March 2007

Organizing the Electronic Century

Richard N. Langlois
University of Connecticut

Follow this and additional works at: https://opencommons.uconn.edu/econ_wpapers

Recommended Citation
https://opencommons.uconn.edu/econ_wpapers/200707
Organizing the Electronic Century

Richard N. Langlois
University of Connecticut

Working Paper 2007-07
March 2007
Abstract
This paper’s title is an echo of Alfred Chandler’s (2001) chronicle of the electronics industry, *Inventing the Electronic Century*. The paper attempts (A) a general reinterpretation of the pattern of technological advance in (American) electronics over the twentieth century and (B) a somewhat revisionist account of the role of organization and institution in that advance. The paper stresses the complex effects of product architecture and intellectual property regime on industrial organization and technological change. Whereas large research-oriented multidivisional firms always played a crucial role in the industry’s history, such firms proved most adept at systemic innovation, as in the case of television. But, as in the cases of early radio and of the IBM 360 mainframe computer, the multidivisional firm was capable of bottling up within its boundaries (often through intellectual property rights) a relatively modular architecture whose “option value” such firms could not fully exploit. America’s adherence to the model of industrial research within the vertically integrated corporation arguably contributed to the demise of American consumer electronics in the 1970s and 1980s. And America’s subsequent relative success in semiconductors and personal computers — and in today’s converged digital consumer electronics — owes much to the specialized and "fragmented” character of American industry, which could take fuller advantage of competitive global value chains and of the option value of modular architectures.

**Journal of Economic Literature Classification:** L2, L63, N62, O33, O34

**Keywords:** electronics, modularity, product architecture, vertical integration.

Paper for the conference ”Has There Been a Third Industrial Revolution in Global Business?” November 16-18, 2006, Bocconi University, Milan.
Introduction.

Talk of a Third Industrial Revolution presupposes that there has been a Second.\(^1\) Alfred Chandler (2006, pp. 1-2) tells us that the Second Industrial Revolution began in the 1880s, when the railroad, steamships, electricity, and the telegraph and telephone called forth economies of scale and set in motion genuinely multinational enterprises (Chandler 1977, 1990). The revolutionary barricades were manned by a large number of integrated multidivisional firms wielding a multiplicity of technologies. By 1930, that revolution was over, leaving behind the infrastructure of the Industrial Century – the twentieth century. As Chandler (2006, p. 48) reminds us, with what one suspects is a great deal of satisfaction, 98 of the 100 largest industrial enterprises in the U. S. in 1993 had been founded by the early 1930s.

The Third Industrial Revolution, which Chandler tends to call the Electronic or Information Revolution, began just as its predecessor was ending. It would eventually generate the infrastructure for the Electronic Century now upon us. Unlike the Second Industrial Revolution, the Information Revolution bubbled up from a narrow set of technologies – the vacuum tube, the transistor, the integrated circuit, and the microprocessor – and thus involved a smaller set of players (Chandler 2001, p. 12). But the organizational outcome was identical, because it would be the same kind of large multidivisional firms that would commercialize the scientific and technological ideas of the new century. As first movers, and occasionally fast followers, these firms developed an integrated knowledge base from which they could launch innovative products.

\(^1\) Disclaimer: I am not actually interested in what industrial revolutions are, whether they exist, or whether these particular ones are correctly specified. I use the terminology only as a convenient container for my narrative.
Although a “supporting nexus” of smaller, more-specialized firms was crucial to the success of the overall industrial enterprise, it was the multidivisional firms, not the web of specialists, who did the heavy lifting. So long as the pioneering firms employed the virtuous strategy of related diversification and remained on the straight-and-narrow paths of learning the first movers had mapped out, those firms were able to enjoy economies of scale and scope and to become the perpetrators rather than the victims of creative destruction. But when the pioneers strayed from the path, and especially when they succumbed to the temptation of unrelated diversification, they stumbled and fell (Chandler 2001, 2006).

This is clearly a coherent framework for understanding the organization of technological change, and one with a good deal of appeal. Unlike many accounts of organization, especially those emanating from economics departments, Chandler’s framework stresses the knowledge and economic capabilities (Langlois and Foss 1998) that underlie production and the ways in which organization is both shaped by and shapes such capabilities. Nonetheless, one might legitimately wonder whether an explanation honed in the fires of the Second Industrial Revolution retains a sharp edge for the Third. To many observers, the most recent manifestations of the Electronic Revolution, those involving personal computers and the Internet, are notable precisely for the ways in which they have diminished the role of the large multidivisional firm as a generator of innovation and a repository of economic capabilities. There are still large firms with integrated knowledge bases and economies of scale and scope; but those firms – Intel, Microsoft, Dell, Cisco – are far more vertically specialized than firms of old; and it is far from clear nowadays whether it is these large firm or the “supporting nexus” that tends
the industry’s overall path of learning. One implication of this alternative view is a discontinuity not merely between the Second Industrial Revolution and the Third but perhaps even between the two dimensions of the Industrial Century itself, consumer electronics and computers. In this reading, we can perhaps see the path of learning in consumer electronics through the classic Chandlerian lens as a dance of large multidivisional firms like RCA, Philips, Matsushita, and Sony; but understanding the path of learning in the computer industry, especially in its more recent phases, requires an entirely different optic.

My approach will be to steer between – or rather to recombine elements from – these competing accounts of the Electronic Revolution. In accord with the “new economy” view, I will be sensitive to the ways in which changing technology and other factors have affected the nature, role, and scope of both the multidivisional firm and the “supporting nexus.” Indeed, I will concur in the view that the forces of the modern age have led to a widespread “de-verticalization” of production in this and other industries, although, as in previous work (Langlois 2003, 2004, 2007), I will locate the source of that phenomenon less in the specific demands of digital technology and the Internet than in the larger Smithian forces of specialization attendant on a growing economy, increasing globalization, and an expanding base of technological knowledge. At the same time, however, I will not consign the multidivisional firm to the dustbin of history. I will attempt to tell a tale of the electronics industry that is fundamentally Chandlerian in character, as it will focus on the development of technological and economic capabilities and on the paths of learning in the industry. Like Chandler, I will see the dynamics operating in consumer electronics and computers as essentially similar rather than as
dramatically different. The result may or may not please Chandler himself; but it will be my attempt to walk a path of learning that he, as “first mover,” has marked out.

**Capabilities and architecture.**

As Chandler notes, the consumer electronics and computer industries of the twentieth century emerged from a handful of related technologies – the vacuum tube, the transistor, the integrated circuit, and the microprocessor. These technologies arguably qualify as general-purpose technologies (GPTs), and their general-purpose character arguably accounts for the rapid pace of economic growth in and because of the electronics sector (Helpman 1998; Langlois 2002a). At the same time, however, these technologies became useful only when imbedded in larger systems like radios, televisions, and computers. In this respect, the products of the electronics industry have always been what Merges and Nelson (1994) call cumulative systems technologies. A technology is a (complex) system when it is composed of many parts each of which may draw on its own potentially distinctive knowledge base; and a technology is cumulative when “today’s advances lay the basis for tomorrow’s, which in turn lay the basis for a next round, and so on, with the sequence often progressing very far from the original invention starting place” (Merges and Nelson 1994, p. 7). In industries based on cumulative systems technologies, advances do indeed follow a path of learning.

How many of the capabilities necessary to produce a complex system will reside within the boundaries of a (large multidivisional) firm and how many will be left to other

---

2 Chandler (2005, p. 595n1) vents his annoyance at a related attempt by Lamoreaux, Raff, and Temin (2003) to create a new synthesis of business history in light of present-day organizational trends. He accuses them, perhaps a bit unfairly, of imagining a world dominated by small specialized firms and of neglecting the importance and visibility of global capitalism. So far I have escaped Chandler’s notice, or at least his line of fire.
(perhaps more specialized) firms? And – what is not the same question – which pieces of the system will a (large multidivisional) firm produce itself and which will it buy from other (perhaps more specialized) firms? That specialization has its benefits is of course an idea that predates Adam Smith. George Richardson (1972) recast the issue in terms of economic capabilities: the greater the extent to which the complementary capabilities needed to produce a (complex systems) product emerge from distinct bases of knowledge, the more expensive it will be to manage those capabilities effectively within a single organization. This suggests, with Chandler, that successful integrated knowledge bases will consist in capabilities that are related to one another. But how related? And how integrated?

In making their case against broad patent scope, Merges and Nelson (1990, 1994) point out the value of a process of innovation in which many different agents can participate, since competition in ideas can lead to rapid trial-and-error learning (Nelson and Winter 1977). In the case of a complex systems product, patents on crucial components can make it costly to assemble a state-of-the-art system; but broad patents covering large parts of the overall system can be especially damaging, as they allow the patent holders to block or retard the innovative activity of others. The same logic suggests that, even in the absence of explicit patent protection, broad and tightly integrated capabilities within a few first movers can have a similar effect. Baldwin and Clark (2000, 2006) make the argument more formally when they suggest that a given set of innovative activities – of economic experiments – are more valuable in a market than in a (large multidivisional) firm: the value of a portfolio of options in always greater than the value of an option on a portfolio.
A complex systems product is underlain by an *architecture*: a set of parts and a way of fitting those parts together. An *integral* architecture is one in which the parts depend on one another in complex and often unpredictable ways: the system is a tangle of spaghetti. By contrast, a *modular* architecture is one that regularizes the dependencies among the parts, forcing them to interact only in relatively formalized and predictable ways (Langlois 2002b). Such modularity reduces the costs of specialization and permits actors to participate with only a limited repertoire of capabilities. In the case of a complex systems product, a certain degree of modularity is inevitable to the extent that the parts of the system call on a wide set of dissimilar technological and economic capabilities.

Although the degree of modularity of an architecture is an important determinant of organizational structure, there is no one-to-one mapping between the architecture of a complex systems product and the industrial structure under which that product is actually produced. For example, a large multidivisional firm can choose to produce internally most of the components of a more-or-less-modular systems product in order to appropriate rents or to maintain the kinds of first-mover advantages Chandler describes. In such a case, the full *option value* of the architecture may lay untapped (Baldwin and Clark 2006). Alternatively, the causality may run in reverse: vertical integration may persist precisely because the firm or firms controlling the architecture are relatively insulated from competition by intellectual property protection or industrial structure (Langlois 2003). It is competition that unleashes the option value of a (potentially) modular architecture. Moreover, architecture is itself a decision variable that may be under the control of a Chandlerian first mover, and such a firm may choose a more-
integral system design as part of a strategy of rent appropriation or because the conditions of competition do not require it to do otherwise.

