' when an industry passes through a 'technology phase shift', suggesting that if they consciously and capably manage the swings between interdependence and modularity, integrated firms ought to have long-term performance advantages over non-integrated firms. Our work supports their finding, and perhaps adds a bit more specificity about the causes of 'technology phase shifts', the 'double helix' pattern that Fine (1999) observed and the 'architectural reconfigurations' that Henderson and Clark (1990) examined.

4.9 Personal Computer Software

Just like the 2.5-inch disk drive market, the personal computer software market when it

Table 2 Contrast in 1998 market shares held by non-integrated versus integrated companies in the over-satisfied 3.5-inch market and the under-satisfied 2.5-inch market

Market shares in the 3.5-inch market (%)		Market shares in the 2.5-inch market (%)	
Integrated firms	8	Integrated firms	97
IBM	4	IBM	67
Toshiba	1	Toshiba	21
Hitachi	2	Hitachi	5
Fujitsu	1	Fujitsu	4
Less-Integrated Firms	87	Less-Integrated Firms	1
Seagate/Conner	33	Seagate/Conner	0
Quantum	23	Quantum	0
Western Digital	23	Western Digital	0
Maxtor	8	Maxtor	1
Others	5	Others	2
Total	100	Total	100

Source: *Disk/Trend Report*, 1999.

coalesced was populated by non-integrated companies. Microsoft's DOS constituted a standard interface into which non-integrated software vendors such as WordPerfect, Borland, Lotus and Harvard Graphics could 'plug' their modules. But within a few years, as customers came to understand what they wanted, a 'performance gap' emerged—PC users began demanding the ability to transport portions of graphics, spreadsheets and word processing files into other types of file. This performance gap demanded integration, and Microsoft responded by creating non-standard, interdependent connections amongst its Windows operating system and its suite of office applications—and later its Internet Explorer. Almost overnight, Microsoft's non-integrated competitors vaporized.[20]

Today, however, the pendulum seems to be swinging in the other direction. The functionality and number of features in most of Microsoft's products have dramatically overshot what most of its customers actually are able to use. Non-integrated software firms writing to disruptive internet protocols and the Java programming language, with their modular architectures, are capturing a dominant share of internet-oriented applications, in a classic disruptive technology fashion. Linux, an operating system whose modular architecture enables open-source devotees independently to maintain and improve elements of the system, is beginning to disintegrate certain tiers of the market as well.

These cases of reintegration constitute what Yin (1984) calls *theoretical* replications of the model proposed in this paper. The model suggests that overshooting the functionality required in a tier of the market precipitates a change in the basis of competition, which in turn causes product or service architectures to evolve from interdependency toward modularity. This in turn causes industry structures to evolve from vertical integration towards specialized stratification. In the cases described immediately above, the emergence of functionality gaps, or 'under-shooting', caused this process to reverse itself towards integration.

5. Shifts in the locus of profits

Our research also suggests that the stages of value-added in which attractive profits can be made tend to differ from the left to the right sides of the disruptive technologies map. During eras characterized by Region A in Figure 2, the largest vertically integrated firms, which engage in designing and assembling architecturally interdependent end-use products whose performance is not yet good enough, tend to capture a disproportionate share of their industry's profits. During the eras of horizontal

---

[20] It seems that the foresight of Microsoft's management team is a common interpretation of why Microsoft made this move toward interdependent architectures, whereas firms that were managed by less aggressive or competent teams, such as WordPerfect, Novell and Lotus, missed this opportunity. To provoke discussion, we are specifically proposing that there is a more fundamental causality behind what happened: the performance gap forced integration, and Microsoft was in the best position to respond.

stratification described as Region B, in contrast, the firms engaged in those same stages of value-added—where more-than-good-enough modular products are designed and assembled—typically find it very difficult to earn more than subsistence profits. Whereas component suppliers tend to struggle to be profitable in Region A, in Region B the firms that supply technologically interdependent subsystems to the assemblers make the lion's share of profit.

The reason why the ability to earn attractive profits flips is that two factors that drive the ability to earn unusual profits—steep scale economics and the ability to create differentiated products—favor designers/assemblers in Region A and subsystem suppliers in Region B. In the next section we will recount in some detail how and why this happened in the disk drive industry, and then suggest how the same phenomenon seems to be occurring in the computer, telecommunications and automobile industries as well.

5.1 The narrowing scope for differentiation

Modularity brings benefits of speed, lower cost and technological flexibility, as Baldwin and Clark (2000) have described. Indeed, adopting modular product architectures is critical to survival in a world where, as suggested in Figure 2, the basis of competition centers upon speed to market and the ability to conveniently customize features and functions to the needs of specific sets of customers. Without modular architectures and disintegrated business models, firms in Region B simply could not compete effectively.

The downside of modularity is that it seems also to narrow the ability of competitors to differentiate their products through superior design. This was demonstrated in the analysis summarized in Figure 3, which described how the variability in the architectural efficiency of disk drives dropped as product architectures became more modular. In addition to losing the ability to differentiate products on the basis of performance, the designers and assemblers of modular products also lose their ability to differentiate on the basis of cost. The cost structure of non-integrated design/assembly firms tends to be dominated by variable, rather than fixed, costs. Because it is high fixed costs that give rise to steep scale economics, assemblers of modular products compete on relatively flat scale curves, meaning that small competitors can enjoy similar costs as larger ones.

