Innovation, competition, and industry structure

James M. Utterback and Fernando F. Suárez
Sloan School of Management, MIT, Cambridge, MA, USA

Final version received June 1991

Why some firms die while others survive? Survival has long been recognized as a basic goal for a manufacturing firm. At least in the long term, survival should be related to various measures of performance, such as market share and profitability. Advocates of population ecology have argued that life chances of organizations are affected by population density at the time of founding. According to this argument, organizations founded during periods of intense competition will have persistently higher age-specific rates of mortality than those founded during periods with lower numbers of competitors. At least for the case of manufacturing firms, there may be more profound causes than competitive turmoil that explain a firm's survival chances. These have to do with the evolution of technology in an industry. Population density may only be a reflection of underlying driving forces based on technological change that determine the form and level of competition, the attractiveness of entry, and ultimately the structure of an industry.

How is changing product and process technology related to competition and industry structure? Although in recent years there has been increased interest in the subject (see, for instance, the work of the MIT Commission on Productivity [23]), few scholars have examined the links between production processes and process improvements and productivity advances and competitiveness. Fewer still have looked at product technology or the relationships between product technology and their enabling processes and corporate strategy and competitiveness. Our contention is that these questions are central to understanding questions of larger scope such as the long term success or failure of firms and even industries. Just as the facets and geometry of a crystal reflect the micro-structure, chemistry and physics of its constituent atoms and molecules, we believe that the competitive structure and dynamics of an industry reflect underlying product and process technologies and innovations. Unlike a crystal, the shapes taken by technological change or by industry structure are not necessarily predetermined. Rather, we contend only that choices made at one level are necessarily reflected at the other. For instance, greater degrees of competition will result in more rapid rates of technological change, while rapidly advancing technology and potentials for broadening application will attract entrants.

In earlier work, Utterback and Abernathy [62] introduced the concept of a dominant product design and suggested that the occurrence of a dominant design may alter the character of innovation and competition in a firm and an industry. A dominant design usually takes the form of a new product (or set of features) synthesized from individual technological innovations introduced independently in prior product variants. A dominant design has the effect of enforcing or encouraging standardization so that production or other complementary economies can be sought. Then effective competition begins to take place on the basis of cost and scale as well as of product performance. Prior to the appearance of a dominant design, we expect to see a wave of entering firms with many varied, experimental versions of the product. Following the dominant design, we expect to see a wave of exits and consolidation of the industry.

Correspondence to: J.M. Utterback, Room E52-541, Sloan School of Management, MIT, Cambridge, MA 02139-4307, USA.
Similar dominant design milestones can be identified in many product lines. These often have the result of drastically reducing the number of performance requirements to be met by a product by making many features implicit in the dominant design and its increasing acceptance. Examples include the common typewriter keyboard, the RCA television standard and the three color black mask picture tube also devised by RCA, the simple four-function calculator and the integrated circuit computer memory chip in its successively larger manifestations. That dominant designs are not necessarily predetermined is easily illustrated by considering that the standard keyboard was designed to minimize the interference of mechanical typewriter keys, or by considering Sony's challenge to the RCA devised tube. A dominant design is generally the product of the experiments, technical possibilities, choices and proprietary positions of its day. Equally, the persistence of the older designs mentioned illustrates the momentum of both established practice and complementary assets such as typing skills and training. Once such a design is accepted it can have a profound impact on both the direction of further technical advance, on the rate of that advance, and on industry structure and competition.

In this article we intend to explore the questions above by looking at a series of examples arrayed over the past century. These include typewriters, the automobile, television sets and picture tubes, transistors, integrated circuits, calculators and supercomputers. Data on industry participation and parallel data on technological change over the full course of an industry's development have been difficult to obtain. Thus, although our sample is not balanced or weighted in any scientific sense, it consists of the most complete sets of data that we have so far been able to discover or synthesize. More fragmentary data from other industries will be used to illustrate particular points. The data available do convincingly point to the fact that a dominant design and product standardization mark a watershed in industry structure and competition in each case examined. In all but the first case (typewriters) and the last (massively parallel computers) information on product technology and industry participation have been derived from independent sources. To date our investigations have been limited to assembled products and essentially to the United States.

**Hypotheses**

We suggest that creative synthesis of a new product innovation by one or a few firms results in a temporary monopoly situation, high unit profit margins and prices, and sales of the innovation in those few market niches where it possesses the greatest performance advantage over other competing alternatives. This is in line with Schumpeter's pathbreaking "creative destruction" model and subsequent studies on the economics of innovation [32,52]. As volume of production and demand grows, and as a wider variety of applications is opened for the innovation, many new firms will enter the market with diverse variations of the product. For example, early versions of the automobile included steam and electric vehicles as well as the now familiar internal-combustion engine. It is common to think of new, innovative entrants as being small firms. We do not agree entirely with the idea that only small firms can be innovative in a young industry. Instead of focusing on firm size, we contend that innovative firms often come from outside the industry in question (this argument is in line with the earlier work of Gilfillan [36] and Schön [51]).

The appearance of a dominant design shifts the competitive emphasis to favor those firms, large or small, which are able to achieve greater skills in process innovation and process integration, and with more highly developed internal technical and engineering skills. The emergence of a dominant design will mark the beginning of a shake-out period in an industry (as we will see in the data presented below). Firms that are notable to make the transition toward greater product standardization and process innovation will be unable to compete effectively and will eventually fail. Others may possess special resources and thus successfully merge with the ultimately dominant firms. Some weaker firms may merge and still fail. Overall, a firm's inability to change its organization structure and practices along with the evolution of technology in the industry will be its major source of failure.

Eventually, we believe that the market reaches a point of stability in which there are only a few
large firms having standardized or slightly differentiated products and relatively stable sales and market shares, until a major technological discontinuity occurs and starts a new cycle again. ¹ During the stability period, a few small firms may remain in the industry, serving specialized market segments, but, as opposed to the small firms entering special segments early in the industry, they have little growth potential. Thus, it is important to distinguish between merely small firms and small firms which are new entrants, and to keep in mind that the term new entrants includes existing firms (large or small) moving from their established market or technological base into a new product area.