On one side of the ledger, then, is the benefit to modularity and vertical disintegration of modular innovation (Langlois and Robertson 1992), the rapid trial-and-error learning that comes from tapping into the larger universe of external capabilities that lie outside the boundaries of the firm. On the other side of the ledger, however, is the potential benefit to integration and integrality of architectural innovation (Henderson and Clark 1990), improvements that come from reorganizing the list of parts and the way the parts fit together. It is the importance of integrative capabilities that resonates in Chandler’s discussion of paths of learning. In the Second Industrial Revolution, first movers emerged and grew into large multidivisional firms because of their ability to create and manage the architecture of a new product (or, more typically, process) the parts for which were not initially available cheaply through arm’s-length transactions (Langlois 1992b, 2003; Langlois and Robertson 1995). Many writers continue to insist on the importance of integrative capabilities even in today’s world of greater modularity and vertical disintegration. Now as then, there remains a need for a systems integration capability (Pavitt 2003), which is why, in general, successful firms “know more than they do” (Granstrand, Patel, and Pavitt 1997; Brusoni, Prencipe, and Pavitt 2001), that is, firms retain internally capabilities not only in systems integration but also in fabricating many of the parts of the system – even though they may actually source those parts from others.

What can we learn from all this? The present-day theory of capabilities has much to say about paths of learning; but it does not prescribe that those path be trodden by
large multidivisional firms alone – or, for that matter, by small highly specialized ones. Rather, it provides a toolkit I will use in nailing together an account of how the infrastructure of the electronic century came to be organized.

**Consumer electronics.**

At the turn of the twentieth century, the strands of technology that would lead to what we now think of as consumer electronics were not yet electronic: they were electromechanical (radio) or strictly mechanical (sound and video reproduction). With John Fleming’s invention of the vacuum-tube diode in 1904, followed quickly by Lee de Forest’s invention of the triode “audion” in 1906 (Hong 2001, pp. 119-120), a genuinely electronic paradigm began to emerge. What would spur the development of that paradigm, however, was not the prospect of consumer demand but rather telecommunications, notably military telecommunications.³

**The radio.**

At the end of the nineteenth century, Guglielmo Marconi had demonstrated the possibility of wireless telegraphy using a spark-gap transmitter. By the turn of the twentieth, he had incorporated in Britain what would become Marconi’s Wireless Telegraph Company, along with a series of similar companies around the world, including one in the United States. Although there were other players, Marconi was a formidable competitor who generally refused interconnection with other networks; American Marconi held a virtual monopoly on wireless telegraphy in the U. S. A major

---

³ Unless otherwise noted, the next few paragraphs draw generally on Maclaurin (1949), Aitken (1985), Graham (1986), and Chandler (2000).
use for wireless was communication with ships at sea, something that became more significant as war broke out; the war also thrust wireless into the land-to-land business as the warring factions cut each other’s undersea telegraph cables. Understandably, this technology did not escape the notice of the Navy, which in the U.S. effectively nationalized American Marconi for the duration.

The spark-gap transmitter was a relatively crude device that spewed electromagnetic radiation indiscriminately over the spectrum. Marconi continued to depend on this technology, however, as he concentrated on the geographical expansion of his empire at the expense of technological development. By the First World War, American companies, along with some in Europe, had developed electromechanical approaches to the transmission of cleaner waveforms. But these companies, including GE, were in the equipment business not the radio-transmission business. At war’s end, the Navy, fearing dominance of a crucial military technology by what was ultimately a British company, wanted badly to nationalize wireless telegraphy.\(^4\) When it became clear that Congress wouldn’t go along, the Navy changed tack and pushed for the “Americanization” of American Marconi through an organizational alliance with GE (Maclaurin 1949, p. 103; Howeth 1963, chapter 27). GE would become the major shareholder in a new entity – the Radio Corporation of America – that would absorb most of the assets of American Marconi, including personnel and patents. The new company would then provide radio services using American (GE) equipment.

\(^4\) The case in favor of nationalization was pressed at a 1919 Congressional hearing by the Assistant Secretary of the Navy, one Franklin Delano Roosevelt (Aitken 1985, p. 386n42). A year earlier, the case against had been eloquently argued at another hearing by a rising star in the American Marconi organization, David Sarnoff. Navy ownership, Sarnoff argued, would have a chilling effect on innovation (Graham 1986, p. 36).
RCA received its Delaware charter in 1919. Within the next two years, a number of significant players joined in with capital and technology: AT&T, United Fruit, and Westinghouse. In part, RCA attracted these partners because of the systemic nature of the service. In the case of AT&T, for example, RCA needed access to AT&T’s lines as feeders for its service. The more compelling motive, however, was intellectual property. RCA was a way to pool the many patents that would be needed to produce the complex systems products of radio transmission and reception.\(^5\) Seeing vacuum tubes as important for amplifiers in its wire-based transmission system, AT&T had purchased the de Forest triode patent. Marconi had acquired the Fleming diode patent. Before the U. S. entered World War I, development of vacuum-tube technology had been at a standstill as the two firms dueled in court in a classic case of blocking patents\(^6\) (Merges and Nelson 1990, pp. 892-893; Aitken 1985, pp. 248-249). By 1920, RCA controlled both patents. United Fruit, which had developed wireless technology for use with its extensive fleet of ships, contributed patents and facilities, thereby also ridding itself of concerns far away from its core competences. The last piece of the puzzle was Westinghouse, America’s other electrical-equipment giant, which contributed among other things the rights to the crucial heterodyne principle.\(^7\) As a result of this agglomeration, as well as numerous international cross-licensing agreements, “RCA obtained rights to over 2,000 issued

\(^5\) A 1919 Navy memorandum had “found that there was not a single company among those making radio sets for the Navy which possessed basic patents sufficient to enable them to supply, without infringement, … a complete transmitter or receiver.” (Quoted in Maclaurin (1949, p. 105).) Maclaurin (1949, p. 97) also reports that there were 20 major issues of patent infringement between AT&T and GE between 1912 and 1926.

\(^6\) During World War I, the U. S. Government effectively inactivated all radio-related patents for the duration and assumed financial responsibility for any resulting infringements (Reich 1977, p. 214).

\(^7\) The heterodyne technique involves adding two waveforms in order to produce two new signals at different wavelengths. The superheterodyne receiver became the eventual dominant design in radio.
patents, including practically all the patents of importance in the radio science of the day” (Maclaurin 1949, p. 107). It remains an open question why all this took the form of an equity joint venture rather than a contractual patent pool alone, though most commentators point to the role of the government in pushing for the development of what they saw as a national champion in radio technology (Maclaurin, pp. 100ff). In the words of founding chairman Owen D. Young, RCA for its part “was anxious to create an industry in which competition would be ‘orderly and stabilized’” (Maclaurin 1949, p. 105).

Wireless telegraphy was RCA’s original business model. Everyone recognized that vacuum tubes would eventually be important in this enterprise, even if they were not initially capable of providing the power levels of the electromechanical alternator. Vacuum tubes could produce a cleaner waveform, something of great use if one were to modulate the signal to transmit and receive sound and voice not just dots and dashes. But few at the time saw any potential in sound broadcasting. There was, however, one crucial exception, and even he underestimated its potential. David Sarnoff was a Russian immigrant who rose through the ranks from office boy and key operator to a management position at American Marconi, where he had conceptualized what he called a “Radio Music Box” (Graham 1986, p. 32). The business model was to make money from selling receivers to a mass market while providing broadcasting content free of charge as a kind of advertising. American Marconi had no interest in the idea; but, when Sarnoff moved over to RCA, he eventually found a warmer reception.
<table>
<thead>
<tr>
<th>Year</th>
<th>Number</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1922</td>
<td>100</td>
<td>$60,000</td>
</tr>
<tr>
<td>1923</td>
<td>550</td>
<td>136,000</td>
</tr>
<tr>
<td>1924</td>
<td>1,500</td>
<td>358,000</td>
</tr>
<tr>
<td>1925</td>
<td>2,000</td>
<td>430,000</td>
</tr>
<tr>
<td>1926</td>
<td>1,750</td>
<td>506,000</td>
</tr>
<tr>
<td>1927</td>
<td>1,350</td>
<td>425,600</td>
</tr>
<tr>
<td>1928</td>
<td>3,281</td>
<td>690,550</td>
</tr>
<tr>
<td>1929</td>
<td>4,428</td>
<td>842,548</td>
</tr>
<tr>
<td>1930</td>
<td>3,827</td>
<td>496,432</td>
</tr>
<tr>
<td>1931</td>
<td>3,420</td>
<td>300,000</td>
</tr>
<tr>
<td>1932</td>
<td>3,000</td>
<td>200,000</td>
</tr>
<tr>
<td>1933</td>
<td>3,806</td>
<td>300,099</td>
</tr>
<tr>
<td>1934</td>
<td>4,084</td>
<td>350,000</td>
</tr>
<tr>
<td>1935</td>
<td>6,027</td>
<td>370,000</td>
</tr>
<tr>
<td>1936</td>
<td>8,248</td>
<td>500,000</td>
</tr>
<tr>
<td>1937</td>
<td>8,065</td>
<td>537,000</td>
</tr>
<tr>
<td>1938</td>
<td>6,000</td>
<td>350,000</td>
</tr>
<tr>
<td>1939</td>
<td>10,500</td>
<td>375,000</td>
</tr>
<tr>
<td>1940</td>
<td>11,800</td>
<td>584,000</td>
</tr>
<tr>
<td>1941</td>
<td>13,000</td>
<td>610,000</td>
</tr>
</tbody>
</table>

Table 1
Sales of home broadcast radio sets (in thousands).

Source: Maclaurin (1949, p. 139).

We are accustomed nowadays to the rapid penetration of new technologies – the DVD, the cell phone, the Internet – and maybe accustomed to thinking of such rapid penetration as unique to our age. But the speed with which the American home adopted the radio was on a par with anything our age has to offer; the penetration of the personal computer proceeded at a snail’s pace by comparison. Sarnoff apparently thought that the total cumulative demand for radios would be about a million sets at $75 each. In the event, the industry sold $60 million worth in the first year (1922) alone; $136 million in

---

8 In the first two years after introduction, IBM sold a total of 750,000 personal computers, and that represented 26 per cent of the market for PCs (Langlois 1992, p. 23).
1923; and $358 million in 1924. RCA's own sales were $11 million in 1922; $22.5 million in 1923; and $50 million in 1924. (See Table 1.) Over the next decade, 60 per cent of American homes came to possess radio receivers (Scott 2001). Sarnoff saw his sales staff burgeon from 14 people in 1921 to 200 nationwide offices the next year – and saw himself catapulted into the position of vice-president and general manager (Graham 1986, pp. 38-39).

