Recombinant Uncertainty in Technological Search

Lee Fleming
Harvard University, Graduate School of Business, Morgan Hall T97,
Boston, Massachusetts 02163
lfleming@hbs.edu

While the course of technological change is widely accepted to be highly uncertain and unpredictable, little work has identified or studied the ultimate sources and causes of that uncertainty. This paper proposes that purely technological uncertainty derives from inventors' search processes with unfamiliar components and component combinations. Experimentation with new components and new combinations leads to less useful inventions on average, but it also implies an increase in the variability that can result in both failure and breakthrough. Negative binomial count and dispersion models with patent citation data demonstrate that new combinations are indeed more variable. In contrast to predictions, however, the reuse of components has a nonmonotonic and eventually positive effect on variability.
(*Invention; Search; Recombination; Negative Binomial Dispersion Model*)

Introduction

While many have acknowledged the pervasive uncertainty of technological change (Rosenberg 1996), the ultimate sources of that uncertainty remain poorly understood. Many scholars in the product life-cycle tradition have observed that uncertainty peaks early and decreases following convergence on a dominant design (Anderson and Tushman 1990, Klepper 1997). Radical and destabilizing change and the sources of the life cycle are particularly difficult to predict and often attributed to luck or individual genius (Tushman and Anderson 1986, Ayres 1988, Mokyr 1990). Clark (1985) proposes that technologies evolve and uncertainty decreases as inventors iterate between the needs of customers and the technological logic of their current trajectory. Much research remains agnostic about the causal sources of uncertainty and simply models it as stochastic draws (Nelson and Winter 1982, Klepper 1996). Proponents of bounded rationality often characterize technological change as an intrinsically uncertain and, to varying degrees, blind search process (Nelson and Winter 1982, Vincente 1990, March 1991). While the assumptions of bounded rationality may be more accurate than those of classical economics, the metaphor of search remains informal and was only recently developed empirically (Stuart and Podolny 1996). Despite this widespread acknowledgment of the importance of uncertainty, most research gives the topic brief consideration en route to other issues, and little work has attempted to identify and empirically validate the causal sources of uncertainty.

Part of the difficulty in explaining the ultimate source of uncertainty is the conceptual and empirical conflation of its many sources, and, in particular, technological invention and commercial innovation. There exist many causes of uncertainty in technological change besides purely technological sources, including the adoption and diffusion of new technologies, market and customer acceptance, and competitors' strategic actions (Rosenberg 1996). Arguments for intentional conflation of these sources notwithstanding (Ruttan 1959), our confusion could be reduced by returning to the classic differentiation between technological invention and commercial innovation. Schumpeter (1939) defined *innovation* as the com-

0025-1909/01/4701/0117$5.00
1526-5501 electronic ISSN

mercial application or adoption of an *invention*. He argued that, "the making of the invention and the carrying out of the corresponding innovation are, economically and sociologically, two entirely different things" (p. 85). While these sources remain interdependent, their conflation confuses our explanations of their respective processes and contributions to uncertainty. To avoid such confusion, this paper will focus on invention and purely technological sources of uncertainty.

Attempts to understand the source of technological uncertainty have also been frustrated by a lack of understanding of the process by which inventors create new technologies. A theory of the process of invention would facilitate our understanding of the sources of technological uncertainty. To address this issue, this paper synthesizes two classic perspectives on the sources of technological novelty: first, invention is a process of recombination, and, second, invention is indeed an inherently uncertain and, therefore, typically local search process.