In an attempt to illustratively measure the flattening of scale economics in a modular world, we collected data on the unit volumes, total costs and product line complexity for each disk drive manufacturer and built a regression model that allowed us to estimate each manufacturer's cost, during each year, to produce a drive of a given capacity. The equation takes the form

$$\ln(\text{Product Cost}) = B_0 + B_1 \ln(\text{Drive Capacity}) + B_2 \ln(\text{Total Units Produced}) + B_3 \ln(\text{Product Line Complexity})$$

The variables are defined as follows: *Product Cost* is calculated by dividing the total operating costs in the company, exclusive of interest and taxes, by the number of disk drive units produced. Hence, we call this measure *fully allocated product cost. Drive Capacity* is the weighted average capacity of the disk drive units shipped each year by the company. This is an important variable, because higher-capacity drives are more costly to produce. We expected the coefficient of this variable to be positive. *Total Units Produced* is the total number of disk drives shipped during the year. We expected the coefficient of this variable to be negative, positing that as scale increased, unit costs would fall. *Product Line Complexity* is the number of product families produced by the company in the year. We expected the coefficients of this variable to be positive—overhead costs per unit would increase as increasing complexity of the product line would demand higher management overheads.

All of the data required for this calculation were taken from *Disk/Trend Report.* The equation for the early 1980s, when modular architectures were just beginning to penetrate the industry, was

$$\ln(\text{cost/unit}) = 296.39 - 0.146(\text{year}) - 0.370\ \ln(\text{unit volume}) +$$
$$0.126\ \ln(\text{no. of families}) + 0.511\ \ln(\text{weighted mean MB/unit})$$

t-statistics: (–3.44) (–4.70) (1.68) (6.23) $R^2 = 0.88$

The equation for the early 1990s, when modular architectures had become pervasive, was

$$\ln(\text{cost/unit}) = 322.22 - 0.160(\text{year}) - 0.15\ \ln(\text{unit volume}) +$$
$$0.014\ \ln(\text{no. of families}) + 0.544\ \ln(\text{weighted mean MB/unit})$$

t-statistics: (3.41) (–0.52) (0.12) (4.20) $R^2 = 0.88$

Note how the scale coefficient fell and became statistically insignificant, as did the complexity coefficient (no. of families). We would expect both in a regime of modularity.

We estimated this equation for each year. The coefficient B_2 constituted a measure of the steepness of the scale economics in disk drive manufacturing at each point in the industry's history. Using this equation, we could then estimate what it would cost each manufacturer to make a drive of a given capacity, given the scale at which it produced in any year, and the complexity of its product line, measured by the number of product families. The scale curve as it looked in the late 1980s is shown in Figure 4, which charts the fully allocated costs on the vertical axis and each firm's production volumes on the horizontal axis for a typical modular 100 MB 3.5-inch drive. It shows that the scale economics in design and assembly of modular disk drive designs were flat—small competitors could add this value almost as cost-effectively as the largest firms—because the cost structure was dominated by variable rather than fixed costs.

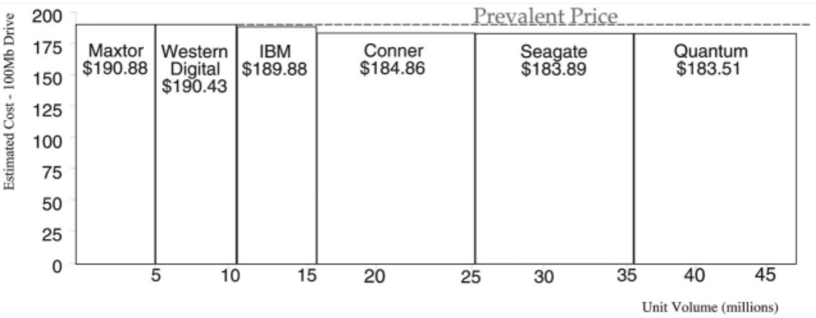

Figure 4 Scale economics in the design and assembly of modular 3.5-inch drives in the late 1980s.

For the makers of modular disk drives, competition in the absence of performance and cost differentiability has been difficult and unrewarding. It is the users of disk drives, not the manufacturers, who have reaped the benefits of lower cost, more flexibility and greater speed in product development that result from modularity. Typical gross margins for high-volume drives used in desktop personal computing fell from 35% in 1984 to become mired in the 10–15% range in the last years of that decade.

5.2 The profitability of component manufacture

While the business of design and assembly of modular drives was a wearying race on an accelerating treadmill in Region B of Figure 2, the business of making heads and disks during this same period shifted progressively to Region A of Figure 2. In order to force the cost of disk drives ever-lower, designers could never be satisfied with the functionality (recording density achievable) of heads and disks—because the higher the recording density, the fewer the number of disk platters and heads required in the drive. As a consequence, over this period heads and disks themselves become more technologically interdependent assemblies of materials, creating substantial scope for product differentiation. The high fixed costs of designing and manufacturing thin film disks and MR heads steepened the scale economics in this slice of value-added as well. Figure 5 shows, for example, that Komag, the largest independent disk manufacturer (which also made the highest-performance disks in the industry),[21] had substantially lower costs than its competitors. Because products were differentiable and scale economics were steep, the leading independent component makers were highly profitable. In contrast to the 11% annual rate of return that the shareholders of non-integrated disk drive assemblers received between 1988 and 1994, the leading component makers, Komag and Read-Rite, returned 38% annually to their shareholders.

---

[21] The sources of these data were *Trend/Focus*, an annual market research report on the industry supplying components to the disk drive assemblers, and the engineering staff at Komag. Figure 5 was constructed with an economic model built in conjunction with the engineering staff at Komag, which estimated the cost of producing a disk of a given quality at various volumes. By inserting *Trend/Focus* data into the model, we developed estimates of the production costs of various competitors.