Mueller and Tilton [47] were among the first to present this hypothesis in its entirety. They contend that a new industry is created by the occurrence of a major process or product innovation and develops technologically as less radical innovations are introduced. They further argue that the large corporation seldom provides its people with incentives to initiate a development of radical importance. Thus, these changes tend to be developed by new entrants without an established stake in a product market segment. In their words, neither large absolute size nor market power appears to be a necessary condition for successful development of most major innovations.

Mueller and Tilton contend that once a major innovation is established, there will be a rush of firms entering the newly formed industry, or adopting a new process innovation. They hold that during the early period of entry and experimentation immediately following a major innovation, the science and technology upon which it depends is often only crudely understood, and that this reduces the advantage of large firms over others. However, the authors suggest that as the number of firms entering the industry increases and more and more R&D is undertaken on the innovation, research becomes increasingly specialized and innovations tend to focus on improvements in small elements of the technology. This clearly works to the advantage of larger firms in the expanding industry and to the disadvantage of smaller entrants. Product differentiation will be increasingly centered around the technical strengths and R&D organization of the existing firms. Strong patent positions may have been established by earlier entering firms that are difficult for later entrants to completely circumvent.

Burton Klein [41] suggests a profound connection between industry structure and technological change in his seminal work on dynamic economics. Klein portrays each firm's investments and product introductions as experiments which provide corrective and stimulating feedback to that firm and to the industry about product and market requirements. Thus, the earliest period in the development of a product line or industry in which few firms participate would necessarily be a period of relatively slow technical progress and productivity advance. As larger numbers of firms enter the arena, thus broadening the range of experimentation and the definition of the product technology, Klein expects greater innovation with correspondingly greater technological progress and productivity advance. Finally, as a few firms come to dominate the industry due to superior product technology and productivity, both experimentation and progress will slow. Renewal or broadening of competition would seemingly be required for more rapid progress to recur. In reviewing earlier work, Klein finds no case in which a major advance, one which established a new and more rapid trajectory for technological progress, came from a major firm in the industry in question. From this evidence, he concludes that the process of moving from a dynamic organization to a static one, of going from a period of rapid organizational learning to a period of slow or no progress, appears to be highly irreversible.

Some economists interested in technological innovation have proposed models that parallel many of the features of our model. Gort and Klepper [34] present a five-stage product life cycle model which they then contrast with data on 46 industries. The implications of their model for industry structure over time anticipate to ours. Klepper and Graddy [43] present an analytical model that attempts to capture the main “regularities” discussed in Gort and Klepper [34]. Gort and Konakayama [35] estimate probabilities of entry and exit based on a logistic model. Al-

¹ The successful entry of large Japanese firms into mature industries in the last few decades has provided a counter-example to this hypothesis. We discuss the Japanese strategy later on in the paper.
though important contributions bridging economics and the management of technology, these studies omit important insights coming from a deeper understanding of an industry’s technological evolution. For instance, the data on innovations are not divided between product and process innovation, and thus do not allow the authors to test whether the locus of innovation changes over time. The identification of each stage in prior studies is basically done by looking at the net entry data itself. The work presented here attempts to isolate the identification of the different stages from the industry’s net entry data by examining the evolution of technology through independent sources. The emergence of a dominant design, for instance, is seen as an unequivocal sign that a new stage has begun.

Our work and that of Mueller and Tilton and of Klein claim that as an industry stabilizes – that is, as technological progress slows down and production techniques become standardized – barriers to entry increase. The most attractive market niches will already be occupied. As process integration progresses, the cost of production equipment rises dramatically. Product prices and production costs will fall, so that firms with the largest market shares will be the ones to benefit from further expansion.

An existing distribution network may also be a powerful barrier to entry, particularly to foreign firms. However, a strong distribution network may also be a disadvantage. Underwood, for example, thought that its position in the market for business typewriters was secured by nearly one thousand proprietary outlets. But when IBM entered this market with electric typewriters it was able to overwhelm Underwood’s advantage by marketing through its 2300 office equipment dealerships. Similarly, when Sony introduced the transistor radio most radios where sold and repaired through manufacturers’ own distributorships and retail stores. Sony simply circumvented this seeming barrier by selling directly through mass merchandisers. Transistor radios were simply mailed to a repair center or replaced if problems were encountered.

A hallmark of stability is a concerted drive among the surviving firms toward tightening the control over the value chain. This process can take one or both of the following forms: (1) an improvement in the relationships with suppliers and distributors, toward a more integrated and cooperative relationship, and (2) the pursuit of vertical integration, i.e. direct ownership of the different stages in the value chain.

Earlier versions of the model presented here (e.g. Abernathy and Utterback [3]) considered vertical integration during the period of stability as an inevitable outcome of technological evolution in an industry. Here, we claim that what surviving firms really seek is “control over the value chain”. Vertical integration is but one possibility to achieve such control. Today, firms are increasingly relying in improved supplier and distributor relationships to achieve control over the value chain [7,29,50]. The emergence of a set of more or less captive suppliers of equipment and components linked to a large firm is commonplace. Although closer relationships with distributors may have similar benefits, most of the literature has focused on supplier relationships, particularly on describing the Asian (mostly Japanese) model. Indeed, the level of cooperation achieved by Asian firms and their suppliers is so high, that often technical teams from the parent firm work at a supplier’s site for extended periods of time, solving a specific production problem [5,18]. Suppliers are often invited by the parent company to participate in the design of a new product from its outset [17,23].

Suppliers may play a creative role if the parent firm is able to generate enough loyalty and cooperation. Indeed, such a close relationship suggests that we should probably re-think our very concept of a firm and its boundaries. Only when relationships with suppliers are not cooperative enough will there be a drive among producing firms to capture those elements of supply which create the greatest uncertainties for them — i.e. a drive toward vertical integration.