Since radio constitutes what economists now call a hardware-software network, what actually catalyzed the takeoff of radio was less RCA's entrance into receiver manufacturing than the impetus RCA's entry gave to the launching of broadcast stations. The broadcast of voice and music probably goes as far back as 1906, all in the context of amateur radio, which constituted a large "hobbyist" sector akin to what would later drive the early personal computer industry.

By the 1920's wireless had become the hobby of thousands of young Americans. No other modern industry has been supported by so many ardent participants. It is hard today [i. e., in 1949] to recapture the spirit of this period: amateur clubs were started in every state, comprising all types and classes – schoolboys, professors, electricians, and ex-servicemen who had operated radios during the war. Radio was a new toy, not only technically interesting, but the means by which people could reach out into unknown regions and communicate with new-found friends. (Maclaurin 1949, p. 112.)

In 1920, Westinghouse set up a radio station (under the call sign KDKA) at its East Pittsburgh plant to cater to this ham radio market and earn some cheap good will. The effect was to alert other commercial enterprises, including AT&T and RCA, to the potential that already existed because of amateur radio (Aitken 1985, pp. 469ff.). The number of broadcast stations on the air leapt from five in 1921 to 556 in 1923, leveling
off to 765 by 1940 (Sterling and Kittross 1978, cited in Scott 2001). Most early stations did indeed broadcast for free to earn goodwill (and maybe sell receivers); but by 1927, the tide had turned in favor of the advertising-revenue model.

Most of the first commercial producers of radio receivers were hobbyists and garage-shop operations. Between 1923 and 1926, by one estimate, an average of 187 new firms entered the business every year, most of whom failed quickly (Maclaurin 1949, p. 134). Despite its formidable capabilities, RCA was not in a good position to compete on price with the garage-shops, as it had costly and often unwieldy supply relations with GE and Westinghouse for parts. As the architecture of the radio receiver matured, RCA and its owner-suppliers struggled with the problem of standardization for parts like vacuum tubes (Graham 1986, pp. 39-40). Like personal computers decades later, radios were in fact relatively inexpensive to assemble; and increased standardization and the emergence of a dominant design quickly eroded the rents one could earn from selling assembled receivers. RCA initially tried to extract rents at the level of the vacuum tube, since this was indeed the high-tech core of the radio; but they did this by insisting on full-line forcing and exclusive dealing in their arrangements with distributors, practices of dubious value that were in any event struck down by the Federal Trade Commission (Graham 1986, p. 40; Chandler 2001, p. 19). In the late 1920s, RCA’s market share had slipped to between 18 and 20 per cent (Chandler 2001, p. 21).

By 1927, however, courts had affirmed the validity of RCA’s dominant patent portfolio, which opened the door to what would be the company’s strategy for the next three decades: package licensing. Chandler (2001, p. 18) takes pains to point out that by the end of the roaring ‘twenties, garage-shop radio manufacturers had faded away in
favor of significant firms founded before 1920 in related technologies like batteries, automotive equipment, telephone equipment, or electrical equipment.\(^9\) The superior capabilities of incumbents may indeed have won out in the end; but in this case the demise of the newcomer was mostly the result of RCA’s patent policy: only 25 large assemblers would initially have the rights to RCA’s patents, in exchange for a sizeable royalty of \(7\frac{1}{2}\) per cent plus back damages for infringement.\(^10\) The licensing was a package in the sense that an assembler had to pay royalties on RCA patents for all relevant parts of the radio even if the assembler didn’t use all those parts.

![Figure 1. Effect of RCA patent licensing on entry into radio](image)

**Source:** Maclaurin (1949, p. 134).

Although RCA did later extend the deal to others and reduce royalty demands somewhat, it was nonetheless RCA’s control of the patent portfolio that gave shape to the

---

\(^9\) The two exceptions were startups Zenith and Raytheon.

\(^{10}\) RCA had initially wanted to limit licenses to customers whose royalties would amount to at least $100,000, though this minimum was never enforced (Maclaurin 1949, p. 135).
radio industry. In part, this meant more rapid consolidation. More importantly, as Graham has noted, “the most enduring consequence of the [package-licensing] policy was that it made it uneconomic for most other companies to do radio-related research, because they could not recoup the investment. This left control of the rate and direction of technological change in the radio industry largely in the hands of RCA. For RCA, the effect was to make licensing fees the major payoff of its research activity. RCA was effectively in the business of selling research”\(^\text{11}\) (Graham 1986, p. 41). Thus in radio it was not the case that an integrated path of learning within a large firm gave rise to innovation; it was rather that innovation, channeled within a particular structure of property rights, contained the path of learning within a single large firm. One might indeed wonder whether, far from representing the optimum optimorium of capability-building, RCA’s integrated structure failed ultimately to tap the option value of what was potentially a powerful modular architecture.

Sarnoff understood that, since the firm was not a cost leader and had essentially no source of rents in the receiver value chain other than its patents, RCA needed to look elsewhere for future sources of rents. This meant pouring money into research. It also meant a move away from hardware to “software”: broadcasting and content. Taking over the broadcasting assets of AT&T, RCA famously created the National Broadcasting Company (NBC) in 1926, from which the American Broadcasting Company (ABC) would eventually spin off in 1942.

\(^{11}\) On this point see also Reich (1977).
Sound recording.

In addition, Sarnoff began to see the relevant architecture as one of “consumer electronics” broadly rather than of radio as an appliance.\(^{12}\) The core of all consumer electronics in the era was the vacuum-tube amplifier, which could be linked to electromagnetic speakers to reproduce sound. The amplifier could receive its input from a radio tuner. But it could receive other inputs as well. At the system level, then, consumer electronics in the age of sound was potentially highly modular, and at the high end explicitly so. The principal source of input to an amplifier other than radio was of course pre-recorded sound. In 1930, Sarnoff purchased the Victor Talking Machines Company, a prominent maker of phonographs based on the techniques of Edison.\(^{13}\) Initially, the phonograph was entirely mechanical: a needle picked up vibrations in the tracks of a recording and a horn then amplified the sound acoustically. But as technology advanced, it became possible to capture the vibrations electronically and transmit them to a vacuum-tube amplifier. The research department at Western Electric (later to become Bell Labs) developed electric sound recording in the early 1920s, but Victor initially refused to adopt the technology.\(^{14}\) A company called Brunswick was first to market with an electronic phonograph it called the Panatrope. Sarnoff clearly understood the potential of such technological convergence, and RCA began to offer not only electronic Victrolas but also a device called the Duo Jr., a $9.95 record player that plugged into an existing radio. Smaller producers of audiophile equipment pursued an even more forceful

\(^{12}\) This paragraph and the next draw on Robertson and Langlois (1992).

\(^{13}\) At the same time, Sarnoff also engineered the sale of both GE’s and Westinghouse’s stakes in RCA, thus bolstering the company as an independent entity (Chandler 2001, p. 22).
modular strategy, offering (and often specializing in) separate amplifiers, pre-amplifiers, tuners, phonographs, and speakers. A principal element of the standardized interface within this modular system is still today called an RCA plug.

Because of the integrated structure of RCA, however, meaningful competition by specialists was decidedly limited. But there did emerge one relatively integrated rival; and one can argue that the ensuing competition did much to spur technological change in the industry. Columbia began its life in the late nineteenth century as a distributor of Edison’s phonographs and cylinders in the Washington, DC area (hence the name), and by the 1920s was Victor’s principal rival in phonographs. Columbia also owned a small network of radio stations that had formed as an outlet for talent snubbed by Sarnoff’s NBC. In 1927, a young Philadelphia cigar maker called William Paley bought the whole operation on the strength of his enthusiasm for radio advertising. The Columbia Broadcast System (CBS) competed with RCA’s networks and was also integrated into phonograph records. An excellent example of the benefits of Columbia’s competition with RCA is the famous battle of the speeds. Adapted to the acoustic phonograph, standard records of the time were made of shellac and spun at 78 rpm. The new electronic technology suggested slower speeds and a new material, vinyl. RCA experimented with 33-rpm recording, but was unable to increase playing time significantly and abandoned the project at the onset of the Depression. Columbia took up

---

14 Supposedly on the grounds that consumers were accustomed to the tinny sound of the Victrola and would find the new sound unpleasant (Robertson and Langlois 1992, p. 330).

15 In fact, the Columbia record label had spun off from CBS after Paley purchased the Columbia networks. Columbia merged with the British firm Gramophone (controlled by RCA) in 1931 to form EMI, but, because of antitrust concerns, the American assets of Columbia were not part of the deal. CBS eventually reacquired the Columbia label in 1938.
the idea under the direction of Peter Goldmark and Edward Wallerstein, the latter the erstwhile general manager of RCA’s Victor Division. As a result, Columbia stole a march on RCA, whose attempt to retake the standard with the large-spindle 45-rpm disk fell short. After World War II, phonographs adopted a gateway technology that enabled a user to play records at all three speeds.

*Television.*

Advancements in radio and phonograph technology proceeded largely in modular fashion through improvements in components. As many of these, notably vacuum tubes, were general-purpose technologies, their manufacturers benefited from scale and learning economies from demand – including military demand – outside of the consumer sector. But the next big advance in consumer electronics required innovation that was much more systemic in character.

Dating back to the nineteenth century, there had been numerous attempts to transmit pictures. All of these required some kind of electromechanical apparatus, and the results were never satisfactory. By the mid-twenties, Vladimir Zworykin was making progress on a fully electronic method of scanning and transmitting images. His work attracted Sarnoff’s attention, and RCA began funding the work, first at Westinghouse and then at RCA’s own facilities. Whereas radio did not strictly depend on vacuum tubes and could benefit from the independent development of tubes and other components, television depended on the simultaneous development of a design architecture (for both transmitter and receiver) and many new and specialized components, including highly complex electron tubes. Moreover, television raised many more issues of technical
standardization than had radio. Zworykin evidently believed that television could in fact develop in an incremental, modular, and decentralized way (Graham 1986, p. 53), but he grossly underestimated the resources that development would require. And, as the rapid development of technology made early standard-setting problematic, it is likely that Sarnoff was right to see the technology as an ideal fit with the model of the well-funded industrial research laboratory.

Although entrepreneurial startups – like those created by Philo Farnsworth and Alan B. Du Mont, who had invented technology parallel to Zworykin’s – did attempt to develop television, it was more substantial enterprises like CBS, Philco, and Zenith who provided a challenge to Sarnoff, even if these competitors were obliged to license much of the technology from RCA. Far more than up-front costs, it was, however, the systemic nature of the innovation in its early (or “pre-paradigmatic”) stages that created what Chandler (2001, p. 25) refers to as barriers to entry. The initial dearth of programming reduced television’s attractiveness to early adopters, as Sarnoff discovered after the first RCA TVs rolled out at the 1939 World’s Fair. But RCA’s integration into a broadcast network placed it a more favorable position than competitors, who had banded together to push a standard different from RCA’s. At the behest of the Federal Communications Commission (FCC), the radio manufacturers trade association formed the National Television Standards Committee (NTSC), which, in 1941, promulgated standards largely identical to those RCA favored.