Invention As a Process of Recombinant Search

Many scholars have proposed that recombination provides the ultimate source of novelty (Gilfillan 1935, Usher 1954). Schumpeter (1939, p. 88) observed that "innovation combines components in a new way, or that it consists in carrying out New Combinations." Nelson and Winter (1982, p. 130) state that "the creation of any sort of novelty in art, science, or practical life—consists to a substantial extent of a recombination of conceptual and physical materials that were previously in existence." Henderson and Clark (1990) argue that the mere rearrangement of previously used components can itself cause destabilizing industrial change. Hargadon and Sutton (1997) describe how a design consulting firm creates novelty by brokering ideas and technologies between their clients' industries.[1] Basalla (1988) develops an analogy from Kroeber's (1948) image of a tree of cultural artifacts. Unlike genetic trees, however, the branches of the tree of technology can fuse together. "Separate types or branches fuse together to produce new types, which merge once again with still other branches" (Kroeber 1948, p. 138). Similar to the theory of natural evolution before the discovery of DNA, the analogy breaks down at the genetic level. "We who postulate theories of techological evolution likewise have our Darwins but not our Mendels" (Kroeber 1948, p. 210).

Although the idea of recombination as the source of novelty has been widely discussed, the implications of the idea remain undeveloped. Previous work suggests, however (Schumpeter 1939, Henderson and Clark 1990), that an invention can be defined as either a new combination of components or a new relationship between previously combined components. While knowledge, science, algorithms, culture, applications, and manufacturing processes also influence invention, they are not part of an actual artifact and are not actually instantiated in an invention. They strongly influence the process of inventive search, however, by inspiring, aiding, explaining, or constraining the use of particular components or combinations. Colloquial usage of the word "components" implies known, available, and commercially available hardware. In this paper, however, "components" will denote the constituents of invention, along the lines of what Schumpeter calls "factors" (1939, p. 88).

The Scope of Potential Recombination

Even though the number of potential components overwhelms the imagination, there are no restrictions on the scope of their recombination. Components are not like genes and "similar" technologies are not like species (Basalla 1988). In contrast to variation processes within genetically isolated populations, inventors can recombine any components within their purview. Perceptions that certain technologies or components "belong together" develop through social construction and previous association. For example, if an electrical engineer of the 1940s had been asked about his profession's use of sand and aluminum, he probably would have replied with a blank stare. Today, he or she probably would reply that they are the most common basic materials of semiconductors

[1] While the concept of recombinant search as the source of novelty may well apply outside the strictly focused realm of invention—indeed, to cuisine, innovation (Abernathy and Utterback 1978, Kogut and Zander 1992, Levinthal 1998), and process improvement (Romer 1993)—I can only comment empirically on invention.

and the focus of much research investment. Clearly, no technology evolves independently of the entire world of made things. At any point in technological evolution, any component is at risk of being recombined with any other component. The made world evolves as a holistic, continuous, and interdependent web, and not as a disjoint assortment of separate trajectories or product life cycles.

Investors constantly import previously untried components from outside the extant made world, for example, the use of medicinal substances from tropical jungles or the exploitation of petroleum and natural gas, substances that had little application to technology prior to the 19th century. Investors also create new components through encapsulation and hierarchical modularization of component sets (Simon 1996, Baldwin and Clark 1999). Such "black box" engineering efforts make the underlying components less salient but do not destroy the recombinant potential of the components or the encapsulating module. As these processes make additional components available for further recombination, the diversity of the "made" world increases. Because all previously used components and inventions provide potential constituents for future inventions, the potential risk set for recombination is the entire extant made world. Because inventors also scan outside the made world, anything not yet derived from the natural world can also be considered as a potential constituent of invention, as part of the theoretical risk set. Recombination usually occurs, however, between components that are salient, proximal, and available for the inventor.

These ideas prompt us to look backward in time to consider the components and untried combinations that were available at the time of invention; they prompt us to look forward to predict which inventions are more likely to motivate further recombination. Prediction of further recombination depends on the fundamental tension between exploration of untried possibilities and exploitation of previous successes (March 1991).

Cognitive, Social, and Technological Influences on Recombination

Because the agents of recombination are people, the process of invention remains strongly influenced by cognitive and social phenomena. The most fundamental influence is a limitation on the number of potential components and combinations that an inventor can simultaneously consider. Because every invention can be incorporated in further recombinations, inventors' combinatoric potential has grown explosively (Weitzman 1996). The set of potential combinations and, a fortiori, the possible ways that each set of potential combinations can be combined has become essentially infinite. It has become impossible for individual inventors, groups, even entire communities of inventors to have more than an infinitesimal understanding of all these potential combinations and relationships. As a result of this combinatoric explosion, inventors and their organizations and communities must focus and recombine locally from a limited set of components and combinations.