All of the factors above point to the hypothesis that large amounts of capital would be needed by a new firm entering late. Thus, it is no wonder that most radical innovations occur within new entrants attempting to break into an established set of competitors, rather than within firms whose capital and resources are tied up in the existing technology. Indeed, it is just at the point of stability that firms may become locked into narrow positions that may ultimately increase their vulnerability. An existing distribution network may suddenly be threatened by a new technology
that requires sharply reduced servicing or maintenance, or by the entry of a large firm with an even stronger distribution network or a broader product line. Existing production management techniques may become superseded by more effective practices. An existing patent may expire. Mueller and Tilton contend that industries become stable when patent positions expire, and Klein contends that vertical integration tends to buffer larger firms from new competitors. Our hypothesis is that a period of stability in industry structure and market share is more likely to be a harbinger of invasion of the industry by a functionally superior but somewhat more costly technology. Following a period of stability a new wave of product and process change — or, in a few cases, the revitalization of the dominant technology itself via significant product improvement or the use of new production techniques and technology, may be expected to occur.

In summary, we expect the development of a set of competitors to begin with a wave of entry gradually reaching a peak at about the time that the dominant design of the major product emerges, and then rapidly tapering off. The wave of entry will be followed by a corresponding wave of exits of firms from the industry. The sum of the two curves will yield the total number of participants in the product market segment at any point. This total curve will usually start with a gentle rise representing the first few fluid productive units entering the business followed by a much sharper rise which represents a wave of imitating firms. The point at which a dominant design is introduced in the industry is expected to be followed by a rather sharp decline in the total number of participants until the curve of total participants reaches the stable condition with a few firms sharing the market. A major technological discontinuity would start a whole new cycle again.

Related concepts

Other contemporary concepts related to the model proposed here are those of technological paradigms and trajectories, technological guideposts, discontinuities, product design hierarchies, and notions from the population ecology literature. Let us discuss briefly how they relate to our work.

Population ecology predicts that the surviving organizations in a environmental niche would be those best fitted to resist a process of “natural selection” [32,37]. If we consider business firms as the organizations in this theory and think of a market as the environmental niche, population ecology predicts a strikingly similar pattern to the one proposed by our model. In our terminology, the population ecology model predicts a rapid increase in the number of firms at the beginning of the industry; a peak is reached later on and thereafter the population of firms declines. This is also what the model proposed here predicts. Indeed, the graphs based on data from the population ecology literature (see, for instance, Caroll and Hannan [14]) look much like the ones the reader will find in this paper.

There is one major difference between population ecology models and the one presented here, however. Population ecology, at least in its stricter presentations, assumes that organizational inertia is always present. Organizations do not change, they are selected out of a niche by better-fitted competitors. In contrast, organizational change is a hallmark of our model. For us, organizational change is driven by technological change in the industry. Organizations are not born to win or lose a natural selection process. Instead, some of them evolve through changes in their administrative and production structures triggered by technological change to become the surviving ones.

Population ecologists distinguish two basic types of organizations: specialists and generalists. Roughly, generalists will tend to be selected out in stable environments, where specialists are better fit. Conversely, specialists will tend to be selected out in volatile environments, where adaptation capabilities are an asset, and thus generalists have a clear advantage over more rigid organizations. This taxonomy is not inconsistent with our model. We could think of firms in

---

2 The notion of an “industry” is always somewhat obscure, for its limits are difficult to define (this problem is also present, and perhaps more strongly, in the notion of niche). We think of an industry as composed by a product class, i.e. a group of similar products that serve the same market need and thus compete directly in the market-place.
the early stages of an industry as generalists, and of those which survive after the emergence of a dominant design as specialists. Our model predicts that the successful firms in an industry will be those which manage to mutate from generalists to specialists, in concert with the evolution of technology in that industry.

Other scholars have developed related concepts. Devendra Sahal, for instance, has coined the concept of “technological guideposts” [50b]. For Sahal, technological change is characterized by technological guideposts, which are major advances in technology capable of setting a direction to be followed by more incremental developments. In his framework, technological guideposts are chosen (among many alternatives) essentially by chance. Only when a guidepost is established, the more rational, predictable process of advancement along the line set by the guidepost begins.

Giovanni Dosi, borrowing heavily from Kuhn’s 1962 work on scientific advances, has proposed the concepts of “technological paradigms” and “technological trajectories” to refer to the same phenomenon [24]. Roughly, a technological paradigm is a pattern of solution of selected technological problems. Technological paradigms define some idea of technological progress, i.e. they point out specific technological trajectories. Technological trajectories are relatively minor technological developments along the pattern set by a paradigm.

Both Dosi’s and Sahal’s concepts are related to Nelson and Winter’s notions of “technological regimes” and “natural trajectories”, and to Clark’s notion of “design hierarchies” [16,48]. Clark’s notion is the most closely related to the concepts of the model proposed here. In Clark’s view, patterns of innovation are the result of the logic of problem solving in design and the formation of concepts underlying customer choice. Both processes are seen as imposing a hierarchical structure on the evolution of technology. The choice of a core technical concept (e.g. internal combustion engine in the auto industry) establishes an “agenda” for a product’s technical development. Similarly, problem solving on the customer side also helps establish the agenda for technology development. The well-known model T, for instance, came into being not only because of the outcome of technical decisions made earlier by designers. It was also the result of the evolution of customers’ awareness and preferences about automobiles, in particular the demand for a durable, reliable and low cost means of “basic transportation” [1].