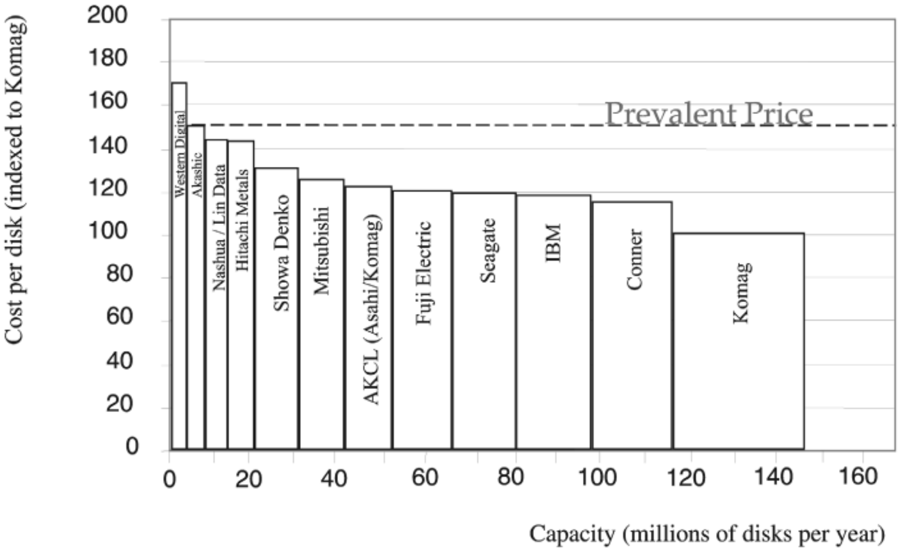

Figure 5 Supply curve for the thin film disk industry, 1994.

In response to this shift in the stage of value-added in which attractive profits could be earned, some leading assemblers of drives began producing their own components. Figure 5 makes it easy to see why. Although their costs of production were higher than Komag's, the price at which assemblers were buying disks from Komag was determined where the supply and demand curves intersected—at the cost of the *marginal* supplier in the market—Akashic. As long as their scale enabled them to produce at greater volumes than the marginal supplier, and as long as the scale curve sloped steeply enough, these companies found disk-making to be a more attractive slice of value-added in which to participate than was design and assembly (see footnote 21). Hence, one of the leading non-integrated assemblers, Conner Peripherals, built an independent business to manufacture disks; and another, Seagate, developed businesses to make both disks and heads, which involved even steeper scale economics. These firms initially made components for internal consumption only, but ultimately found it compelling to sell to competing assemblers as well.[22] A number of analysts, in fact, reported that for many years in the 1990s over 100% of these firms' profits could be attributed to the value they captured in component manufacturing operations (meaning that they lost

---

[22] By the late 1990s, Komag's profits had plummeted. But in an odd way, its collapse supports the thesis of this section. So many disk drive assemblers had been enticed into making their own disks because of the profitability of disk-making relative to final product assembly, that little merchant-market volume was available to Komag by the late 1990s. Its costs consequently rose and its profitability was decimated. But it was the profits that attracted the assemblers into disks, and not the requirement for technological and managerial coordination that drove this move to vertical integration. A similar fate befell Read-Rite, at least temporarily. Especially with the advent of magneto-resistive heads in the mid-1990s—an immature technology with steep scale economics—the design/assembly firms that could afford it brought as much head-making capacity in-house as possible.

money in design and assembly). IBM subsequently has followed suit in selling components into the market as well.

It is important to note that while economists might call these firms' migration into making components 'vertical integration', it would be more accurate to say that they established managerially and technologically independent business positions in the component-manufacturing stage of value-added. This stemmed from a very different motivation and entailed a very different management structure than did the requirement to create new, interdependent product architectures, built through unstructured technical dialogue in response to the performance gaps described above.

Our conclusion, in essence, is that attractive profits tend to be earned where performance is not yet adequate relative to the needs of user in the next stage of value-added; and where, therefore, product architectures are likely to be proprietary and interdependent in character. Because the pace of performance improvement typically outstrips the ability of customers to utilize that improvement, the places in the value chain that presently enjoy attractive profitability are likely to lose their ability to continue those levels of profitability, and vice versa. In the following, we will summarize how this same shift in the ability of designers/assemblers versus component suppliers to capture attractive profits occurred in computers, and is beginning to happen in the automobile industry.

5.3 Shift in the locus of profitability in the computer industry

The largest firms that designed and assembled computers in the technologically interdependent era—particularly IBM and Digital Equipment—captured extraordinary profits because of the differentiability of their products and the high fixed costs (and consequently steep scale economics) in design and manufacturing. They wielded such power that most of their parts suppliers survived at subsistence levels of profitability. But, as personal computers with modular architectures came to dominate mainstream markets, the tables turned. How does a designer of a modular personal computer in a firm such as Compaq create a better product than competitors such as Dell, Gateway, Hewlett Packard or IBM? Incorporate a faster microprocessor? A higher capacity hard drive? More megabytes of DRAM? In a consummately modular product there are so few degrees of design freedom that the only way to offer a better product is to offer higher-performance components, which competitors can also offer. When most costs are variable, the scale curve flattens substantially: it becomes difficult to assemble at lower costs too. As speed to market and the ability to mass-customize become the only dimensions along which assemblers of modular products can compete, firms in this stage of value-added can find competition to be an unrewarding race on an accelerating treadmill.[23]

---

[23] The first firms to identify how the basis of competition amongst assemblers of modular products in over-served tiers of the market shifts to speed and customization can, in fact, do well for a time. This was certainly the case with Dell Computer and Chrysler, for example. The operations-based abilities to compete in these terms against other assemblers, however, can be replicated, as Porter (1996) notes.

The impact on the profitability of design and assembly in the computer value-added chain has been predictable. ‘In 1986, companies that built and sold computer systems captured about 80% of the total profits being generated in the computer industry. By 1991, however, systems makers were getting just 20%. The market re-allocated profits to the component makers’ (‘Deconstructing the Computer Industry,’ *Business Week*, 23 November 1992: 90–96). Indeed, in the modular era it has been interdependent subsystem makers such as Intel, Microsoft and Applied Materials—whose products themselves are technologically interdependent, involve high fixed costs and consequently enjoy steep scale economics—that have captured a disproportionate share of industry profits.