TV broadcasting began in earnest later that year – just in time for the American entry into World War II and the consequent ban on the manufacture of commercial television equipment. Nonetheless, the war proved immensely beneficial for the
development of the television. The military demanded technology closely related to that used for television, notably the cathode-ray tubes crucial to radar. As a result, the end of the war presented commercial industry with numerous advances in both product and process technology, including techniques for mass producing cathode-ray tubes that halved the post-war price of TVs (Graham 1986, p. 59). Component makers, working with glass-making firms, continued the trajectory of improvement in product and process after the war (Graham 2000, p. 144). A year after the war ended, RCA was selling a mass-market TV set bundled with a contract for installation and service\textsuperscript{16} (Graham 1986, p. 60).

RCA encouraged competition in the production of sets, even to the extent of holding technical seminars for competitors.\textsuperscript{17} This ought not to be surprising, as RCA was in a position to benefit from its portfolio of television patents, its position as key producer of a crucial bottleneck component (picture tubes), and its sale of complementary “software” through the NBC network. Indeed, in the early 1950s, NBC turned handsomely profitable (Graham 1986, p. 60). In the period 1952-1956, RCA received some $96 million in TV patent revenues – some 77 per cent of all industry revenues (Levy 1981, p. 124) – which Levy (1981, p. 162) estimates to have raised the price of televisions by 2.26 per cent. As it had in the case of radio, RCA licensed its TV patents

\textsuperscript{16} This was arguably the cheapest way to provide what consumers really want – reliable television services – when technology is both unreliable and unfamiliar to consumers. Since reliability was the central issue, and since knowledge of television technology was not widely diffused to independent dealers, it was cheaper to provide such contracts through vertical integration (as RCA did) than through contracts with independents.

\textsuperscript{17} “In 1947, Frank Folsom, then president of RCA Victor, invited representatives of all other television makers to visit the RCA plant in Camden, New Jersey. He then proceeded to give them a tour of RCA’s production facilities and presented each one of them with a copy of the blueprints for RCA’s most popular television model.” (Levy 1981, p. 129, citing Fortune, September 1948, p. 81.)
(as well as its patents in other technological areas) as a package – until a 1958 antitrust consent decree mandated a patent-pool among American (but not foreign) producers.

The American public adopted the TV with an enthusiasm that rivaled its earlier infatuation with radio. Production ramped up to three million units by 1949, peaking at 7.79 million in 1955, and then declining slightly as the market for black-and-white sets reached saturation (Levy 1981, p. 99). In 1950, five years after the War, nine per cent of American homes could boast a television set; after 10 years, 64.5 per cent could; and by 1960, the figure was 87 per cent (Levy 1981, p. 116). Klepper and Simons (2000) document the resulting flurry of entry of new firms into the market for black-and-white TV. American manufacturers peaked at 92 in 1951, after which a shakeout reduced the number to 38 in 1958.18 Of these RCA and Zenith accounted for about a third of sales in the ‘fifties, and the top four firms held more than half the market (Levy 1981, p. 86).

The move to color television required systemic innovations that were more expensive and arguably more complex than those of monochrome TV, and the attendant standardization issues were equally daunting. Here again, RCA took the lead (Levy 1981; Graham 1986; Chandler 2001). Systemic innovation of this sort was well suited to RCA’s corporate research apparatus, which had grown in capability through war-related research as well as commercial TV research. And, once again, RCA’s possession of NBC reduced the coordination costs of launching a product with substantial network effects. But, as happened in the battle of the speeds, CBS quickly emerged as formidable competitor, having since added tube manufacture to its portfolio of broadcasting, records,

18 Klepper and Simons (2000) also show that the producers who survived the shakeout almost all tended to be those with prior experience in radio.
set assembly, and corporate research (Graham 1986, p. 61). In 1947, Goldmark petitioned the FCC to approve a partially mechanical system that CBS had been tinkering with since before the war. RCA was at work on a fully electronic system compatible with the existing black-and-white standard, but the company couldn’t work out the bugs fast enough to forestall FCC approval of the CBS system in 1951. Unfortunately for CBS, however, the Korean war put a hold on production of color (but not black and white) technology. By war’s end, growth in black-and-white sales had made compatibility a bigger issue; moreover, the likes of Zenith, Philco, and Philips much preferred an all-electronic system.

Another NTSC convened, and RCA not CBS earned the imprimatur. The latter slunk off to concentrate on broadcasting and records, leaving the field largely to RCA. But the technology remained expensive and finicky, and RCA and was forced to absorb high overheads in research, in production, and at NBC, overheads on which competitors chose to free ride until color reached its tipping point. Unsurprisingly, the rate of penetration of color TV was much slower than that of black-and-white: five years after introduction, only half a percent of American homes had a color set; ten years after, only 2.9 per cent did (Levy 1981, p. 115). Nonetheless, RCA was eventually able to capture rents from color TV through the production of sets, the sale of picture tubes (of which it was initially the sole producer), and the broadcasting of color programs by NBC.19

19 The 1958 antitrust consent decree limited RCA’s ability to collect rents from color TV patents, as these became subject to a patent pool in which all comers could dip for free so long as they tossed in some of their own patents (Levy 1981, p. 159).
Figure 2

(Thousands of units.)

**Source:** Levy (1981, p 99).

**Note:** U. S. production includes units manufactured in the U. S. by foreign-owned firms.

So far we have seen evidence to support Chandler’s insistence on the importance of integrated capability-building in large research-equipped multidivisional firms. Both the monochrome and the color TV – if not, however, the radio – were systemic innovations well suited to early development through strong corporate research; and all three technologies partook of network externalities that could be partly internalized in an integrated structure. But we have also seen evidence of the importance to technical advance of vibrant competition and of innovation at the level of suppliers.

*The demise of American consumer electronics.*

The major integrated firms, notably RCA and Zenith, continued to provide most of the product innovations in television through the 1970s (Klepper and Simons 1997, p. 421),
though much process innovation came from suppliers and from advances elsewhere in electronics, like the wave-soldering techniques developed in connection with military applications (Levy 1981, pp. 64ff; Klepper and Simons 1997, p. 428). As the technology of television, including color television, began to mature in the 1960s and 1970s, RCA’s rents from production of color picture tubes began to decline as other firms entered tube production, and NBC’s early lead in color programming dissipated as all three networks began broadcasting almost entirely in color. Advantage began to shift to those who could produce receivers cheaply. And that increasingly meant foreign, especially Japanese, firms. (See Figure 2.)

With the exception of Sony, which pursued a strategy of product innovation akin to that of American firms like RCA and Zenith, most Japanese firms entered the American market as low-cost producers dependent on American technology, notably that of RCA (Levy 1981, p. 97). They concentrated on black-and-white sets, portable sets, and private-label production for retailers like Sears, eventually moving up to higher-end branded products. Figure 2 understates the penetration of foreign firms into the American market, as many “American” producers – indeed half of all American producers in 1979 (Levy 1981, p. 82) – were actually subsidiaries of foreign (mostly Japanese) firms.20 By 1971, Sony had set up an assembly plant in California, with a picture-tube plant to follow in 1974. In the same decade, Sharp and Hitachi also set up de novo American facilities. Meanwhile, many existing American producers were gobbled up by foreigners. In 1974, Matsushita bought Motorola’s television business. The same year Philips acquired Magnavox, and by 1981 had picked up the remnants of
Philco and Sylvania. In 1976, Sanyo formalized its growing hold on Sears’s OEM business by buying out Warwick, which had been Sears’s American assembler. In 1986, the mighty RCA sold its television operations (including NBC) to its quondam parent GE, which promptly swapped the electronics part for the medical-technology business of France’s Thomson. Zenith, the last domino, fell in the early 1990s (Perry 1988).

Why? Chandler (2001) lays great stress on the strategic mistakes. RCA (and other integrated American firms) strayed from the virtuous strategy technological development and related diversification that would have maintained paths of learning on which they had originally embarked. Instead, they succumbed to the temptation of conglomerate diversification, thus ultimately destroying the integrated knowledge base on which success depended. It is certainly true that RCA went in for unrelated diversification in a big way. The process began under David Sarnoff in 1966 with the purchase of Random House and continued apace after the accession of David’s son Robert to the helm of RCA in 1968. Acquisitions included golf equipment, car rental, paper manufacture, frozen food, real estate development, and carpet manufacture.

Economists tend to be less inclined than historians or management scholars to see strategy as an independent causative force, especially when all firms in the environment seem to be adopting the same one. The strategy of massive unrelated diversification in this reading is merely the proximate cause rather than the ultimate cause. As Michael Jensen (1986) has taught us, unrelated diversification is one possible symptom of “free cash flow,” corporate windfalls that allow managers to pursue their own interests and

---

visions without the short-run discipline of product markets and financial markets. Although it is not often remarked on, the heyday of the large American multidivisional firm coincided with a period of relative economic isolation. The Depression, tariffs, and wars of the first half of the twentieth century constitute what economic historians now see as a massive collapse of nineteenth-century globalization (James 2001; Bordo, Taylor, and Williamson 2003). And, with the destruction of the German, Japanese, and other economies in World War II, that isolation continued for the better part of two decades. In a general sense, then, the post-war golden age of the large multi-divisional American firm was one in which competition was relatively relaxed by later globalized standards – and in which managers found themselves with sources of free cash flow. The resulting (relatively) slack environment not only encouraged diversification (as Chandler insists) but also reinforced the multidivisional form itself, a form of which unrelated diversification is the logical if extreme extension, and arguably isolated that form from economic realities to which it was increasingly ill adapted (Langlois 2003, p. 370-371).