These arguments follow the assumption of bounded rationality and local search (March and Simon 1958, Nelson and Winter 1982, Cohen and Levinthal 1990, March 1991, Kauffman 1993, Stuart and Podolny 1996). *Localness* corresponds to inventors' familiarity with their recombinant search space. Local search or *exploitation* (March 1991) occurs when an inventor recombines from a familiar set of technology components or refines a previously used combination. A pastry chef searches extremely locally when he mixes previously used dyes in a new proportion to create a novel frosting color. Distant search or *exploration* (March 1991) occurs in the opposite situation, when inventors try completely new components or combinations. The early auto industry provides many examples of successful and unsuccessful distant search including various power sources, pneumatic tires, brakes for each passenger, four-wheeled diamond configurations, and, in the 1930s, a combined car and plane configuration (Basalla 1988).

One would expect that local recombination is more certain and, on average, more successful. To the extent that inventors draw from familiar component sets and refine previous combinations, they are less likely to develop a completely useless invention. They also decrease their upside potential, however, of developing a radically different invention that is of much greater impact. Local recombination is more certain because inventors learn from past failures.

Inventors learn which components failed in previous inventions and stop using them. They learn to avoid the combinations and architectures that failed in the past. They winnow and bound the less successful regions of recombinant space (Vincenti 1990). This winnowing and bounding improves the average usefulness of inventive efforts but also decreases the possibility of wildly successful inventions.

As inventors reuse components they begin to understand and characterize those components. They begin to understand which components are more or less useful in different contexts. Such knowledge enables selection and exploitation of more appropriate components in future inventions. Understanding and knowledge increase with use; the greater the use of a component, the greater the knowledge of and familiarity with it. Knowledge can also be forgotten, however, and is therefore more potent the more recently it has been gained (Argote et al. 1990). More recent and frequent usage therefore implies greater knowledge and familiarity. Inventions that incorporate familiar components should be more useful because inventors can select more appropriate components. Because they can better predict the performance of the included components, incorporation of familiar components should also decrease inventive uncertainty and, hence, the variability of outcomes.

Inventors can draw on others' knowledge and experience in addition to their own. While social proximity certainly increases the ease and likelihood of sharing, knowledge about component use will still diffuse between organizations (Allen 1977) and technological communities (Bijker 1987). Many mechanisms facilitate knowledge diffusion, including personnel movement, personal friendships, organizational merger, education, reverse engineering, technical literature, and strategic alliances (Ahuja 2000). Although much knowledge will be lost or changed in its diffusion, sharing will take place at all social levels, between individuals, organizations, and communities. Taken together, these arguments imply the following.

Hypothesis 1. *Recombination of familiar components will increase an invention's usefulness.*

Hypothesis 2. *Recombination of familiar components will decrease inventive uncertainty.*

Parallel arguments hold for the refinement and improvement of combinations. Engineers gain experience with particular combinations by using them. As inventors learn which combination relationships or architectures are less useful, they avoid those approaches. Such learning and knowledge helps them improve their inventive efforts on average and decrease inventive uncertainty. As with knowledge about individual components, knowledge about combination relationships increases with frequency of use and decreases since last use. As with individual components, knowledge of combinations and their optimal relationships will diffuse throughout the made world. These arguments imply combination hypotheses that parallel the individual component hypotheses.

Hypothesis 3. *Refinement of familiar combinations will increase an invention's usefulness.*

Hypothesis 4. *Refinement of familiar combinations will decrease inventive uncertainty.*

Practitioners have long recognized the value of reuse and refinement (Mead and Carver 1980) and the difficulty of exploring new regions of the essentially infinite design space. For example, Altschuler (1998) explicitly recommends searching previous inventions for universal analogies and possible applications to new contexts, and Goldenberg et al. (1999) propose and empirically validate a method to identify and link profitable dependencies between previously used components. Unfortunately for inventors, however, these benefits of familiarity do not last forever. This results from the technological and social-psychological exhaustion of potential refinements, given a particular combination. Both influences run counter to the positive effects of local search.