Strategists often use a ‘five forces’ framework to describe where in the value chain competitive advantage and attractive profits can be built. We believe that the mechanism described here may define the dynamic, causal mechanisms behind the somewhat static characterizations of market power described in Porter’s (1980, 1985) work.

This implies, of course, that through ongoing processes of overshooting and disruption, we are likely to see yet further shifts in the locus of power to earn profit in this industry. For example, as the functionality of operating systems, microprocessors and MR heads becomes more than good enough, and as disruptive hand-held wireless computing/communication devices emerge (which are not yet good enough), it is very possible that the power to earn attractive profits will migrate away from the ‘back-end’ locations where it has resided to the stage of value-added where the end-use product is designed and assembled. We would welcome any efforts by other scholars to evaluate this hypothesis.[24]

5.4 The world automotive industry

Our final ‘case study’ is not historical, but is predictive: we will use the model to project how the structure of the world auto industry might evolve in response to the tendencies we have chronicled here. We hope that this case study will help other scholars visualize

---

One of the foremost prophets of time-based competition, George Stalk, recognized the same unattractive end-game for modular assemblers (Stalk and Hout, 1990; Stalk, 1993).

[24] We have an additional hypothesis. Assemblers of modular products, such as Compaq Computer, increasingly outsource more and more value-added to contract manufacturers. Firms such as Solectron, Celestica and Flextronics, for example, began as circuit board assemblers. They then integrated forward into assembling computer motherboards. They then advanced into assembling the entire computer; then into managing inbound and outbound supply chain logistics; and most recently, into the design of the products themselves. Why would the contract manufacturers find it attractive to integrate *into* the very stages of value-added that the computer companies found it attractive to get *out* of? The computer makers need to keep improving return on assets. When the assemblers of modular products cannot differentiate in performance or cost, the cannot improve the numerator of the return-on-assets ratio (ROA). They only have leverage over the denominator—they improve ROA—and they do this by outsourcing asset-intensive stages of value-added. This accentuates the ‘modularity trap’ that Chesbrough and Kusunoki (2001) describe.

better the implications of the mechanisms described above, and assess how they might play out in other industries that presently are dominated by large integrated firms. Theory can only be built cumulatively if scholars' explanations of cause and effect can be falsified, when used to predict what we are likely to see under various circumstances (Kuhn, 1962; Kaplan, 1986). We hope that future scholars, looking through the lenses of our model, can see anomalous phenomena as the auto industry evolves, and bring better theory to the academy.

The performance of automobiles has overshot the ability of most customers to utilize it, on several dimensions. Autos today routinely go 150,000 miles and more. They often go out of fashion before they wear out. Many car owners simply cannot utilize even longer-lasting cars. While technology enables comfortably-sized cars to travel 30 and more miles per gallon of fuel (some new hybrid gas–electric vehicles get 80 m.p.g.), consumers' rush toward less efficient sport-utility vehicles suggests that the car makers have overshot on this dimension of performance as well. Although autos are capable of cruising at speeds exceeding 90 miles per hour, traffic laws will not allow it. The list could go on.

The major disruptive innovators in the world automotive industry in the last 30 years have been Japanese firms, such as Toyota, which entered North American and European markets with low-priced offerings, and subsequently moved up-market in a classic disruptive fashion. Analysts have noted that a key tool used by these disruptive innovators to control costs and accelerate their product design cycle has been their use of a tiered supplier system (Dyer, 1996). Rather than designing and manufacturing their own components and performing all system design and assembly in-house, as General Motors and Ford traditionally had done, the disruptive Japanese innovators procured subsystems from a limited number of 'Tier 1' suppliers, such as Nippon Denso. Their more modular product architectures, and the supplier infrastructure that mirrored them, helped the Japanese disruptors bring new designs to market much more rapidly than their American and European competitors.

Just as in computers, the basis of competition in the mainstream tiers of the automobile market is changing as the functionality of cars has overshot what is actually utilized by customers. Speed to market is increasingly important (Stalk and Hout, 1990; Clark and Fujimoto, 1991). Design cycles, which often took 6 years in the 'Region A' era of the industry's history, have been shortened to 2 years today and are converging on 18 months. The ability to conveniently customize the features of each car for specific customers is emerging as a critical differentiator. For example, Toyota recently announced that its customers could custom-order cars for delivery within 5 days (Simison, 1999). The acceleration in time-to-market and improving ability to customize conveniently is being enabled by a steady modularization of the architecture.

To stay abreast of the frenetic pace of product development in the industry, General Motors has disintegrated by bundling its component-making companies into Delphi Corporation and spinning it off. Ford followed suit by spinning off Visteon.

To date, the industry seems to have evolved in a way that is quite consistent with the early steps in our model. If it continues to evolve along the path predicted by the model, we might expect the following developments in the coming years.

Ever-more-modular automobiles, comprised of increasingly standardized subsystems provided by Tier 1 suppliers, will likely take root initially at the lowest tiers of the market, amongst disruptive auto makers whose only hope for gaining a competitive edge is to introduce models faster than the competition, and who want to flatten the scale curve.[25] These are firms that will want to be able to design and assemble autos with the lowest ratio of fixed to variable costs possible.[26] We would expect auto makers in the more sophisticated tiers of the market to remain more vertically integrated, longer into the future, than would auto makers at the low end.

As the Tier 1 module suppliers experience the freedom to make trade-offs within the subsystem in order to minimize cost and optimize performance, subject to the interface constraints specified by the car companies, the internal architecture of the subsystem will become progressively more interdependent. The Tier 1 suppliers will begin to see interactions amongst the components that they were unable to see when they were fragmented suppliers of individual parts. Hence, the performance of subsystems from different suppliers is likely to become more differentiated.