In the case of RCA, as we saw, free cash flow came importantly from patent licensing. In addition, government, especially military, contracts came to prop up RCA’s research laboratories, thus increasingly insulating them from contact with the market. Both the post-war federal research climate and RCA’s own addiction to patent royalties moved the labs in the direction of more basic research (Graham 1986, pp. 68-71). Sarnoff believed that profit lay in continuing RCA’s founding strategy of staying abreast of – or ahead of – all knowledge in electronics. Thus RCA’s ultimately disastrous foray into digital computers.
As Adam Smith might have predicted, however, this strategy would become increasingly problematical as both the extent of such knowledge and the extent of the market expanded after World War II. As technological leaders in electronics, American firms found themselves with a broad menu of product and process options; and, because of the ultimate limit of capabilities within even a large organization, those firms became the prey of competitors who could pick off and specialize in pieces of technology or market. However valuable the integration-cum-research structure had been in the generation of systemic innovations like television, it became a liability in a globalizing world that could take advantage of outsourcing, low-cost production abroad, and withal the option value of increasingly mature and thus relatively modular technology. Moreover, for Japan in particular, specialization in production led to the creation of technological, organizational, and institutional approaches (notably in miniaturized electronic and electro-mechanical devices) that constituted a kind of general-purpose technology that could be applied to newly emerging devices like video-tape and optical-disk storage peripherals. Importantly, as we will see again presently, the diffuse focus of American systems houses like RCA, GE, and Philco made it difficult for them to compete with specialized American firms in the fabrication of semiconductor devices and with Japanese firms in the incorporation of such devices into electronic systems (Klepper and Simons 1997, p. 421).

Digital technology.

Even though semiconductors and computers were born in the years immediately following the second world war, their institutional origins were quite different. The invention of the computer involved both universities and direct government research
funding. By contrast, the transistor, the basic building block of the semiconductor industry, emerged from private research at AT&T’s Bell Labs. Because of its success and its secure status as the nation’s telephone monopoly, AT&T was able to pursue a policy of research that, while arguably more focused toward commercial ends than basic research at universities, was nonetheless willing to indulge basic science and to envisage a research agenda quite far from commercial fruition.\footnote{For a classic account of how the research environment at Bell Labs led to the transistor, see Nelson (1962).}

*The transistor.*

The Bell System was facing even more imminently the long-run problem that had motivated the acquisition of the de Forest vacuum-tube patent: the difficulty of expanding a switching system based on electromechanical relays. By the 1930s, Mervin Kelly, the research director at Bell Labs, was voicing the opinion that electromechanical relays would eventually have to be replaced by an electronic alternative in order to handle the growing volume of traffic. William Shockley, one of the three Bell scientists to receive the Nobel Prize for the transistor, was impressed by this observation, and believed that the objective would be best realized with solid-state technology (Shockley 1976).

Bell Labs announced the transistor in December of 1947.\footnote{For detailed histories of the invention of the transistor, see Braun and Macdonald (1978), Morris (1990), and Nelson (1962).} Almost immediately transistor technology began spilling out to other firms. This was not, however, a process in which slippery knowledge leaked unintendedly to others but rather a deliberate and systematic attempt by AT&T to disseminate know-how through inexpensive licenses,
technical symposia, and site visits (Tilton 1971, pp. 75-76; Braun and Macdonald 1978, pp. 54-55). The main driver of this policy was the consent decree AT&T had just signed with the Antitrust Division of the U. S. Justice Department, which specified how the company was to treat technology outside the scope of the company’s primary mission.\textsuperscript{23}

But there is also reason to think that AT&T pursued a strategy of dissemination because, like RCA in the case of color television, the company saw profit in the widespread adoption of the technology. AT&T was still primarily concerned with the usefulness of transistors to its own line of business, telephone switching. The company believed that if it allowed access to the transistor, telephony would reap the benefits of spillovers from the development of the capabilities of others in the electronics industry to an extent that would outweigh the foregone revenues of proprietary development\textsuperscript{24} (McHugh 1949; Bello 1953; Braun and Macdonald 1978, p. 54; Levin 1982, p. 76-77).

Unlike the triode vacuum tube, which had been entangled in patent litigation and then formed part of RCA’s onerous package licensing, the transistor became easily available at relatively low royalties. The result was a large cohort of entrants (Mowery and Steinmueller 1994). Existing producers of vacuum-tubes, as well as Bell Labs itself, continued to be major sources of transistor innovations through the 1950s, especially in the realm of process and materials. The work of this period led ultimately to a pivotal innovation that did allow for rapid experience-based improvements and cost reductions:

\textsuperscript{23} AT&T’s strategy of dissemination may also have been motivated in part by a desire to preempt any thought the military might have had of classifying the technology (Levin 1982, p. 58).

\textsuperscript{24} An AT&T vice president put it this way. “We realized that if this thing [the transistor] was as big as we thought, we couldn’t keep it to ourselves and we couldn’t make all the technical contributions. It was to our interest to spread it around. If you cast your bread on the water, sometimes it comes back angel food cake.” Quotation attributed to Jack Morton, in “The Improbable Years,” \textit{Electronics} \textbf{41}: 81 (February 19, 1968), quoted in Tilton (1971, pp. 75-76).
the planar process, a development arguably responsible for the increasing-returns trajectory upon which the semiconductor industry now finds itself. But the planar process was not developed by Bell Labs or by any of the established vacuum-tube firms. Instead, in a pattern characteristic of the American semiconductor industry, the new approach was developed by a small start-up organization.

Among the many Bell Labs researchers who had struck out on their own in the 1950s was Shockley, who returned home to the San Francisco peninsula to found Shockley Semiconductor Laboratories. Apparently prompted by dissatisfaction with the company’s orientation toward product breakthroughs at the neglect of the commercially richer area of process technology (Braun and Macdonald 1978, p. 84; Holbrook 1999), eight of Shockley’s team defected in 1957, and, with the backing of Long Island entrepreneur Sherman Fairchild, founded the semiconductor division of Fairchild Camera and Instrument Corporation. The Fairchild group mounted an ambitious plan to produce silicon mesa transistors using technology developed at Bell Labs (Malone 1985, p. 88; Lydon and Bambrick 1987, p. 6). In attempting to overcome some of the limitations of this transistor design, one of the eight defectors, Jean Hoerni, found a way to create a device by building up layers on a flat surface — a “planar” device (Dummer 1978, p. 143; Braun and Macdonald 1978, p. 85; Morris 1990, p. 38). The planar structure made it easy for Fairchild to devise a way to replace the mesa’s clumsy wires with metal contacts deposited on the surface.
By 1961, two Americans, Robert Noyce of Fairchild and Jack Kilby of Texas Instruments (TI), had created prototype ICs. Unlike Kilby, who had started with the monolithic idea and then sought to solve the problem of fabrication and interconnection, Noyce began with a process for fabrication and metallic interconnection — the planar process — and moved easily from that to the idea of the integrated circuit. Under pressure from the industry, TI and Fairchild forged a cross-licensing agreement in 1966 under which each company agreed to grant licenses to all comers in the range of two to four per cent of IC profits (Reid 1984, pp. 94-95). This practice served to reproduce and extend the technology-licensing policies of AT&T, again broadly diffusing the core technological innovation to all entrants and thereby reasserting the principle that innovative rents should flow to those who could commercialize and improve upon the key innovation.

As important as the innovation of the IC was, the planar process is arguably the more important technological breakthrough, not merely because it underlay the IC but because it provided the paradigm or technological trajectory the industry was to follow. By either etching away minute areas or building up regions using other materials, semiconductor fabrication alters the chemical properties of a “wafer,” a crystal of silicon. Each wafer produces many ICs, and each IC contains many transistors. The most dramatic economic feature of IC production is the increase in the number of transistors that can be fabricated in a single IC. Transistor counts per IC increased from 10 to 4,000 in the first decade of the industry's history; from 4,000 to over 500,000 in the second

25 The idea of the integrated circuit was probably first propounded in 1952 by G. W. A. Dummer of the British Royal Radar Establishment (Braun and Macdonald 1978, p. 108).
decade; and from 500,000 to 100 million in the third decade. The ten-million-fold increase in the number of transistors per IC has been accompanied by only modest increases in the cost of processing of a wafer, and almost no change in the average costs of processing the individual IC. This factor alone has been responsible for the enormous cost reduction in electronic circuitry since the birth of the IC. Electronic systems comparable in complexity to vacuum-tube or transistor systems costing millions of dollars can be constructed for a few hundred dollars, a magnitude of cost reduction that it is virtually unprecedented in the history of manufacturing. The cheapness of electronic functions has reduced the costs of electronic systems relative to mechanical ones and lowered the relative price of electronic goods in general — developments that have had a major effect on the industrial structure of the electronics and IC industries.

Langlois and Steinmueller (1999) have pointed to the critical role of end-use demand in shaping industrial structure and competitive advantage in the world-wide semiconductor industry throughout its history. In the early years, demand in the United States came first from military sources and then importantly from the computer industry. Government procurement demand proved so valuable to the development of the industry not only because of its extent but also because of the military’s relative price-insensitivity and its insistence on reliability (Dosi 1984). Commercial demand eventually grew more rapidly than military, however, and, by the mid 1970s, government consumption had declined to less than ten per cent of the market (Kraus 1971, p. 91).

26 Canonical sources here are Abernathy and Utterback (1978) and Utterback (1979).
The American government also pushed the transistor and the IC through support of R&D and related projects. But scholarship on the subject is essentially unanimous that this activity was far less important for, and less salutary to, the industry than was the government’s procurement role. All the major breakthroughs in transistors were developed privately with the military market (among others) in mind. And, although the government tended to favor R&D contracts with established suppliers, notably old-line systems houses like RCA, it *bought* far more from newer specialized semiconductor producers (Tilton 1971, p. 91). The pragmatic policy of awarding work to those firms that could meet supply requirements was particularly important for encouraging new entry.

A significant feature of the transition to the IC was the virtual disappearance of those vertically integrated American electronics companies that had led in the production of vacuum-tubes. Although these firms had been able to stay in the race during the era of discrete transistors, their market shares began to plummet in the era of the integrated circuit. Why? Wilson, Ashton, and Egan (1980) point out that the new leaders were either specialized startups or multidivisional firms (like TI, Fairchild, and Motorola) in which the semiconductor division dominated overall corporate strategy and in which semiconductor operations absorbed a significant portion of the attention of central management. By contrast, the semiconductor divisions of the integrated system firms were a small part of corporate sales and of corporate strategy, thereby attracting a smaller portion of managerial attention and receiving less autonomy.
The digital computer.

The history of the digital computer has much in common with that of semiconductor technology, even if there are a number of important differences. Like the transistor, the digital computer was developed with a specific bottleneck in mind. But, unlike the transistor, the digital computer was developed not privately but at universities, with explicit government subsidy from the start.

During World War II, the U. S. Army contracted with J. Presper Eckert and John W. Mauchly of the Moore School at the University of Pennsylvania for a device “designed expressly for the solution of ballistics problems and for the printing of range tables”\(^{27}\) (Stern 1981, p. 15). By November 1945, they had produced the Electronic Numerical Integrator and Computer (ENIAC), the first fully operational all-electronic digital computer — a behemoth occupying 1,800 square feet, boasting 18,000 tubes, and consuming 174 kilowatts of electricity. Universities continued to play an important role throughout the early life of the technology, helping to create the wholly new discipline of computer science. Indeed, Rosenberg and Nelson go so far as to call the computer “the most remarkable contribution of American universities to the last half of the twentieth century” (Rosenberg and Nelson 1994, p. 331).