Technological exhaustion occurs because most of the possible relationships between a set of components have already been tried. As Kim and Kogut (1996) argue, "The repeated application of a particular set of technologies or organizing principles eventually exhausts the set of potential combinations." The argument generalizes Sahal's (1985) demonstration of the decreasing returns to physical scaling. For example, if semiconductor inventors restricted their usage to their original materials of aluminum and bipolar transistors, progress in the field would have halted long ago.

However, because inventors began using new materials such as copper interconnect and new combinations such as metal oxide semiconductors, semiconductor chips have continued to shrink and the trajectory has repeatedly avoided exhaustion.

Exhaustion in technological potential also stems from social and psychological sources and can ultimately be traced to inventors' frames and "imaginary life cycles" (Henderson 1995). If a community of inventors accepts or reifies a particular combination as mature, then they are less likely try new combinations or broaden their component set. Henderson describes three such attitudes in her study of the premature obituaries for optical lithography in 1977: Inventors failed to consider the possibility of new architectures, improvement in their components, or changes in users' application of the technology. These technological and social-psychological effects run counter to the improvement that results from familiarity and learning.

Hypothesis 5. *Cumulative use of a combination will decrease an invention's usefulness.*

Methods

Thinking of invention as a recombinant search process implies that an invention's usefulness and uncertainty can be predicted from previous usage of its components and particular combination of components. I test these ideas with patent data and estimation of a statistical model. The models predict future prior art citations to a given patent based on previous usage of its assigned subclasses and particular combination of subclasses. The hypotheses place additional constraints on the empirical work, however. If engineers can recombine any technology with any other, the data should include recombination across technological communities. If invention is a process of search, learning, and exhaustion, the data should include the history of component and combination use. Finally, if the uncertainty of invention decreases with learning and the use of familiar components and combinations, the statistical model should estimate the impact of such learning and familiarity on the variability of inventive outcomes.

Data

The unit of analysis is an U.S. patent granted during May or June of 1990 ($n = 17{,}264$).[2] The data came from the MicroPatent (1996) product. While patent data are admittedly imperfect records of technology (Levin et al. 1987) and conditional on a successful patent application, these data still represent a large portion of failed and successful invention. Independent measures (recombinant history) came from data before July of 1990 and dependent measures (citations) came from data after July 1, 1990. Citations are used as controls and a dependent variable only—they are not used to trace the path of combination reuse, learning, or the diffusion of knowledge. The design is cross-sectional and does not consider multiple time periods. While the sample is large enough to provide very significant results for all the hypothesis tests, future work should consider other time periods to minimize the chances of idiosyncratic sampling.

I observe recombination across all technological communities by looking simultaneously at all patents granted during the two-month time period, thus satisfying the first modeling constraint. If the data did not consider all communities simultaneously, I would need to define the boundaries between particular technological communities. Because such boundaries are transient, permeable, overlapping, and nested, it would be empirically intractable to determine them for even a fraction of the made world. Finally, two hundred years of patent data satisfy the need to observe recombinant history of component and combination use.

Measures

Table 1 lists and describes all variables. For the dependent variables of mean and variance, I measure an invention's usefulness as the number of prior art *citations* that it receives from subsequent patents. To be granted a patent, the applicant must establish the novelty of her invention relative to all previous inventions. She establishes her claim to novelty by identifying (almost always after the fact of invention) and citing similar "prior art." The patent examiner then reviews and usually supplements these citations (Carr 1995).

[2] I chose the months of May and June at random.