We also view the emergence of a dominant design as the result of the interplay between technical and market choices. Indeed, our notion of a dominant design can be represented using Clark’s design hierarchies approach. Figure 1 below illustrates a simplified case of a design hierarchy. A technological trajectory is the path of technical progress established by the choice of a
core technical concept at the outset (there are two trajectories in the figure). A dominant design is the outcome resulting from a series of technical decisions about the product constrained by prior technical choices and by the evolution of customer preferences. A dominant design often does not represent radical change, but the creative synthesis of the available technology and the existing knowledge about customers preferences. It may not be an ideal choice in a broader context of optimality, but rather a design, such as the familiar QWERTY typewriter keyboard, that becomes an effective standard. We now know how to design better keyboards, but the difficulties of moving from a well established trajectory to a different one clearly can be substantial [20,21].

Clark correctly points out that "radical" technological change will be associated with a movement up the design hierarchy, i.e. when existing core concepts are challenged. Along the same vein, the notions of technological paradigms and guideposts are easily associated with the well known concept of technological discontinuities [30,62b]. In our model, technological discontinuities correspond to the origin of what we have called an industry (see footnote 2). The model we have outlined in the previous section holds for a given technological paradigm, guidepost, or discontinuity. A discontinuity is considered in our model as the beginning of a "new" industry. For instance, our data below only contain information for the mechanical typewriter industry. The advent of the electric typewriter, a discontinuity, changed completely the rules of the game for the mechanical typewriter firms. A new industry was born, which would also be subject to the postulates of our model.

Recently, Tushman and Anderson have borrowed the Abernathy–Utterback concept of dominant design and that of technological discontinuities to provide further evidence for the hypotheses considered here [6]. They have gathered valuable data on the minicomputer, glass, and cement industries and performed tests of the model. Their work not only provides more data to support the Abernathy–Utterback model, but also enhances the latter in several respects. They provide, for instance, additional insights on the emergence of dominant designs, by looking at this problem as the result of a political process (the authors have yet to fully explore this interesting idea). They also make an insightful distinction between competence-enhancing and competence-destroying discontinuities, an idea based on the taxonomy proposed by Abernathy and Clark [2]. A competence-enhancing discontinuity builds on existing know-how in the industry, while a competence-destroying one renders existing knowledge obsolete. Tushman and Anderson argue that each type of change will have a different effect on the dynamics of competition in an industry.

Most of the concepts and frameworks mentioned in this section — including ours — can be considered manifestations of a more general "punctuated equilibria" model [24b]. Such a general model is basically composed of relatively long periods of stability (equilibrium), punctuated by compact periods of qualitative, metamorphic change (Gersick [33]; the author reviews applications of the model to different fields). Although the construction of such "grand theory" of change is a challenging intellectual enterprise, we shall limit our claims in this paper to manufacturing industries. It is our hope that insights gained through more focused studies like this one can provide the right pieces to assemble a larger intellectual puzzle.

We will now turn to a discussion of some specific examples which exhibit the waves of change our model predicts: the manual typewriter industry, the auto industry, the electronic calculator, semiconductors, television, television tubes, and massively parallel supercomputers. We will then turn to the question of successive waves of entry, and again present specific examples that illustrate the applicability of the hypotheses presented in the previous section. Though we have a rather small number of observations (industries), each example examined so far conforms in its major features to the hypotheses stated above.

Summary of data

In general, each of the eight industries studied here presents a similar pattern of firms' entry and exit. American firms, usually small entrepreneurial firms, enter at the dawn of a new industry at a moderate pace. Later, a rapid wave of entry occurs, raising the slope of the curve of total number of firms in an industry. This can be seen in fig. 2, which summarizes such curves for the
different industries in this study. After a dominant design is established, the total number of firms declines steadily until it reaches a point of stability — a few large firms remaining in the industry supply most of the demand. Successful firms often enter the industry early.

Table 1
A list of dominant designs by industry

<table>
<thead>
<tr>
<th>Industry</th>
<th>Dominant design</th>
<th>Date</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typewriter</td>
<td>Underwood’s Model 5; Hess’s innovations</td>
<td>1906</td>
<td>Engler [27]</td>
</tr>
<tr>
<td>Automobile</td>
<td>All steel, closed body</td>
<td>1923</td>
<td>Fabris [28]; Abernathy [1]</td>
</tr>
<tr>
<td>Television</td>
<td>21-inch set; adoption of RCA’s technical standards</td>
<td>1952</td>
<td>John Rydz; Television Factbook [57]</td>
</tr>
<tr>
<td>TV tubes</td>
<td>All-glass, 21-inch tube</td>
<td>1956</td>
<td>John Rydz; Television Factbook [57]</td>
</tr>
<tr>
<td>Transistor</td>
<td>Planar transistor</td>
<td>1959</td>
<td>Tilton [58]; Braun and MacDonald [11]</td>
</tr>
<tr>
<td>Integrated circuit</td>
<td>Multiple designs in rapid succession</td>
<td></td>
<td>Dataquest; Flaherty</td>
</tr>
<tr>
<td>Electronic calculator</td>
<td>Calculator on a chip, D.D. not yet identified, but we speculate it will be a variant of the hypercube architecture</td>
<td>1971</td>
<td>Majumdar [45]; Afuah and Utterback [4]</td>
</tr>
</tbody>
</table>
A list of the industries with their respective dominant design dates can be found in table 1. Our sample includes both large and relatively small industries. In terms of total revenue, the automobile and televisions are the largest industries. It is interesting to note that the peak in number of firms in these two industries occurs at a higher level than all other industries. Larger industries seem to attract more firms. The smallest industry in terms of revenue, massively parallel supercomputers, shows the smallest number of firms at its peak. Also, the sample contains industries with both long and compressed periods of competitive turbulence. Purely mechanical typewriters existed for nearly 70 years before the electric typewriter appears in the market. In contrast, transistors, integrated circuits, and supercomputers are cases in which the competitive turbulence we have observed appears to have occurred in a shorter period of time, though the latter two industries have yet to reach a state of few participants with stable market shares.