At the outset, the subsystems from which the modular autos will be built are unlikely to interface with each other according to industry standards—the interface specifications will continue to be detailed by the car designers. But there will likely come a point when cars come to be designed around the interface specifications that Tier 1 suppliers articulate.

The number of suppliers of subsystems in the world industry is then likely to drop, as a result of steepening scale economics in design and manufacturing of those interdependent products. In contrast, the scale economics in the design and manufacture of cars are likely to flatten, as the job of designing and assembling modular cars becomes ever simpler. In fact, it will become increasingly possible to design and accurately simulate the performance of cars on a computer. The number of auto brands, therefore, is likely to maintain steady or even increase, as barriers to entry erode.

Ironically, even though the functionality of cars has overshot what consumers can actually utilize, car makers will continue to strive to offer products that are better than the competition. Customers will be reluctant to pay for the superfluous functionality,

---

[25] Maynard (1998) points out that even General Motors is seriously considering implementing modular designs and assembly lines at the *low* end of its product line.

[26] Chrysler's recent entry into the Brazilian automobile market is an example of this. Because Chrysler was the thirteenth company to enter the Brazilian market, it could not justify the typical investment (usually hundreds of millions of dollars) required to build a traditional assembly facility, given the initial market share it could reasonably expect to capture. In order to achieve profitability at low volumes, Chrysler's strategy has been to modularize both the vehicle design and the assembly process. A few suppliers design and build major subsystems in their own plants. They deliver these major modules to the Chrysler line, where the modules fit together in far fewer steps, with far less capital investment, than typically required (White, 1998).

but, given the choice between equally priced autos, consumers will always accept the one with better performance, even though they will pay little for it. This means that car makers that refuse to keep racing up-market will lose share and profits. Because the designs are modular, however, the only way for car companies to differentiate the performance of their products from those of competitors will be to offer the best subsystems. This means that, by definition, suppliers of subsystems will be in Region A, even as their customers, the car makers, are in Region B. Hence, the internal architecture of the subsystems will increasingly become interdependent. It will therefore be difficult for focused firms that make only one or a few of the components comprising the subsystem to survive.

As the automobile becomes progressively more modular and the subsystems more interdependent, the ability to capture a disproportionate share of the industry's profits will migrate from the car makers to the Tier 1 subsystem suppliers. This probably will not be the case with all subsystems in the automobile, however. Extraordinary profits will accrue to those suppliers whose subsystems are in Region A, which forces their designs toward the interdependent end of the architectural spectrum. Subsystems which themselves perform beyond what the car makers need, and are therefore architecturally modular, are unlikely to generate abnormally attractive profits. Hence, the strategies that integrated manufacturers such as General Motors and Ford have followed in spinning off their components operations in order to be more cost- and speed-competitive in the stage of value-added where attractive profits formerly were made, mirror almost exactly IBM's decision to out-source the microprocessor and operating system to Intel and Microsoft, so that it could continue to design and assemble personal computers.

5.5 Implications for integration and outsourcing strategies

We hope this model can add insights to two pieces of prevailing wisdom about industry structure and outsourcing. The first is about the general trend seen in most industries, where dominant, integrated firms over time give way to a horizontally stratified population of specialized firms (Chesbrough and Teece, 1996; Grove, 1996). Our contribution is that the causal mechanism that precipitates the vertical disintegration of industries may be the overshooting of the functionality that actually is utilized in certain tiers of the market. Overshooting precipitates a change in the basis of competition towards speed to market and the ability to conveniently customize features and functions. This, in turn, requires the modularization of product architectures which, finally, enables industry disintegration or deconstruction (Langlois and Robertson, 1992). This means that we would expect integrated firms to remain strong in tiers of the market that are under-served by the functionality of prevailing products, and that industries will trend toward reintegration when shifts in what customers demand cause performance gaps to emerge.

The second insight relates to the simple rule that managers and consultants use in making outsourcing decisions—that firms should outsource components or services if it is not their core competence, or if somebody else can do it at lower cost. This logic almost always makes compelling sense on the surface. But this research suggests that this logic can lead a firm to outsource those pieces of value-added in which most of the industry's profit will be made in the future—and to retain activities in which it is difficult to create enduring, differentiable advantages versus competitors. Although these hypotheses require further study, it appears that the assemblers of modular items at any stage of the value chain—whether they be end-use products, subsystems or components—are likely to struggle to achieve competitive advantage and to earn attractive profits. Attractive profitability seems to flow from the point of customer contact back through the product system to the point at which unsatisfied demand for functionality, and therefore technological interdependency, exists. Hence, these dynamics can cause the point of attractive profitability to shift from the system provider to the subsystem or component providers—from the front end to the back end to the front end again—as these dynamics work through an industry.

Address for correspondence

Clayton M. Christensen, Harvard Business School.

References

Abernathy, W. and J. Utterback (1978), 'Patterns of industrial innovation,' *Technology Review*, **50**(June–July), 40–47.

Adner, R. (1999), 'A demand-based view of the emergence of competition: demand structure and technology displacement,' working paper, Insead.

Adner, R. and D. Levinthal (2001), 'Demand heterogeneity and technology evolution: implications for product and process innovation,' *Management Science*, **47**, 611–628.

Adner, R. and P. Zemsky (2001), 'Disruptive technologies and the emergence of competition,' working paper, Insead.

Baldwin, C. Y. and K. B. Clark (1997), 'Managing in an age of modularity,' *Harvard Business Review*, **75**(September–October), 84–93.

Baldwin, C. Y. and K. B. Clark (2000), *Design Rules: The Power of Modularity*. MIT Press: Cambridge, MA.