Government, especially military, support for the computer remained significant throughout the 1950s, and government funding helped spur important technical developments like ferrite-core memory, which emerged from the military-funded Whirlwind project at MIT (Redmond and Smith 1980; Pugh 1984). But, as Bresnahan
and Malerba (1999, pp. 89-90) argue, government research support had little to do with the success of the commercial computer industry. Moreover, much of government policy, notably in the areas of R&D funding and antitrust, was actually aimed at forestalling the emergence of IBM as a dominant “national champion” in computers. As in semiconductors, however, the military’s pragmatic approach to procurement favored those firms who could deliver the goods, and in computers that meant IBM (Bresnahan and Malerba 1999, p. 90; Usselman 1993).

By the mid 1960s, however, IBM found itself riding herd on a multiplicity of physically incompatible systems — the various 700-series computers and the 1400 series, among others — each aimed at a different use. Relatedly, and more significantly, software was becoming a serious bottleneck. By one estimate, the contribution of software to the value of a computer system had grown from eight per cent in the early days to something like 40 per cent by the 1960s (Ferguson and Morris 1993, p. 7). And writing software for so many incompatible systems greatly compounded the problem. In what Fortune magazine called “the most crucial and portentous — as well as perhaps the riskiest — business judgment of recent times,” IBM decided to “bet the company” on a new line of computers called the 360 series. The name meant to refer all the points of the compass, for the strategy behind the 360 was to replace the diverse and incompatible systems with a single modular family of computers (Flamm 1988, pp. 96-99). Instead of having one computer aimed at scientific applications, a second aimed at accounting applications, etc., the company would have one machine for all uses. This was not to be

27 In the event, the end of the war reduced the urgency of this goal, and the first major task given the ENIAC was actually to perform calculations for the development of the hydrogen bomb (Stern 1981, p. 62).
a homogeneous or undifferentiated product; but it was to provide a framework in which product differentiation could take place while retaining compatibility.

As Timothy Bresnahan suggests, the 360 was the first major computer platform, by which he means “a shared, stable set of hardware, software, and networking technologies on which users build and run computer applications” (Bresnahan 1999, p. 159). To put it another way, the 360 was a modular system, albeit one that remained mostly closed and proprietary despite the efforts of the “plug compatible” industry to pick away at its parts. Carliss Baldwin (2006) argues that IBM never understood the tremendous option value implied in the 360 architecture. Had the company opened the architecture up to the market while retaining control of key bottlenecks in the system, they could have created considerably more value.

As the market for computers picked up speed, the symbiosis between computers and semiconductors became stronger: competition among computer makers drove the demand for ICs, which lowered IC prices by moving suppliers faster down their learning curves, which in turn fed back on the price of computers, etc. The result was a self-reinforcing process of growth for both industries. Indeed, the falling prices of semiconductor logic fueled a second computer revolution, that of the minicomputer. Minicomputers were smaller than mainframes and geared toward specialized scientific and engineering uses. Digital Equipment Corporation (DEC), founded in 1957, was the pioneer in the field. Among the other firms to enter the minicomputer market were Scientific Data Systems, Data General (founded in 1968 by defectors from DEC), Prime Computer, Hewlett-Packard, Wang, and Tandem (Flamm 1988, p. 131).
Japanese challenge and American resurgence.

Japan had responded to American competitive advantage with high tariffs, and in addition imposed quotas and registration requirements\(^{28}\) (Tyson and Yoffie 1993, p. 37). The Japanese government essentially forbade foreign direct investment, which forced American firms to tap the Japanese market only through licensing and technology sales to Japanese firms rather than through direct investment.

In Japan, the principal producers of transistors in the 1950s and 1960s were diversified systems houses, including firms that had previously produced vacuum tubes, rather than companies that were principally specialized into semiconductors. Moreover, the main end-use for transistors in Japan was consumer products rather than the military. But the Japanese vacuum-tube firms were much smaller than their American or European counterparts at the beginning of the transistor era. As Tilton (1971, p. 154) notes, the small size and rapid growth of the Japanese firms “also helped create a receptive attitude toward change on the part of the [Japanese] receiving tube producers by reducing the risks associated with new products and new technologies and by increasing costs, in terms of declining market shares, to firms content simply to maintain the status quo.” In many ways, then, Japanese systems firms faced many of the same constraints, and adopted many of the same approaches, as the aggressive American merchant firms rather than those of the American systems houses. The Japanese also sought licenses primarily from the American merchants rather than the American systems firms.

\(^{28}\) This is in contrast to European policy, which featured high tariffs but no prohibition on foreign direct investment. As a result, much of the European demand for semiconductors was satisfied by European subsidiaries of American companies. Japanese companies have typically supplied some 90 per cent of the Japanese semiconductor market, whereas American firms — through imports or foreign direct
Despite their early success in transistors, Japanese firms found themselves in a weak position by the 1970s. These firms were slow to make the transition to batch-produced silicon devices in the early 1960s, and, when they turned later in the decade to the production of bipolar ICs, they could not compete with the likes of Texas Instruments and National Semiconductor; some Japanese firms accused the Americans of “dumping” (Okimoto et al. 1984, pp. 14-15). After 1967, indeed, the purchase of American ICs created a Japanese trade deficit in semiconductors (Malerba 1985, p. 136).

How did Japanese industry move from this weak position in the 1970s to its dominant position by the mid-1980s? Until recently, the tacit assumption of most commentators had been that Japanese success was the result of some combination of (1) Japanese industrial structure, understood as superior to American industrial structure in a very general or even absolute sense, and (2) Japanese industrial policy, understood as a highly intentional — and even prescient — system of government-industry planning and control. Langlois and Steinmueller (1999) suggests a somewhat different picture. Although both industrial structure and government policy played important roles in the rise of the Japanese semiconductor industry, the benefits of that industrial structure were far less timeless than commentators supposed, and the effects of government policy were far less intentional, and perhaps somewhat less significant, than the dominant accounts suggested.

---

investment — have supplied between 50 and 70 per cent of the European market (Tyson and Yoffie 1993, p. 34).
As in the earlier rise of the American semiconductor industry, the pattern of end-use demand was crucial in shaping the bundle of capabilities that Japanese industry possessed, as well as in narrowing and limiting the choices the Japanese firms had open to them. In this case, that end-use demand came largely from consumer electronics and, to a somewhat lesser extent, from telecommunications, especially purchases by NTT, Japan’s national telephone monopoly. (See Table 2.) Consumer demand helped place the Japanese on a product trajectory — namely MOS and especially CMOS ICs — that turned out eventually to have much wider applicability. Moreover, Japanese firms adopted a strategy of specialization in high-volume production of one particular kind of chip. The DRAM, or dynamic random-access memory chip, is a technology that benefited from increasing returns to scale not only because of the volume effects of mass

<table>
<thead>
<tr>
<th>End-use</th>
<th>United States</th>
<th>Japan</th>
<th>Western Europe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer</td>
<td>40 45</td>
<td>22 36</td>
<td>25 20</td>
</tr>
<tr>
<td>Telecommunications</td>
<td>21 10</td>
<td>10 13</td>
<td>20 29</td>
</tr>
<tr>
<td>Industrial</td>
<td>11 10</td>
<td>17 6</td>
<td>25 19</td>
</tr>
<tr>
<td>Military and Aerospace</td>
<td>17 18</td>
<td>0 0</td>
<td>5 7</td>
</tr>
<tr>
<td>Consumer</td>
<td>11 16</td>
<td>51 45</td>
<td>25 25</td>
</tr>
</tbody>
</table>

Table 2
Demand for integrated circuits by end-use market, United States, Japan, and Western Europe, 1982 and 1985, in percent.

Note: Includes captive consumption.
production but also because it is arguably a general-purpose technology of considerable importance — a device that can store digital information for a wide variety of purposes.\textsuperscript{30}

Established American firms, accustomed to providing customized devices, were slow to recognize the cost-reduction advantages of a standardized memory chip (Wilson, Ashton, and Egan 1980, p. 87; Dorfman 1987, p. 193). Two new firms — National and Intel — quickly gained advantage over their established competitors in the merchant market by moving more quickly into the production of high-volume standardized devices. Both firms were spin-offs from Fairchild — two of the first of what came to be called the “Fairchildren” (Lindgren 1969). In pushing standardized DRAM chips, however, these firms precipitated a “memory race” in which Japanese firms were eventually to prove dominant. American firms led in the early — 1K and 4K — DRAM markets. But an industry recession delayed the American “ramp-up” to the 16K DRAM, which appeared in 1976. Aided by unforeseen production problems among the three leaders, Japanese firms were able to gain a significant share of the 16K market. By mid-1979, 16 companies were producing DRAMs, and Japanese producers accounted for 42 per cent of the market (Wilson, Ashton and Egan 1980, pp. 93-94). (See Table 2). The opportunity opened for Japanese producers in the 16K DRAM market had proven sufficient for them to advance to a position of leadership in the 64K DRAM. Japanese dominance accelerated in the 256K (1982) and one-megabit (1985) generations. Intense price competition, combined with the general recession in the U.S. industry in 1985, caused all

\textsuperscript{29} MOS stands for metal-oxide semiconductor, and CMOS for “complementary” MOS.

\textsuperscript{30} DRAMs are “dynamic” in the sense that the electric charges containing the remembered information decay over time and need periodically to be “refreshed.” This stands in contrast to the static RAM (or SRAM), which does not require refreshing, but which therefore has disadvantages in size, cost, and power consumption because it requires more transistors per memory cell.
<table>
<thead>
<tr>
<th>Device</th>
<th>Maximum Market Share (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>United States</td>
</tr>
<tr>
<td>1K</td>
<td>95</td>
</tr>
<tr>
<td>4K</td>
<td>83</td>
</tr>
<tr>
<td>16K</td>
<td>59</td>
</tr>
<tr>
<td>64K</td>
<td>29</td>
</tr>
<tr>
<td>256K</td>
<td>8</td>
</tr>
<tr>
<td>1M</td>
<td>4</td>
</tr>
<tr>
<td>4M</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 3
Maximum market share in DRAMs by American and Japanese companies, by device.

Source: Dataquest, cited in Methé (1991, p. 69)

but two American merchant IC companies to withdraw from DRAM production\(^{31}\) (Howell et al. 1992, p. 29). In 1990, American market share had fallen to only two per cent of the new generation 4-megabit DRAMs\(^{32}\) (See Table 3.)