The typewriter

According to Engler [27] the typewriter industry began in 1873 with the entry of the Remington Arms Company. Their typewriter was a synthesis of many existing elements. Clockwork suggested the idea of the escapement (to move the carriage one letter at a time). A telegraph sender provided parts for the first model for keys and arms. A sewing machine pedal was used for returning the carriage. The piano contributed the concept of the free and swinging arms and hammers for imprinting the letters. The industry's initial growth was slow, and Remington had essentially a monopoly for the first few years. By 1885, the field had widened to five competitors. A period of rapid entry followed, and by the early 1890s, 30 firms had been established. Of these, Underwood and Smith were the principal innovators. Underwood, a supplier of ribbons and other typewriter materials, entered in 1895. Smith was a 1903 spin-off from the Remington Union Com-

![Fig. 3. Number of firms participating in the typewriter industry in the US. ■, Entry; ◻, exit; ✦, total.](image-url)
pany with a product innovation — a visible, front strike typewriter — that was incompatible with the Remington product line.

In 1904, the Royal Typewriter Company, the last of the four firms which were to dominate the industry, was established. By 1909, almost 40 companies competed in the typewriter industry (data from Engler [27]). Many others which entered and quickly exited, without making any significant penetration, have been left out of the records. Most of the actively competing firms were started between 1886–99, forming the sharp wave of entry hypothesized.

The set of features which was to become the dominant design in the typewriter industry resulted from a fascinating sequence of major innovations. In 1899, Underwood introduced the Model 5. The Model 5 had visible writing, rather than having the page obscured. It was the first to have a tabulator as part of the typewriter, to be able to cut stencils, and to make copies; thus, it did not need expensive attachments to do most of the jobs encountered in the office and consequently won Underwood a large share of the market.

During the years following Underwood’s introduction of the Model 5, Edward Hess, a man with exceptional mechanical abilities, perfected many of the features that were still rough in the Model 5. By rearranging the clutter of knobs, bars, and ribbon mechanism, he was able to deliver “total, uncompromised visibility.” He reversed the linkage in the typebar action, so the action was a pull rather than a push, thus saving energy. He removed much of the friction from the escapement — the toothwheel that links the keys with the carriage and moves it along one space when a letter is struck. These and other innovations gave Hess’ typewriter a light, fast touch welcomed by users. Hess received 140 typewriter patents during his lifetime. One of Hess’ major concerns, and one that has direct implication for our model, was to reduce typewriter production costs by improved design.

After Underwood’s Model 5 and Hess’ innovations, competition centered mainly around features and increasingly on production costs. Figure 3 shows the pattern that our model predicts. The rapid growth in the number of firms halts in 1899, the same year when Underwood’s Model 5 was introduced. After 1906, when most of Hess’ innovations were in place, the number of firms in the industry begins an irreversible decline. Incidentally, Underwood, which had been a major innovator with its Model 5, lagged in bringing out new developments, and within a decade, had lost its dominant position to Royal.

The period between 1906–40 was a period of rapid reduction in the total number of firms. By 1940, there were only five predominant producers: IBM and the four traditional firms. According to Engler [27] each of the four — Remington, Royal, Smith and Underwood — had 20 percent of the market, with IBM having approximately 10 percent, and others (mostly foreign) 10 percent. The efficient size for a single plant was between 10 percent and 30 percent of total demand, or 150,000 to 450,000 units per year. Relative costs of production were substantially higher for a plant below 10 percent of market share.

In summary, more than 90 percent of the firms that had entered the industry had disappeared, either through bankruptcy or, in a few cases, through merger. Only a few early, innovative entrants had survived.

The automobile

More than 100 firms entered and participated in the American automobile industry for a period of 5 years or longer. The data for this industry are presented in graphic form in Utterback [61] based on data from Fabris [28]. The wave of entry began in 1894 and continued through 1950, followed by a wave of exits beginning in 1923 and peaking only a few years later, although it has continued until the present day.

As hypothesized, entry began rather slowly, but then accelerated rapidly after 1900, with participation in the industry reaching a peak of 75 participants in 1923. In the next two years, 23 firms, nearly a third of the industry, left or merged, and by 1930, 35 firms had exited. During the ensuing depression, 20 more firms left. There was a brief flurry of entries and then exits immediately following World War II, but the number of US firms in the industry has been basically stable since 1940.

The number, and scope, of major product innovations is reflected in this pattern of entries and exits. The year with the largest number of
firms, 1923, was the year that Dodge introduced the all-steel, closed body automobile. The large number of exits over the next few years corresponds to the fact that by 1925, 50 percent of the US production was closed, steel body cars, and by 1926, 80 percent of all automobiles were of this type [1]. The post-World War II stability in market shares and number of firms reflects the fact that approximately three-quarters of the major product innovations occurred before the start of the War.

New concepts in product accessories and styling were tested in the low volume, high profit luxury automobile. Conversely, incremental innovations were more commonly introduced in lower priced, high volume product lines. General Motors appears to have led in both types of innovations, particularly for major product changes. In certain years, engines show a higher annual magnitude of changes; these changes, however, occur with less frequency than do those in chassis characteristics; body plants are more flexible and continuously changing than are engine plants, which tend to change occasionally in an integrated and systematic way.

From 1894 to 1918, 60 firms entered and none exited. We do not have data on innovations for this period, but we assume that architectural innovations overwhelmingly predominated. From 1919 to 1929, 22 firms entered and 43 left during this period, 14 of 32 major product changes occurred, nearly half of the total. From 1930 to 1941, 6 firms entered and 29 exited, and 11 major innovations occurred. Finally, from 1946 to 1962, 4 firms entered and 8 left, but only 7 major innovations were introduced. One can see a continual decline in major product innovations over these three periods [28, pp. 85–93].