Bass, M. J. and C. M. Christensen (2002), 'The future of the microprocessor business,' *IEEE Spectrum*, **39**(April), 34–39.

Becker, G. S. and K. M. Murphy (1992), 'The division of labor, coordination cost, and knowledge,' *Quarterly Journal of Economics*, **107**, 1137–1160.

Billington, C. and L. Fleming (1998), 'Technological evolution, standard interfaces, and new market opportunities,' *POMS Series in Technology and Operations Management*, 30–41.

Bower, J. L. and C. M. Christensen (1995), 'Disruptive technologies: catching the wave,' *Harvard Business Review*, January–February.

Chandler, A. D. (1977), *The Visible Hand*. The Belknap Press of Harvard University Press: Cambridge, MA.

Chesbrough, H. W. and K. Kusunoki (2001), 'The modularity trap: innovation, technology phase shifts and the resulting limits of virtual organizations,' in I. Nonaka and D. J. Teece (eds), *Managing Industrial Knowledge*. Sage: London, ch. 10.

Chesbrough, H. W. and D. J. Teece (1996), 'When is virtual virtuous?' *Harvard Business Review*, **74**(January–February), 65–74.

Christensen, C. (1992a), 'Exploring the limits of the technology S-curve (parts 1 and 2),' *Production and Operations Management*, **1**, 334–366.

Christensen, C. M. (1992b), 'The innovator's challenge,' unpublished DBA thesis, Harvard Business School.

Christensen, C. M. (1993), 'The rigid disk drive industry: a history of commercial and technological turbulence,' *Business History Review*, **67**, 531–588.

Christensen, C. M. (1994), 'The drivers of vertical disintegration,' Harvard Business School working paper.

Christensen, C. M. (1996), 'Patterns in the evolution of product competition,' *European Management Journal*, **15**, 117–127.

Christensen, C. M. (1997), *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Harvard Business School Press: Boston, MA.

Christensen, C. M. (2001), 'The law of conservation of modularity,' working paper, Harvard Business School.

Christensen, C. M. and E. Armstrong (1998), 'Disruptive technologies: a credible threat to leading programs in continuing medical education?' *The Journal of Continuing Education in the Health Professions*, **18**, 69–80.

Christensen, C. M. and J. L. Bower (1996), 'Customer power, strategic investment, and the failure of leading firms,' *Strategic Management Journal*, **17**, 197–218.

Christensen, C. M. and R. S. Rosenbloom (1995), 'Explaining the attacker's advantage: technological paradigms, organizational dynamics, and the value network,' *Research Policy*, **24**, 233–257.

Christensen, C. M. and R. S. Tedlow (2000), 'Patterns of disruption in retailing,' *Harvard Business Review*, **78**(January–February), 42–45.

Christensen, C. M. and M. Verlinden (1999), 'Hewlett Packard's Merced decision,' Harvard Business School case study #9-699-011.

Christensen, C. M., F. Suarez and J. Utterback (1998), 'Strategies for survival in fast-changing industries,' *Management Science*, **44**, S207–S220.

Christensen, C. M., R. Bohmer and J. Kenagy (2000), 'Will disruptive innovations cure health care?' *Harvard Business Review*, **78**(September–October), 102–111.

Christensen, C. M., T. Craig and S. Hart (2001), 'The great disruption,' *Foreign Affairs*, **80**, 80–95.

Clark, K. B. and T. Fujimoto (1991), *Product Development Performance*. Harvard Business School Press: Boston, MA.

Coase, R. H. (1937), 'The nature of the firm,' *Economica*, **4**, 386–405.

Demsetz, H. (1988), 'The theory of the firm revisited,' *Journal of Law, Economics and Organization*, **4**, 141–161.

Disk/Trend Report (1999), Disk/Trend Inc.: Mountain View, CA.

Dyer, J. (1996), 'How Chrysler created an American keiretsu,' *Harvard Business Review*, **74**(July–August), 42–56.

Fine, C. (1998), *Clockspeed*. Perseus Press: New York.

Fixson, S. (2000), 'A taxonomy development: mapping different product architectures,' working paper, Technology, Management and Policy Program, Massachusetts Institute of Technology.

Fleming, L. and O. Sorenson (2001), 'Technology as a complex adaptive system: evidence from patent data,' *Research Policy*, **130**, 1019–1039.

Grove, A. S. (1996), *Only the Paranoid Survive*. Doubleday: New York.

Henderson, R. and K. B. Clark (1990), 'Architectural innovation: the reconfiguration of existing product technologies and the failure of established firms,' *Administrative Science Quarterly*, **35**, 9–30.

Hodes, M. S. and G. W. Hall (1999), 'Home run: taking a closer look at internet mortgage finance,' Goldman Sachs Investment Research.

Iansiti, M. (1997), *Technology Integration*. Harvard Business School Press: Boston, MA.

Kaplan, R. (1986), 'The role for empirical research in management accounting,' *Accounting, Organizations and Society*, **11**, 429–452.

Klein, B., R. Crawford and A. Alchian (1978), 'Vertical integration, appropriable rents, and the competitive contracting process,' *Journal of Law and Economics*, **21**, 297–326.

Kuhn, T. (1962), *The Structure of Scientific Revolutions*. The University of Chicago Press: Chicago, IL.

Langlois, R. N. (1994), 'Capabilities and vertical disintegration in process technology: the case of semiconductor fabrication equipment,' working paper, CCC.

Langlois, R. N. and P. L. Robertson (1992), 'Networks and innovation in a modular system: lessons from the microcomputer and stereo component industries,' *Research Policy*, **21**, 297–313.

Macher, J. T. (2001), 'Vertical disintegration and process innovation in semiconductor manufacturing: foundries vs. integrated producers,' working paper, Robert E. McDonough School of Business, Georgetown University.

Maynard, M. (1998), 'GM considers switch to modular assembly,' *USA Today*, December 16, 2B.