Why did the Japanese succeed? In broad terms, circumstances had staked out for the Japanese industry a strategic path that fit well the existing competences of the firms — namely those in mass production and quality control — and supported the thrust of

---

\(^{31}\) The exceptions were Texas Instruments, which produced in Japan, and Micron Technology, which produced in Idaho.
their final products, which, despite government efforts of to create a computer industry (Fransman 1990), were still in consumer electronics and telecommunications.

Rather than feeling that they were on the verge of overtaking American companies, the Japanese saw their computer industry as relatively weak against IBM, and perceived that a key feature of IBM’s advantage was technology, specifically its position in ICs. From the viewpoint of Japanese firms, the American IC industry was enormously innovative but did not share much of the manufacturing culture that had developed in the larger Japanese electronics companies, where quality, systematic capacity expansion, and long term market position were regarded as key variables to control. The fact that Japanese IC producers were large companies in comparison with their American counterparts gave them one particular advantage: they were able to mobilize internal capital resources to make investments in the IC industry in a way that U.S. companies could not.

James March (1991) has pointed out that there is a necessary tradeoff between exploration and exploitation — tradeoff between searching for new ideas and running with the old ones. As the technology leaders, the American firms found themselves with a full plate of alternatives to pursue, in both product and process technology. Sitting somewhat behind the frontier, Japanese firms could pick one item off the plate and run with it. Their morsel was the mass production of DRAMs.

American industry and politics certainly did not let these events go unnoticed, and alarms went up as early as the 64K generation. More worrisome than the loss of the

---

32 These figures do not take into account the sizable captive production at IBM and AT&T.
memory market was the possibility that Japanese dominance in DRAMs would be translated into equal success in other kinds of chips. Although memories constituted at most 30 per cent of the IC market, many believed them to be “technology drivers” essential for continued progress in increasing the number of transistors on an IC. If American firms couldn’t use DRAM production to develop and gain experience in the next generation of technology, then Japanese producers would soon be able to climb up the design-complexity ladder and challenge U.S. positions in logic markets (Ferguson 1985; Forester 1993).

In 1986, Japan’s overall market share in semiconductors slipped ahead of that of the American merchants. Thus, in 1988 the U.S. industry appeared to stand on the brink

---

**Figure 3**

Worldwide semiconductor market shares (in percent), 1982-1998

*Source: Semiconductor Industry Association.*
of oblivion, with no haven in product or process that could be counted to insure its survival into the 1990s. But the predicted extinction never occurred. (See Figure 3.) Instead, American firms surged back during the 1990s, and it now seems the Japanese who are embattled.

Langlois and Steinmueller (1999) argue that this resurgence is not the result of imitating Japanese market structure and policy but rather of taking good advantage of the distinctly American market structure and capabilities developed in the heyday of U. S. dominance. Just as the innovation of, and the growing market for, the standardized DRAM had favored the Japanese, another semiconductor innovation, and the burgeoning market it created, came to favor the Americans. That innovation was the microprocessor, an integrated circuit designed not to store information (like the DRAM) but rather to provide on a single chip the information-processing capability of a digital computer.

In 1969, a Japanese manufacturer asked Intel to design the logic chips for a new electronic calculator. Marcian E. (“Ted”) Hoff, Jr., the engineer in charge of the project, thought the Japanese design too complicated to produce. The then-current approach to the design of calculators involved the use of many specialized hard-wired circuits to perform the various calculator functions. Influenced by the von Neumann architecture of minicomputers, Hoff reasoned that he could simplify the design enormously by creating a single programmable IC rather than the set of dedicated logic chips the Japanese had sought (Noyce and Hoff 1981). By using relatively simple general-purpose logic circuitry that relied on programming information stored elsewhere, Hoff effectively substituted cheap memory (then Intel’s major product) for relatively expensive special-purpose logic circuitry (Gilder 1989, p. 103). The result was the Intel 4004, the first microprocessor. A
sixth of an inch long and an eighth of an inch wide, the 4004 was roughly equivalent in computational power to early vacuum-tube computers that filled an entire room. It also matched the power of a 1960s IBM computer whose central processing unit was about the size of a desk (Bylinsky 1980, p. 7).

Intel gained an early lead in microprocessors that it never relinquished. Early on, Intel did not push patent protection, and, in Hoff’s view, “did not take the attitude that the microprocessor was something that you could file a patent claim on that covers everything” (quoted in Malone 1985, p. 144). Because the microprocessor is a general-purpose computer, there are many different ways to implement the microprocessor idea without infringing on a particular implementation. And the appropriation of rents in microprocessors has always depended on first-mover advantage rather than on patent protection for particular features of the system design or on the ability to produce a microprocessor that could not be emulated technically.

The microprocessor found uses in a wide variety of applications involving computation and computer control. But it did not make inroads into the established mainframe or minicomputer industries, largely because it did not initially offer the level of computing power these larger machines could generate using multiple logic chips. Instead, the microprocessor opened up the possibility of a wholly new kind of computer — the microcomputer.

*The personal computer.*

The first personal computer (or microcomputer) is generally acknowledged to have been something called the MITS/Altair, which graced the cover of Popular Electronics
magazine in January, 1975.\footnote{For a much longer and better-documented history of the microcomputer, see Langlois (1992), on which this section draws.} Essentially a microprocessor in a box, the machine’s only input/output devices were lights and toggle switches on the front panel, and it came with a mere 256 bytes of memory. But the Altair was, at least potentially, a genuine computer. Its potential came largely from a crucial design decision: the machine incorporated a number of open “slots” that allowed for additional memory and other devices to be added later. These slots were hooked into the microprocessor by a network of wires called a “bus.” This extremely modular approach emerged partly in emulation of the design of minicomputers and partly because hobbyists and the small firm supplying them would have been incapable of producing a desirable (that is, more-capable) non-modular machine within any reasonable time. In effect, the hobbyist community captured the machine, and made it a truly open modular system. The first clone of the Altair — the IMSAI 8080 — appeared within a matter of months, and soon the Altair's architecture became an industry standard, eventually known as the S-100 bus because of its 100-line structure.

The S-100 standard dominated the hobbyist world. But the machine that took the microcomputer into the business world adopted a distinctive architecture, built around a Motorola rather than an Intel microprocessor. Stephen Wozniak and Steven Jobs had started Apple Computer in 1976, quite literally in the garage of Jobs’s parents’s house. The hobbyist Wozniak, also influenced by the architecture of minicomputers, insisted that the Apple be an expandable system — with slots — and that technical details be freely available to users and third-party suppliers. With the development of word
processors like WordStar, data-base managers like dBase II, and spreadsheets like VisiCalc, the machine became a tool of writers, professionals, and small businesses. Apple took in three quarters of a million dollars by the end of fiscal 1977; $8 million in 1978; $48 million in 1979; $117 million in 1980 (when the firm went public); $335 million in 1981; $583 million in 1982; and $983 million in 1983.\footnote{Data from Apple Computer, cited in “John Sculley at Apple Computer (B),” Harvard Business School Case no. 9-486-002, revised May 1987, p. 26.}

Existing computer companies were slow to develop competing microcomputers, largely because they saw the machines as a small fringe market. But as business uses increased and microcomputer sales rose, some computer makers saw the opportunity to get a foothold in a market that was complementary to, albeit much smaller than, their existing product lines.\footnote{Few people inside or outside IBM foresaw the sweeping changes the PC would make in computer markets. In April 1981, four months before the official announcement of the IBM PC, IBM gave presentations estimating it would sell 241,683 PCs over five years. In fact, IBM shipped 250,000 PCs in one month alone (Zimmerman and Dicarlo 1999).} By far the most significant entry was that of IBM. On August 12, 1981, IBM introduced the computer that would become the paradigm for most of the 1980s.

In a radical departure, IBM decided to produce the machine outside the control of company procurement policies and practices. Philip Donald Estridge, a director of the project, later put it this way. “We were allowed to develop like a startup company. IBM acted as a venture capitalist. It gave us management guidance, money, and allowed us to operate on our own” \textit{(Business Week, October 3, 1983, p. 86)}. Estridge knew that, to meet the deadline he had been given, IBM would have to make heavy use of outside vendors for parts and software. The owner of an Apple II, Estridge was also impressed by the
importance of expandability and an open architecture. He insisted that his designers use a modular bus system that would allow expandability, and he resisted all suggestions that the IBM team design any of its own add-ons. Because the machine used the Intel 8088 instead of the 8080, IBM needed a new operating system. A tiny Seattle company called Microsoft agreed to produce such an operating system, which they bought from another small Seattle company and rechristened MS-DOS, for Microsoft Disk Operating System.

The IBM PC was an instant success, exceeding sales forecasts by some 500 per cent. By 1983, the PC had captured 26 per cent of the market, and an estimated 750,000 machines were installed by the end of that year. The IBM standard largely drove out competing alternatives during the decade of the 1980s. This happened in part because of the strength of the IBM name in generating network effects, principally because it created the expectation among users that the key vendor would continue to provide services long into the future and that a wide array of complementary devices and software would rapidly become available. But in large measure the “tipping” of the market to the IBM PC standard was a result of the openness of the IBM system, which could be easily copied by others, and the eagerness of Microsoft to license MS-DOS to all comers.

As it had with the 360/370 series, IBM had created a dominant computer platform. But, in the case of the PC, the dominance of the platform would not translate into a dominant market share for IBM. Because of the strategy of outsourcing and the standards it necessitated, others could easily imitate the IBM hardware, in the sense that any would-be maker of computers could obtain industry-standard modular components and compete with IBM. A legion of clones that offered IBM compatibility at, usually, a price lower than what IBM charged. By 1986, more than half of the IBM-compatible
computers sold did not have IBM logos on them. By 1988, IBM's worldwide market share of IBM-compatible computers was only 24.5 per cent. IBM's choice of an open modular system was a two-edged sword that gave the company a majority stake in a standard that had grown well beyond its control. For reasons that are debated in the literature, but that likely have to do both with strategic mistakes by IBM and with the inherently strong positions of key suppliers in controlling their proprietary “bottleneck” technologies — the microprocessor and the operating system — Intel and Microsoft gained control of the standard that IBM had originally sponsored (Ferguson and Morris 1993). The PC architecture is now often referred to as the “Wintel” (Windows/Intel) platform.