**Television and television tubes**

Research leading to the appearance of television started several decades before the first suc-
cessful results were achieved. RCA entered the industry in 1929 after Sarnoff, impressed by a demonstration by the inventor Vladimir Zworykin, decided to hire Zworykin and put him in charge of RCA's Electronic Research Group in Camden. Several other firms or inventors-entrepreneurs entered the infant industry during the 1930s, and all of them contributed to expand the frontier of technical knowledge. Philco, Philo Farnsworth, Louis Hazeltine, American Television, and Allen DuMont are some of the most important names.

The commercial birth of the industry can be traced back to the 1939-40 New York World Fair, where millions of Americans saw television displays for the first time. For the purpose of our analysis, 1939 marks the beginning of the industry. Our data in fig. 4 start only in 1949; however, the data clearly show that the television industry conforms to the hypothesized relationships. The first decade of the industry (the dotted lines are our estimates of the real curves' shape for that decade) obviously witnessed a rapid increase in the number of firms. The wave of entry most likely peaks in 1950, the first year of our entry data, or one year earlier. The total number of firms steadily increased until 1952, the year in which it peaked at 85 firms. Also, in 1951 the exit wave takes off, to peak around 1956.

Several things happened in the early 1950s in the television industry which had a significant impact on the pattern of innovation and competition that was to follow. First, the uncertainty about technical standards for color broadcasting (i.e. UHF vs. VHF) was finally resolved by the Supreme Court in 1951. Later, in 1953, the FCC approved the NTSC system, backed by a group of manufacturers headed by RCA. Several firms which had opposed the RCA technical standards dropped out of business due to this legal verdict. Second, several features of the television sets converged to form a dominant design around 1952. The most important dimension of the dominant design was the size of the screen, and there-

![Fig. 5. Number of firms participating in the cathode ray picture tubes industry in the US. ■, Entry; □, exit; ◆, total.](image-url)
fore the characteristics of the picture tube. The first monochrome set produced by RCA was a 10-inch set. Almost all sets produced in the 1940s had screens smaller than 14 inches. RCA produced its first 21-inch set and other large screen sets around 1952, and they soon became the market standard. Third, during the early 1950s, RCA started to license its television technology to other firms, which further reduced competition and also supports the idea that RCA held rights to most of the key characteristic of the product at that time.

The story of the TV picture and receiving tubes is undoubtedly related to that of the television industry itself. Figure 5 depicts a rapid increase in the total number of tube producers from 1949 to 1956, analogous to that previously discussed for the television industry. In 1956, 4 years after the peak of the television industry's curve, the total number of firms in the tube industry reaches its peak at 66 firms. The entry wave also peaks around that time, registering more than 40 entries in the period 1953–55. The wave of firms exiting the industry takes off slowly during the early 1950s, and reaches a peak in 1958, with 15 firms leaving in that year.

One of RCA's major achievements was the development of the shadow mask color picture tube. The FCC ruled in favor of the RCA compatible color standard on 17 December 1953, and programs were first broadcast in this format in January of 1954. Hazeltine made major contributions to brightness in order to make bright whites. Some problems in manufacture constrained the size of the tubes to 16 inches until CBS laboratories learned how to curve the mask. RCA licensed both of these developments. A problem with the initial tubes was that they were metallic with glass bonded to the front, and this interface proved to be troublesome. For the purpose of this research, the 21 inch, all-glass picture tube is considered the dominant design. The 21 inch, all-glass tube (first black and white, then color) was by far the biggest seller since the mid-1950s.
The first such tube, black and white, was introduced by RCA late in 1955, and captured a significant portion of the TV tube market during the rest of that decade. The advent of reliable color tubes contributed to the difficulties that many tube producers were experiencing during the late 1950s and fostered concentration in the industry. Indeed, on 8 May 1958, RCA publicly announced the first all-glass 21-inch color picture tube, which was to hold a large share of the color tube market. 3

**The transistor**

According to Tilton [58], three firms dominated the receiving tube business in the US in 1950: General Electric, Philco Ford and RCA. Over the ensuing decade, these firms spent more on R&D and received more patents than did new firms entering the industry (counting the expenditures and patents of the Bell Laboratories), but a cohort of new entrants gained nearly two-thirds of the market. By 1966, three of these new entrants — Texas Instruments, Motorola, and Fairchild Semiconductor — together accounted for 42 percent of the market, while the vacuum tube manufacturers’ share of the transistor market had declined to just slightly more than one-quarter of the total. The difference between established and new entrants in the business would be even more dramatic if one included the IBM’s production for its own use (it is believed to have entered production in 1961).

Figure 6 shows that the total number of firms in the industry starts to rise rapidly with the announcement of the invention of the transistor in 1948. The rapid increase in the total number of firms is virtually halted around 1959, the first year in which the planar transistor was in com-
commercial production. Transistors produced through the planar process rapidly became the dominant design in the industry. First introduced by Fairchild Semiconductors, planar transistors presented many advantages over the older mesa transistor technology. In particular, the planar transistor was flat, which meant that electrical connections could be achieved by depositing an evaporated metal film on appropriate portions of the wafer. This was a great advantage over the mesa transistor, whose irregular surface dictated that electrical connections be done laboriously by hand. The planar transistor prompted a drive for producing low-cost transistors typical of after-dominant-design stages of the industry cycle. The reader would note that exists in the industry, almost nil before 1959, become commonplace afterwards.

The years 1963 and 1964 saw the development of epitaxial growth and epitaxial reactors for producing integrated circuits and further development of process integration. We will discuss integrated circuits in more detail below as a separate case.

The integrated circuit

The integrated circuit industry is the only one in our sample that does not clearly conform to our hypotheses. In fact, fig. 7, displaying only data on US firms, shows no clear peak in the number of firms in the industry during our data period. Therese Flaherty, who is currently studying the integrated circuit industry, suggests that no one product of any generation can be easily considered a dominant design. The integrated circuit has kept on changing substantially from generation to generation, which may explain the very broad plateau we observe in total industry participation with continuing underlying entries and exits.