Monroe, D. (1999), 'The end of scaling: disruption from below,' in S. Luryi, J. Xu and A. Zaslavsky (eds), *Future Trends in Microelectronics: Beyond the Beaten Path*. Wiley: New York.

Monteverde, K. (1995), 'Technical dialog as an incentive for vertical integration in the semiconductor industry,' *Management Science*, **41**, 1624–1638.

Pine, B. J. (1992), *Mass Customization: The New Frontier in Business Competition*. Harvard Business School Press: Boston, MA.

Porter, M. (1980), *Competitive Strategy*. The Free Press: New York.

Porter, M. (1985), *Competitive Advantage*. The Free Press: New York.

Porter, M. (1996), 'What is Strategy?,' *Harvard Business Review*, **74**(November–December), 61.

Rosenbloom, R. S. and C. M. Christensen (1995), 'Technological discontinuities, organizational capabilities, and strategic commitments,' *Industrial and Corporate Change*, **3**, 655–685.

Sanchez, R. and J. T. Mahoney (1996), 'Modularity, flexibility and knowledge management in product and organization design,' *Strategic Management Journal*, **17**(Winter special issue), 63–76.

Simison, R. L. (1999), 'Toyota develops a way to make a car within 5 days of a custom order,' *Wall Street Journal*, 6 August, A4.

Smith, A. (1776), *The Wealth of Nations* [Modern Library: New York, 1994].

Stalk, G. (1993), 'Japan's dark side of time,' *Harvard Business Review*, **71**(July–August).

Stalk, G. and T. Hout (1990), *Competing Against Time*. The Free Press: New York.

Stigler, J. (1951), 'The division of labor is limited by the extent of the market,' *Journal of Political Economy*, **59**, 185–193.

Stuckey, J. and D. White (1993), 'When and when not to vertically integrate,' *McKinsey Quarterly*, no. 3, 3–27.

Taguchi, G. and D. Clausing (1990), 'Robust quality,' *Harvard Business Review*, **68**(January–February), 65–75.

Teece, D. (1986), 'Profiting from technological innovation: implications for integration, collaboration, licensing and public policy,' *Research Policy*, **15**, 285–305.

Ulrich, K. (1995), 'The role of product architecture in the manufacturing firm,' *Research Policy*, **24**, 419–440.

Ulrich, K. and S. Eppinger (1995), *Product Design and Development*. McGraw-Hill: New York.

Waid, D. (1989), *Rigid Disk Drive Magnetic Head/Media Market and Technology Report*. Peripheral Research Corporation: Santa Barbara, CA.

Williamson, O. (1979), 'Transactions cost economics: the governance of contractual relations,' *Journal of Law and Economics*, **23**, 233–261.

Williamson, O. (1985), *The Economic Institutions of Capitalism*. Prentice-Hall: Englewood Cliffs, NJ.

Yin, R. (1984), *Case Study Research: Design and Methods*. Sage: Beverly Hills, CA.

Appendix 1: notes on the calculation of architectural efficiency

To measure the abilities of different companies to extract performance from any given set of components, we conducted a multivariate regression analysis of the components used in 4334 disk drives introduced in the industry between 1979 and 1997. The equation estimated in this analysis measured the extent to which the year in which a product was introduced, and the different components (represented by dummy variables) that were used, contributed to the differences in areal recording density (megabits per square inch of disk surface) of different disk drive models. Essentially, the equation derived from this analysis allowed us to estimate, on average for the entire industry, what recording density could be achieved at a given point in time with any set of components. Likewise, the coefficients in this equation measured the improvement in recording density that we would expect the average engineer in the industry to have achieved by using each new component technology.

This equation was used to estimate the areal density that each disk drive manufacturer should have been able to achieve at the time each of its models was introduced, given the set of components used in that model. The ratio of the actual recording density of the product to the predicted density was termed the *architectural efficiency* of the drive. A ratio of 1.2 meant that the company's engineers got 20% greater density out of a given set of components than was average for the industry, whereas one of 0.8 meant that the company succeeded in only getting 80% of the recording density that would have been average for the industry, given the components that were used.

Iansiti (1997) introduced the concept of 'technological yield'. This is a measure of the differences in product performance that stem from clever product design, rather than from use of superior components.

Table A1 presents the coefficients of variables in the equations that were estimated. The dependent variable in each case was the log of areal density. Following Christensen (1992a,b), the reference components were those in common use in 1979. In the database, the use of new-technology components was indicated with a system of dummy variables. The table below lists only those component, software and architectural variables with *t*-statistics during at least one period of >2.00. In addition, a few variables for unusual interfaces were not reproduced in this table, for the sake of brevity. Note that the adjusted R^2 of 0.951 suggests that these variables account for much of the variation in observed areal density among the models of disk drives designed over this period.

Appendix 2: notes on the calculation of the industry supply or scale curve in Figure 4

The purpose of this regression analysis was to calculate the steepness of scale economics in the stage of value-added involving the design and assembly of disk drives at various points in time over the industry's history. The companies whose data were used for these calculations were disk drive companies that were only engaged in design and assembly. Firms that not only designed and assembled, but also manufactured some or all of the components they used could not be included in the study, because reported costs could not be allocated accurately to the various stages of value-added. The data was drawn from *Disk/Trend Report*, as well as from the financial statements of the companies, for the years 1981–1989. The analysis could not be extended beyond 1989 because there were too few surviving firms engaged solely in the business of designing and assembling disk drives. Firms had either exited the industry, integrated into making disk drive components or integrated into making other products.