Table 1 Description of Variables

Variable	Type	Description	Measure
Citations	dependent	invention's usefulness or importance	prior art citations by future patents to focal patent
Mean technology	control	expected citations to technically similar patents	weighted fixed effects by focal patent's sub-class membership in class
Variance technology	control	expected variance in citations to technically similar patents	weighted fixed effects by focal patent's sub-class membership in class
Number of prior art citations	control	patents that cite more heavily should be more heavily cited	number of prior art citations by focal patent
Single-class dummy	control	process of recombination is finer grained than can measure	equals 1 if focal patent assigned to only one subclass
Number of subclasses	control	number of invention's components	number of focal patent's subclasses
Newest subclass	control	artifact of patent classification system	equals the minimum number of previous uses amongst the focal patent's sub classes
Number of classes	control	breadth of patent classifications	number of focal patent's classes
Component familiarity	independent	inventor's familiarity with components of the invention	recent and frequent usage of focal patent's subclasses across all U.S. patents
Combination familiarity	independent	inventor's familiarity with particular combination of components	recent and frequent usage of particular subclass combination across all U.S. patents
Cumulative combination usage	independent	cumulative number of inventive trials with exactly same combination	cumulative number of U.S. patents with exactly same subclass combination

Bibliometric studies have repeatedly demonstrated that future prior citations to a patent correlate with its technological importance and value (Albert et al. 1991, Hall et al. 2000). To make maximal use of the data, I measure prior art citations to a focal patent for 6 years and 5 months after its granting. This period should capture the bulk of citations to a patent as these citations typically reach a plateau after about three years from the grant date (Jaffe et al. 1993). While there appears to be strong correlation between the rates of early and later citations to a patent, researchers are actively pursuing the topic (Hall et al. 2000).

To measure the independent variables, I proxy components with patent subclasses. The patent office categorizes all patentable technologies into some 400 "class references". Each class is also subdivided into very fine divisions of technology or approximately 100,000 "subclasses" (Trajtenberg et al. 1997). The patent office typically assigns each patent into multiple subclasses within and across major classes. The patent office also establishes and updates new classes and subclasses each year, as technology changes (Carr 1995, p. 128). This retrospective updating enables historical consistency in the measurement of components across time.

I do not propose that inventors recombine patent subclasses directly, only that subclasses can be used to observe indirectly the process of recombinant search and learning. I will illustrate the correspondence between components and patent subclasses with patent 5,136,185, coauthored with John Walther. It was (as of Dec. 1996) classified in four subclasses: 326/16 (with test facilitating feature), 326/31 (signal sensitivity or transmission integrity), 326/56 (tristate (i.e., high impedance as third state)), and 326/82 (current driving (e.g., fan in/out, off chip driving, etc.)). Each of these subclasses corresponded to a well-understood digital hardware component at that time. Each of them can be found in (and, for this patent, were drawn specifically from) contemporary textbooks of digital design such as McCluskey (1986). A test facilitate feature (p. 426) has become a necessity with large computer chips due to their millions of gates. To test the innards of a chip, the chip must be operated in a "test mode." A transmission gate (p. 118) simply passes on a signal when enabled.

A tristate driver (back cover diagram of digital components and pp. 105, 119, 142) is a gate that can drive a bus or be turned off to present high impedance to the bus. Finally, "fan in" and "fan out" (p. 104) refers to the number of components that can drive or be driven by a particular gate (high fan out basically means a big gate). While we were not aware of the subclass definitions at the time of the invention, we were very familiar with these basic building blocks of our technological community, namely digital hardware engineers. I propose that subclasses serve as proxies for these building blocks.

With the subclassifications and date of application of each patent (interpolated prior to 1975), I develop the three independent variables for each focal patent: component familiarity, combination familiarity, and cumulative combination usage. *Component familiarity* proxies inventors' familiarity with components, based on the average degree to which the components have been recently and frequently used. The assumption is that inventors will be more familiar with components that have been recently and frequently used. I first calculate an individual measure for each separate subclass of the focal patent. To the extent that a particular subclass has been recently and frequently used, its individual measure will be higher. For each focal patent, I look backward to 1790 and consider each of its individual subclasses in turn. Whenever a particular subclass has been used in any previous invention, I multiply the indicator of occurrence by an exponentially decaying component (1). This exponential component represents the loss and forgetting of knowledge. For example, it is more likely that an inventor will have learned from previous use of a subclass, if that sub-class was used three years prior, instead of thirty. These occurrences are then summed and averaged (2).