There are several generations of integrated circuits. For instance, DRAM, the most important segment of the integrated circuit industry, has had seven generations up to the present time (1K, 4K, 16K, 64K, 256K, 1MB, 4MB). Competition has been tough both within and between generations. No one firm has been able to maintain a leadership position from one generation to another. In general, American firms have been loosing ground to Japanese entrants. The first two generations were dominated by American firms; however, starting with the 16K generation, Japanese firms take a significant share of the market. The industry has grown very rapidly over time. the first generation (1K) had a maximum annual revenue of 152 million in 1977, while the 256K generation — the last one for which revenue data are available — reached an annual level of 1,807 million dollars in 1987. This growth happened in a 15-year time span.

Despite the fact that a quick look at the integrated circuit industry suggests that our model has limited explanatory power here, we would like to point to several issues that cast shadow on such first-sight conclusion. To begin with, most of the entries occur during the early years of the industry. This is especially true for American firms, as can be seen in fig. 7, but it is also true, to a lesser extent, for Japanese or “foreign” firms. Moreover, “enter early” seems to be a winning strategy within each generation, as Flaherty has pointed out. Second, the production capacity of dominant firms has been increasing — relative to total market demand — throughout the generations. The trend is that fewer companies are increasingly able to satisfy most of the demand.

Our model hypothesizes a similar increase in industry concentration. Indeed, it should be noted that exits of American firms from the industry take off in 1985, increasing steadily for the next 2 years. As American entries are nil during this period, the total number of American firms declines in 1987 to one of its lowest levels. Higher concentration and larger firms is the prevailing pattern in the industry today. Finally, although product innovation is still important in later integrated circuit generations, process innovation and production capabilities are increasingly critical as generations pass and greater production volumes

4 We would like to thank Dataquest and Prof. Therese Flaherty of the Harvard Business School for providing us with access to her data for the integrated circuit industry, as well as for her helpful comments and insight about the industry dynamics.

5 The data for foreign firms is not displayed because of space limitations. Non-US firms are not only, but mostly, Japanese in our data.
are required of participant firms. Such production capabilities form an effective barrier to entry in the industry.

The electronic calculator

The American calculator industry in the early 1960s consisted of five major companies manufacturing electro-mechanical machines that controlled nearly 90 percent of the market — Frieden Monroe, Marchant, Victor, and Olivetti. Frieden, Marchant and Monroe each had approximately 20 percent of the market, Victor a slightly smaller share, and Olivetti 10 to 15 percent. These companies were almost completely vertically integrated due to the need for a high degree of precision in the manufacture of many specialized parts. There were strong barriers to entry to new firms. By concentrating on specific segments of the market, the major companies avoided intense competition. They also had reinforced their market dominance by setting up extensive distribution and service networks. In addition, through nearly a century of continuing modification and perfection, the technology of electro-mechanical calculators had reached a high level. Thus, it was not easy for anyone to come up with a dramatic breakthrough that would threaten the status quo.

This situation did not change initially when the electronic calculator entered the market in 1962. The first electronic machines were extremely complex and expensive having more than 2300 discrete parts, and were aimed at specialized scientific and technical market segments. Figure 8 shows (with data from Majumdar [45]) that between 1962–70, 11 firms entered the industry, with ten of them surviving. The wave of entry peaked in 1972 with the entry of 21 firms in a 3-year period, as shown in the figure. However, exits begin in 1971, rising sharply during the next few years. 1971 marks the year of the introduction of the dominant design of the calculator on a chip, which made the assembly of units extremely simple — merely piecing together the chip, display device and keyboard. The entry of semicon-

![Fig. 8. Number of firms participating in the electronic calculator industry in the US. ■, Entry; □, exit; ●, total.](image-url)
ductor manufacturers, such as Texas Instruments and Rockwell in 1972, and National Semiconductor in 1973, further precipitated the departure of firms which were largely assemblers of purchased components. The industry's structure then appears to stabilize, with even a few of the semiconductor makers, such as Rockwell, dropping out, and a small number of the remaining vertically integrated companies dominating the market. Thus, the appearance of a dominant design, and the drive toward vertical integration which normally follows its appearance, were almost concurrent in this highly compressed example.

Supercomputers

Supercomputers, i.e. the most powerful computational systems at any given time, today achieve speeds in the 100 MFLOPS (million floating point operations per second) range. Three major technologies have been used to build supercomputers: sequential, vector, and parallel processing. Sequential computers, whose architecture is often referred to as von Neumann, have only one central processing unit (CPU); they do one thing at a time. Vector processors allow simultaneous computation for some problems, such as problems with vector-like or matrix-like structure. Parallel processing, or more specifically massively parallel processing, is a computer architecture where hundred or thousands of processors are put on the job simultaneously to get the job done faster than more traditional supercomputers and with greater generality.

Traditional supercomputer makers — such as Cray, Fujitsu, Hitachi, IBM, and NEC — produce mostly Von Neumann machines with some having vector processors to boost performance. IBM and Univac are considered the first entrants into the supercomputer industry. Cray Research entered in 1972 to become the presently dominant player in sequential supercomputers. A second set of firms, minisupercomputer makers, use the Von Neumann architecture with the associated incremental innovations of pipelining and vector processing, but build less powerful machines which target low-end applications with price-sensitive customers. Massively parallel computer (MPC) makers are the latest entrants in the supercomputer industry. Firms such as Thinking Machines, Intel, Floating Point Systems, and Meiko started production around 1985, while the MPC “pioneers” Ametek, Myrias, and Good year Aerospace entered the industry only as far back as 1983 (data from Afuah and Utterback [4]).

There are two issues of interest in the supercomputer industry in the light of our hypotheses. First, at a more aggregate level, we suggest that the massively parallel architecture will become the dominant design in supercomputers. Therefore, it is likely that we see some exit of traditional firms from the industry in the future, and large players such as IBM or Cray turning to the MPC architecture. Second, although numerous MPC design exist today, Afuah and Utterback [4] forecast that some variation of the hypercube configuration will prevail.