Langlois (1992) has argued that the rapid quality-adjusted price decline in microcomputers resulted not only from the declining price of computing power attendant on successive generations of Intel processors but also from the vibrant competition and innovation at the level of hardware components and applications software that resulted from the open modular design of the PC. A decentralized and fragmented system can have advantages in innovation to the extent that it involves the trying out of many alternate approaches simultaneously, leading to rapid trial-and-error learning. This kind of innovation is especially important when technology is changing rapidly and there is a high degree of both technological and market uncertainty (Nelson and Winter 1977). Moreover, the microcomputer benefited from technological convergence, in that it turned out to be a technology capable of taking over tasks that had previously required numerous distinct — and more expensive — pieces of physical and human capital. By the early 80s, a microcomputer costing $3,500 could do the work of a $10,000 stand-alone word-
processor, while at the same time keeping track of the books like a $100,000 minicomputer and amusing the kids with space aliens like a 25-cents-a-game arcade machine.

The personal computer grew rapidly in a niche that existing mainframes and minicomputers had never filled. Quickly, however, the microcomputer’s niche began to expand to encroach on the territory of its larger rivals, driven by the rapidly increasing densities and decreasing prices of memory chips and microprocessors. In the early 1980s, a class of desktop machines called workstations arose to challenge the dominance of the minicomputer in scientific and technical applications. As in the case of personal computers, the workstation market was driven by open technical standards and competition within the framework of what was largely a modular system (Garud and Kumaraswamy 1993; Baldwin and Clark 1997). Initially, these workstation used microprocessors and operating systems different from those of personal computers. By the early 1990s, however, the same process of increasing power and decreasing cost began pushing the Windows-Intel platform into what is today a dominance of the workstation space. At the same time, workstations hooked together (or hooked to personal computers) began to take over many of the functions of larger minicomputers and mainframes. By the 1990s, networks of fast, cheap smaller machines were

---

36 So-called traditional workstations are built around Reduced-Instruction-Set-Computing (RISC) microprocessors and run variants of the UNIX operating system. Intel-platform workstations use high-end versions of the same microprocessors used in personal computers and typically run Microsoft’s Windows NT or Windows 2000, which are compatible with Microsoft’s operating systems for personal computers.
widespread, a development accelerated by the spectacular growth of the Internet.\footnote{In some respects, the demand for large websites created by the Internet has spurred demand for large central servers. Increasingly, however, even these servers are frequently networks of high-powered personal computers rather than traditional mainframes or minicomputers.} This growth had a significant negative effect on the makers of larger computers, notably the Boston-area minicomputer makers. Many went bankrupt; and, in a telling development, the flagship maker of microcomputers — DEC — was acquired by Compaq, a maker of microcomputers. Bresnahan and Greenstein (1996, 1997) refer to this encroachment of smaller computers as the “competitive crash” of large-scale computing.

Former Intel CEO Andy Grove (1996) has famously described the evolution of the computer industry as a transition from a structure of vertical “silos” in the days of IBM and DEC to a horizontal structure today. Once large multidivisional firms undertook virtually all stages of production internally, and captured rents at the level of the system. Nowadays computers – and electronics more generally – are the product of multiple independent suppliers competing at every stage of production. Such competition drives down costs and spurs modular innovation (Langlois and Robertson 1992). But, because of the PC’s relatively open modular structure, the assembly of computer systems themselves is not an obvious source of economic rent, and few assemblers prospered in a consistent way. Compaq, which had gobbled up the remains of DEC, was itself eaten by Hewlett-Packard. Gateway, an early mail-order success, has flirted with bankruptcy. And, in late 2004, the originator of the IBM PC sold the entirety of its PC operations to Lenovo of China (Williams and Kallender 2004). The one spectacular success has been Dell Computer, which has been able to use the modular structure of the personal computer to its advantage by making good use of the “external” capabilities of a
worldwide network of suppliers (Curry and Kenney 1999; Kraemer and Dedrick 2002; Field 2004). More than any other major producer, Dell has abandoned the traditional model of the integrated electronics firm in favor of what Baldwin and Clark (2006) call a small-footprint strategy. Dell’s source of rents lies in not in any physical assets it owns but rather in the way it organizes the PC value chain, including through its own innovative logistics system.

Among suppliers, a principal beneficiary of the rise of the personal computer was the American semiconductor industry. The abandonment of the DRAM market by most American firms — including Intel — was a dark cloud with a bright silver lining. When Intel led the world industry in almost all categories, it and many of its American counterparts faced a full plate of product alternatives. With the elimination of mass memory as a viable market, these firms were impelled to specialize and narrow their focus to a smaller subset of choices. The areas in which American firms concentrated can generally be described as higher-margin, design-intensive chips. For such chips, production costs would not be the sole margin of competition; innovation and responsiveness would count for more. And innovation and responsiveness were arguably the strong suit of the “fragmented” American industry. As in the case of the personal computer industry, the decentralized structure of the American semiconductor industry permitted the trying out of a wider diversity of approaches, leading to rapid trial-and-error learning (Nelson and Winter 1977). And the independence of many firms from larger organizations permits speedier realignment and recombination with suppliers and customers. Building on existing competences in design (especially of logic and specialty circuits) and close ties with the burgeoning American personal computer industry,
American firms were able to prosper despite the Japanese edge in manufacturing technology (Ferguson and Morris 1993).

The most important area of America specialization has been microprocessors and related devices.\textsuperscript{38} Between 1988 and 1994, a period in which merchant IC revenues grew by 121 per cent, revenues from the microprocessor segment grew much faster than did memory revenues (ICE 1998). This evolution of the product mix in the industry has strongly favored American producers. In the microprocessor segment of the chip market, American companies accounted for 72 per cent of world production in 1996, compared with a 21 per cent share for Japanese companies. (See Figure 3.)

\textsuperscript{38} This segment includes not only microprocessors but also microcontrollers (less sophisticated microprocessors that are used in embedded applications) and related “support” chips, such as memory controllers, that are necessary to assembling a microprocessor system.
The importance of the microprocessor segment has meant that a single company, Intel, is responsible for much of the gain of American merchant IC producers. In 1996, Intel accounted for 43 per cent of world output in the microprocessor segment. (See Figure 4.) Intel’s strategy for recovery, begun in the 1980s, has proven remarkably successful (Afuah 1998). In the late 1980s, the firm consolidated its intellectual-property position in microprocessors by terminating cross-licensing agreements with other companies and, more importantly, began extending its first-mover advantage over rivals by accelerating the rate of new product introduction. These developments pushed Intel into the position of the largest IC producer in the world, with 1998 revenues of $22.7 billion — more than the next three largest firms combined. (See Table 5.) Although Intel dominates the microprocessor market, it is not entirely without competitors; and it is
significant that its principal competitors in microprocessors are also American companies, notably AMD and Motorola.

Another aspect of specialization that benefited the American industry was the increasing “decoupling” of design from production, a result in this case of growth in the extent of the market, which brought with it the development of computerized design tools (Hobday 1991) and the standardization of manufacturing technology (Macher, Mowery, and Hodges 1998). On the one hand, this allowed American firms to specialize in design-intensive chips, taking advantage of a comparative advantage that arguably arises out of the decentralized and “fragmented” structure of American industry. On the other hand, it also allowed many American firms to take advantage of growing production capabilities overseas. This “modularization” of the semiconductor industry is spurring the kind of decentralized innovation from which the personal computer industry has benefited.

Another area in which American suppliers have prospered is, of course, software, with Microsoft sitting prominently at at least one critical bottleneck. Effectively, the personal computer relies for its modular structure on three major technological standards. One involves the microprocessor, where, as we saw, Intel (and now AMD) control the standard. Another involves the architecture of the “bus” along which the various pieces of the computer communicate with one another; this is in the public domain, shepherded by a committee of an industry trade group. And the third standard is the operating system, which regulates how the hardware communicates with the higher-level (or
“applications”) software that yields the services consumers ultimately demand. Here Microsoft came to earn significant rents as first MS-DOS and then the Windows operating system, over whose underlying code Microsoft owns copyright, became the dominant technical standard. American firms also remain strong in other key segments of the value chain, including disk drives (Kraemer and Dedrick 2002). And, although PC production takes place in a global network of suppliers (Angel and Engstrom 1995), American firms have in general retained those aspects of the production process requiring high skill levels and paying high wages (Dedrick and Kraemer 2007, p. 22).

\textit{Convergence and conclusion.}

By the mid 1980s, by most accounts, America had “lost” consumer electronics and was in imminent danger of losing semiconductors and computers.\textsuperscript{40} Contemporary analysis ran within Chandlerian channels. Innovation and production both necessarily emanate from large multidivisional firms. The leading American versions of these had either failed (RCA) or were embattled (IBM), while Japan’s large multidivisional firms were on the ascendant. Ultimate reasons were sought in the realms of culture and government policy. If organization had a lesson, it was that American firms should become more like Japanese firms.

As it has a habit of doing, however, history failed to conform to predicted trends. The most striking development at the end of the electronic century was the convergence of consumer electronics into digital technology. As we saw, digital technology

\textsuperscript{39} Perhaps surprisingly, the mid-1980s — that dark period for American fortunes — was actually the most fertile period in history for the startup of new semiconductor firms, by a large margin. Most of these new firms were involved in design-intensive custom devices and ASICs (Angel 1994, p. 38).
proceeded in important respects within an open modular framework in which specialized American players were often better positioned than large multidivisional firms to capture pieces of value. As digital technology developed, its various strands began to converge and to encompass what had been consumer electronics. Coupled to the Internet and cellular communications networks, the microprocessor has become the focal point for the generation and transmission not only of “data” (as we used to understand that term) but also of voice, music, text, and video. Sometimes the microprocessor lies within a conventional personal computer; but increasingly it is the core of a device that combines the functions of a handheld computer, a phone, a camera, a music player, an arcade game, and even a video player.

The network that produces such devices, their components, and the attendant software (now increasingly broadly understood) is genuinely international, with significant players in Japan, Korea, Europe, and elsewhere, notably China. But whereas Japan was arguably the hotbed of innovation in consumer electronics in the 1970s and 80s – video-tape recorders and cameras, optical disks, and miniaturized devices like the Sony Walkman – by century’s end the United States had regained a significant measure of its stature in that field. At the turn of the millennium, Apple’s iPod was arguable the signature device in converged digital consumer electronics.

What does all this imply? Paths of learning are not thoroughfares excavated by large multidivisional firms with entourage in train. They have always been, and are perhaps increasingly, trails beaten out by a variety of specialists working in cooperation

40 And maybe even software (Cusumano 1991).
and competition. The Chandlerian model works well for producing systemic innovations in their early stages (television was a prime example) and occasionally for generating fundamental new ideas (like the transistor). But many if not most important developments – from the vacuum tube to the planar process, from the radio to the personal computer – were the product of specialists within the network. (The digital computer was the product of a special kind of specialist, the university.) Moreover, by taking advantage of a range of capabilities far wider than the boundaries of what even the largest firm can encompass, a network of specialist suppliers and competitors is better able to exploit the value of a complex and potentially modular product architecture.
References.