The equation that was estimated was of the form

Table A1

	Coefficient	*t*–statistic	Original technology to which the new technology's performance is compared
Constant	–0.913	–3.04	
Year of introduction	0.199	57.80	NA: this is a continuous, not dummy variable
Disk diameter	–0.00154	–9.69	none: continuous variable, in centimeters
New head and disk technologies			
MIG head	0.047	2.00	ferrite head
Thin film head	0.251	13.50	ferrite head
MR head	0.838	22.70	ferrite head
Thin film disk	0.188	9.45	particulate oxide disk
Interaction of thin film head and thin film disk	–0.295	–3.35	
Actuator technologies			
Stepper motor	–0.428	–18.30	voice coil motor
Torque motor	–0.157	–4.93	voice coil motor
Rotary actuator design	0.063	3.55	linear actuator design
Optical positioning system	–0.380	–3.75	stepper positioning
Dedicated surface servo	0.020	–1.27	stepper positioning
Embedded servo	0.169	8.67	stepper positioning
Recording/error correction codes			
2,7 RLL recording code	0.207	9.23	modified frequency modulation (MFM) code
1,7 or 8,9 RLL recording code	0.439	14.40	MFM code
0,4,4 PRML recording code	0.585	11.90	MFM code
0,6,6 PRML recording code	0.775	13.00	MFM code
Interfaces			
PC/AT	0.073	2.64	ST412 interface
SCSI	0.099	3.83	ST412 interface
SCSI2	0.197	6.43	ST412 interface
SCSI3	0.324	2.80	ST412 interface
SMD	0.356	11.80	ST412 interface
ESDI	0.214	6.96	ST412 interface
ANSI	0.226	2.53	ST412 interface
IBM	0.287	6.55	ST412 interface
Proprietary interfaces	0.113	3.54	ST412 interface
Other interfaces	0.314	10.30	ST412 interface
Other technologies			
Zone-specific bit recording rate	0.127	6.35	uniform rate regardless of distance from center
Ramp loaded heads	0.106	1.61	heads rise from surface of the disk
Number of observations	4334		
Adjusted R^2	0.951		

$$\ln(\text{cost/unit}) = B_0 + B_1(\text{year}) + B_2 \ln(\text{unit volume}) + B_3 \ln(\text{no. of product line complexity}) + B_4 \ln(\text{drive capacity})$$

The choice of these variables as the ones most likely to impact total cost was grounded in research conducted by the first author that has been synthesized in the Harvard Business School teaching case, 'Michigan Manufacturing Corporation' (HBS case no. 9-694-051).

The variables were defined as follows: *cost/unit* was calculated by dividing the total operating costs in the company, exclusive of interest and taxes, by the number of disk drive units produced. Hence, this measure is the *fully allocated product cost* for each company, for each year. *Drive capacity* is the weighted average capacity of the disk drive units shipped each year by the company. This is an important control variable, because each company's product mix was differently distributed across tiers of the market, and higher-capacity drives are more costly to produce. We expected the coefficient of this variable to be positive. *Unit volume* is the total number of disk drives shipped during the year. We expected the coefficient of this variable to be negative, positing that firms with larger production scale would enjoy lower costs, and that as any firm's production scale increased, its unit costs would fall. *Product line complexity* is the number of product families produced by the company in the year. We expected the coefficients of this variable to be positive—overhead costs per unit would increase as increasing complexity of the product line would demand higher management overheads. The definition of a product family was that used in Christensen (1992b).

We estimated coefficients for the equation for panels of years: 1981–1984 and 1986–1989. An alternative approach, to include a dummy variable for each year during this period, was not feasible because in some years there were fewer than thirty observations. The equation for the years 1981–1984, when modular architectures were just beginning to penetrate the industry, was

$$\ln(\text{cost/unit}) = 296.39 - 0.146(\text{year}) - 0.370 \ln(\text{unit volume}) + 0.126 \ln(\text{no. of families}) + 0.511 \ln(\text{weighted mean MB/unit})$$

t-statistics: (–3.44) (–4.70) (1.68) (6.23) $R^2 = 0.88$

The equation for the years 1986–1989, when modular architectures had become much more pervasive in 3.5-inch drives used in desktop computers, was

$$\ln(\text{cost/unit}) = 322.22 - 0.160(\text{year}) - 0.15 \ln(\text{unit volume}) + 0.014 \ln(\text{no. of families}) + 0.544 \ln(\text{weighted mean MB/unit})$$

t-statistics: (3.41) (–0.52) (0.12) (4.20) $R^2 = 0.88$

Several comparisons between these measurements merit mention. First, the year term is essentially a 'catch-all', whose coefficient represents the year-to-year reduction in cost attributable to engineering and product technology improvements. The relative stability of the coefficients measured in the two time panels suggests that the variables in the equations vary independently, and that probably no other important explanatory variables are missing from these estimations, which have interactions with the variables shown. The coefficient of the 'weighted average megabytes per unit' variable was similarly stable, as we would expect: adding an extra megabyte of capacity to a drive ought to result in a predictable increment to cost.

Note how the coefficient of the unit volume variable was negative and statistically significantly different from zero, suggesting rather steep scale economics in the 1981–1984 period. The coefficient was statistically insignificant in the later period, suggesting that scale economics were not a significant driver of cost: the scale curve seems to have flattened. Similarly, the product line complexity variable, which was modestly significant during the era of architectural interdependency, was insignificant when modular architectures were more pervasive—reflecting the fact that modularity facilitates increased product variety without the significant cost penalties incurred when architectures are interdependent.

One reviewer of this paper noted that because of the large standard error of the coefficient of the scale variable in the second period, we actually cannot reject, based upon statistical analysis along, the null hypothesis that scale economics might still have been steep during this period. Nonetheless, industry executives who have reviewed the work uniformly felt that there were almost no differences in cost across the five largest firms. We take this statistical analysis, therefore, to be consistent with the theory (a high proportion of variable to fixed costs flattens the scale curve), as well as consistent with the perceptions of industry executives.