Individual component familiarity of patent i's subclass $j \equiv$

$$I_{ij} = \sum_{\substack{\text{all patents } k \text{ granted} \\ \text{before patent } i}} 1\{\text{ patent } k \text{ uses subclass } j\} \times e^{-\left(\frac{\text{application date of patent } i - \text{ application date of patent } k}{\text{time constant of knowledge loss}}\right)}. \tag{1}$$

Average component familiarity of patent

$$i\text{'s subclasses} \equiv F_i = \frac{\sum_{\substack{\text{all subclasses} \\ j \text{ of patent } i}} I_{ij}}{\sum_{\substack{\text{all subclasses} \\ j \text{ of patent } i}} 1}. \tag{2}$$

I set the time constant of knowledge loss at five years. This constant in the denominator of the exponential implies that approximately one-third of the knowledge remains after five years, or a yearly loss rate of 18%. Argote, Epple, and Darr (Argote et al. 1990, Epple et al. 1991, Darr et al. 1995) have estimated a much higher geometric loss parameter for manufacturing and service organizations, between 40% to 97% per year. It is unlikely that design technology would experience such a high rate of loss, however. Design knowledge is far less contextual and more easily articulated than manufacturing or service experience. Design knowledge is more likely to have been recorded in trade journals, firm documentation and, of course, patents. It may have been actually realized in prototypes and products and is far more likely to have required substantial personal effort and investment on the part of its designers. All of these influences would argue for a slower loss rate than manufacturing and service organizations, but the precise estimation certainly constitutes a valid research question.

I calculate *combination familiarity* and *cumulative combination usage* similarly. The cumulative usage measures how many times since 1790 a particular combination of subclasses has been used. Combination familiarity proxies inventors' familiarity with the combination, based on the degree to which the combination has been recently and frequently used. For each focal patent, I consider its particular combination of subclasses. I then look back in time at all other patents that used an identical combination of subclasses. For each previous patent that used an identical combination of subclasses, I increment the cumulative combination usage (3). For the combination familiarity variable, I multiply each indicator count by the exponential component, to reflect the loss and forgetting of experiential knowledge regarding the combination (4). All the measures remain indirect because they assume that learning occurs with use and that knowledge diffuses throughout technological communities and the made world. Given that all the independent variables are highly skewed, I took the square root of their original value. This minimized the effects of outliers and enabled more parsimonious modeling.

Cumulative combination usage of patent $i \equiv C_i =$

$$\sum_{\substack{\text{all patents } k \text{ granted} \\ \text{before patent } i}} 1\{\text{patent } k \text{ uses identical combination of subclasses as patent } i\} \tag{3}$$

Combination familiarity for patent $i \equiv R_i =$

$$\sum_{\substack{\text{all patents } k \text{ granted} \\ \text{before patent } i}} 1\{\text{patent } k \text{ uses identical combination of subclasses as patent } i\} \times e^{\left(\frac{\text{application date of patent } i \,-\, \text{application date of patent } k}{\text{time constant of knowledge loss}}\right)} \tag{4}$$

Given the theoretical interest in recombinant search, the models should ideally control for other differences between inventions, for example, the differences in citation patterns across different technology classes. I calculate *technology mean and variance* by measuring citations to a technology class prior to the dependent measures period. I first calculate the expected number of citations to each class in the five and a half years prior to July of 1990 (5). Equation (6) illustrates similar calculations for the variance of each patent class. Using these numbers for each class, I control for the expected mean and variance of citations to each focal patent based on a weighted average of its assigned technology class references. I calculate this by multiplying the proportion of a patent's subclass assignments within a particular class by the average cites/class of that class. As an additional control of differences in citation activity, I include the *number of prior art citations* that a focal patent makes to previous patents. Although none of these measures differences across industry directly, in aggregate they provide some control of nonrecombinant heterogeneity, to the extent that industries have similar technologies and citation patterns.