Summary

Bearing in mind that the data reviewed here are derived essentially from US firms producing assembled products, the appearance of a dominant design does indeed seem to shift emphasis in an industry from predominantly entrepreneurial product innovation toward process development, scale of production, production management techniques, elaborate research programs aimed at planned incremental change, and correspondingly to advantages for larger firms. The dominant design which emerges is not necessarily the result solely of technical potentials, but also of timing, collateral assets and other circumstances. Once a degree of standardization is accepted, however, major innovations from within an industry seem less and less likely to occur short of a wave of new entrants and increasing competition.

Strategies for entry, development, and resource allocation should, we believe, bear a relationship to the current state of the technical development of an industry's product and process, as well as the degree to which these are integrated. Counter strategies are likely to be highly risky, and thus should be thought out with great care. While the process described is not necessarily irreversible, the evidence to date indicates that it is highly directional. Surviving to become a dominant firm is an improbable event.
(the outcome for only 5–10 percent of US entrants in the cases examined) doubtless requiring superb execution of product design, process development and organization, and timing in what is at least initially a highly uncertain environment.

Failures of firms in our analysis seem to be a matter of weakness of technical resources or slowness in development, not simply of lack of scale or market power. Thus, mergers of weaker firms are most often seen in the cases above to be followed by failure. Few of the mergers noted during the waves of exit shown above appear to be truly complementary ones which extend product lines and markets. Normally in the cases studied, mergers which occur after a dominant design appears quickly fail. The pattern shown by the emerging dominant firms appears to be one of growing internal technological and manufacturing strength rather than one of successive mergers with competitors. Development of close and technically creative supplier relationships appear to be keys to successful, continuing dominance in many cases. There is also the potential that such closed relationships may result in even more rigidity ad resistance to change by entrenched firms.

Questions for further research

There is suggestive evidence, both in the literature and in the cases examined, that product performance and cost are strongly dependent on entry, exit and growth of competing firms. In particular, performance improvement seems to reach a maximum as firms pour into an industry during its formative or architectural period, prior to the dominant design. Klein has shown this to be true in the case of autos and aircraft engines. Modis and Debecker [46] clearly demonstrate that periods of greatest improvement in the performance of microcomputers correspond to waves of newly entering firms. Conversely, cost reduction may reach a maximum as firms struggle for dominance during the time of most rapid exit from an industry. The peak period of exits is a period of rapid revenue growth both for the industry as a whole, and especially for the emerging dominant firms as they increase their scale of operations. These are important questions bearing centrally on issues of technology strategy. They may be well worth the effort needed to gather data in greater detail on firms' market shares, product performance and costs where possible.

The summary curves of industry participation in fig. 2 show striking differences in the total numbers of firms entering different industries. Does the total number of firms vary with the size of the market for each product? It seems reasonable that the total markets for autos, televisions and tubes, and typewriters are larger than those at any corresponding point in development for calculators or supercomputers.

Does a rather flat summit or plateau in total industry participation, as is the case for typewriters and integrated circuits, reflect either continuing technological change in a product or many rapidly succeeding generations? Conversely, does the sharpness of a peak in total industry participation reflect strong economies of scale as in autos and calculators, and complementary assets such as software written or a particular architecture, as in the case of supercomputers? Can we infer the occurrence of a dominant design from the occurrence of a peak in participation and the beginnings of consolidation in an industry? This certainly seems to be the case for massively parallel supercomputers, as Afuah and Utterback [4] have predicted.

Late entry has not seemed to be a viable strategy for US firms. However, late entry and tenaciousness seem to be the common strategy for Japanese entrants in the four cases for which we have data (Japanese data were not presented in this paper). Are highly integrated Japanese firms better able to absorb the heavy investments and learning that such a strategy implies and to remain prosperous survivors than US firms? Some observers have pointed to the fact that all major Japanese semiconductor producers support consumer electronics or computer production as an explanation for their apparent ability to survive low levels of profitability and return. In other words, businesses are judged as strategic assets contributing within a larger complex, rather than as independent business entities. Some of the implications of our work here for organization structure on a macro-scale seem to be well worth further examination. Also with respect to Japan, it remains to be seen whether Japanese firms will continue entering late now that their research
capabilities are becoming comparable with those of American firms.

Finally, we have not examined cases of highly capital intensive, process driven industries such as petrochemicals, fibers or glass. We suspect that much different patterns may exist in such cases due to high barriers to entry. For example, with regard to the glass industry the peak in peaches due to high barriers to entry. For example, with regard to the glass industry the peak in industry participation for plate and flat glass appears to have occurred prior to the introduction of the Float process. However, the Float process developed by Pilkingtons did lead to a rapid wave of exits and consolidation both in the glass industry and in product types. The technological innovator licensed firms such as LOF and PPG, but while the Float process diffused in this manner, there were no new firms formed to enter the business. The Float process created a major step in economic plant scale at the same time.

If our conclusions are correct when more generally examined and tested, then any action based on a static analysis of a firm’s strengths and strategy will probably reduce that firm’s chances for long run survival, much as highly specialized animals fail to survive slight shifts in climate and habitat while generalized foragers prosper. This is the more strikingly clear if a firm has as its ambition surviving a generational shift in technology. Most firms fail to survive, at least as players in a particular product arena, even in the competitive shake out seemingly precipitated by the synthesis of a dominant product design. The number able to survive a generational shift is apparently vanishingly small. The patterns observed here clearly imply that a firm should suboptimize in the short term in order to build the flexibility, skills, and resources it will almost surely need if it is to become a dominant survivor in the longer term.

References

[24] Giovanni, Dosi, Technological Paradigms and Technological Trajectories: A Suggested Interpretation of the


[62b] James M. Utterback and Linsu Kim, Invasion of a Stable Business by Radical Innovation, in: Paul Klein-