Average citations to a patent classified within class $i \equiv \mu_i =$

$$\frac{\displaystyle\sum_{\substack{\text{all patents } i \text{ granted} \\ \textit{from } 1985-1990.5}} 1\{\text{patent } j \text{ cites a patent within class } i\} * (\text{proportion of cited patent within class } i)}{\displaystyle\sum_{\substack{\text{all patents } k \text{ granted} \\ \text{from } 1985-1990}} 1\{\text{patent } k \text{ classified within class } i\} * (\text{proportion of patent } k \text{ within class } i)} \tag{5}$$

Variance in citations to patents classified within class $i \equiv \sigma_i^2 =$

$$\frac{\displaystyle\sum_{\substack{\text{all patents } k \text{ granted} \\ \textit{from } 1985-1990.5}} 1\{\text{patent } k \text{ classified within class } i\} * (\text{proportion of patent } k \text{ within class } i) * (\text{cites}_k - \mu_i)^2}{\displaystyle\sum_{\substack{\text{all patents } k \text{ granted} \\ \text{from } 1985-1990.5}} 1\{\text{patent } k \text{ classified within class } i\} * (\text{proportion of patent } k \text{ within class } i)} \tag{6}$$

Given the theoretical assumptions of bounded rationality and component search, the models should control for the number of the invention's components. These can be proxied by the focal patent's *number of subclasses*. Because it becomes increasingly difficult for an exact match to occur as the number of components increases, I include a second-order term for this as well.[3] Because the recombinant history of patents with only one subclass remains unobservable, I also include a dummy variable for *single* subclass patents (they comprise 8.1% of the data).

Finally, the models should control for artifacts of the patenting system. The patent office periodically updates technology classifications and creates new subclasses (these analyses use the classifications in effect as of November 1996). Because such reclassification essentially recognizes newly successful technologies, the first few patents of a retrospectively identified technology stream are likely to be highly seminal. Hence, patents that include a newly designated subclass will probably be highly cited. I include a *newest*

[3] I would like to thank an anonymous reviewer for this suggestion. In addition to adding the squared term, I checked the sensitivity of the results by estimating the model separately on subsets of the data, for patents classified in two and three subclasses. The results agreed with those presented below.

Table 2 Descriptive Statistics for All U.S. Patents, May/June 1990 ($n = 17{,}264$)

Variable	mean	stnd.dev.	minimum	maximum
citations	3.80	4.88	0.00	82.00
mean tech control	1.19	0.41	0.33	3.03
variance tech control	4.65	2.65	0.36	20.55
number of prior art citations	7.63	6.99	0.00	110.00
single subclass dummy	0.08	0.27	0.00	1.00
number of subclasses	4.21	3.31	1.00	130.00
newest subclass	107.07	103.97	0.00	1874.00
number of classes	1.78	0.95	1.00	12.00
component familiarity	0.60	0.27	0.00	2.28
combination familiarity	0.28	0.75	0.00	11.25
cumulative comb usage	0.62	1.60	0.00	26.76

subclass control variable that equals the minimum of the least used subclass (for example, if a patent had three subclasses and each subclass had been previously used in 56, 2, and 43 patents, the variable would equal 2). This variable is very similar to the dependent variable and hence probably decreases the effects of the independent variables. I also include the *number of classes* to which a patent belongs to reduce the effect of additional citations to a patent, similar to that which occurs when a scientific paper straddles fields.

Tables 1, 2, and 3 list the variables, descriptive statistics, and a correlation matrix. Combination familiarity and cumulative usage demonstrate high correlation in Table 3. This is not a problem except for inflated standard errors, given the desirable large sample properties of maximum likelihood estimators (Greene 1993, p. 133).

Negative Binomial Count Models

The dependent variable of citation counts takes on only whole number values (that is, 0, 1, 2, etc.). The use of a linear regression model on such data can yield inefficient, inconsistent, and biased coefficient estimates (Long 1997). Explicit count models can avoid these problems. Researchers often use Poisson models to analyze count data, but Poisson models assume that the mean and variance of the observed distribution are equal. These data, like most count data, exhibit over-dispersion—the variance is greater than the mean. Negative binomial regressions explicitly accommodate this over-dispersion, however, by enabling the variance to be greater than the mean. Recent developments (King 1989, Jorgensen 1997) also support independent estimation of effects on the mean and the variance of the predicted mean, as functions of potentially differing sets of variables.

To begin with, consider a model that estimates the mean number of citations that a patent should receive, given its independent variables. The observed citations to the patent will not correspond exactly to the prediction, however, and will be distributed with some variance around the expected mean. I operationalize the uncertainty of invention by estimating the effects of substantive variables on this variance. For example, to the degree that some independent

Table 3 Correlation Matrix for All U.S. Patents, May/June 1990 ($n = 17, 264$)

	cites	mean	variance	prior	single	numsub	newest	class	comp	combin
mean tech control	0.31									
variance tech control	0.30	0.94								
number prior art cites	0.12	0.02	0.00							
single subclass	−0.06	0.00	0.00	−0.07						
number subclasses	0.11	0.01	0.07	0.08	−0.29					
newest subclass	−0.08	−0.08	−0.07	−0.03	0.24	−0.23				
number of classes	0.08	−0.02	0.01	0.06	−0.24	0.51	0.07			
component familiarity	0.16	0.30	0.35	−0.01	−0.09	0.27	0.01	0.14		
combination familiarity	0.01	0.06	0.05	−0.04	0.47	−0.27	−0.05	−0.23	0.17	
cumulative comb usage	−0.05	−0.03	−0.03	−0.05	0.65	−0.30	−0.05	−0.26	0.06	0.85

variable decreases inventive uncertainty, its effect on the variance of this error distribution should be negative.

Most derivations of the negative binomial start from a basic Poisson model (7). The basic Poisson model estimates the probability of an observed count, conditional on an expected mean μ_i. To avoid negative (i.e., undefined) expected values for the mean μ_i, Poisson models typically parameterize explanatory variables as an exponential function (8). The method of maximum likelihood is then applied to the joint frequency formed from the product of the marginal frequencies of (7), to determine the coefficient values that are most likely to result in the observed counts.

$$\Pr(y_i|\mathbf{x}_i) = \frac{e^{-\mu_i}\mu_i^{y_i}}{y_i!}. \tag{7}$$

$$E(y_i|\mathbf{x}_i) = \mu_i = e^{(x_i\beta)}. \tag{8}$$

The negative binomial model replaces the Poisson mean μ_i with the random variable $\tilde{\mu}_i$ (9). This replacement enables the inclusion of an error term $\delta_i = e^{\varepsilon_i}$—and allows the predicted mean to vary according to the distribution of the error term. Substitution of $\tilde{\mu}_i$ for μ_i in (7) results in (10).

$$\tilde{\mu}_i = e^{(x_i\beta+\varepsilon_i)} = \mu_i\delta_i. \tag{9}$$

$$\Pr(y_i|\mathbf{x}_i, \delta_i) = \frac{e^{-\mu_i\delta_i}(\mu_i\delta_i)^{y_i}}{y_i!}. \tag{10}$$

When $\tilde{\mu}_i$ replaces μ_i in (10), the probability of the observed count becomes conditional on the error distribution. This conditioning can be removed, however, by specifying the error distribution and integrating with its probability density function to obtain the marginal density. Most formulations specify a gamma distribution for δ_i with parameter ν_i and probability density function $g(\delta_i)$ as in (11) (Hausman et al. 1984, Cameron and Trivedi 1986, King 1989, Long 1997). Although the error term can take other distributions, this parameterization is flexible, computationally tractable, and can be derived from a variety of assumptions. While other versions of the gamma distribution take two parameters, this derivation (from Long 1997, p. 232) sets both to ν_i, which forces the mean of δ_i equal to one and the variance of δ_i equal to $1/\nu_i$. Integrating (10) with the density function of (11) gives the probability of the negative binomial of (12).

$$g(\delta_i) = \frac{\nu_i^{\nu_i}}{\Gamma(\nu_i)}\delta_i^{\nu_i-1}e^{(-\delta_i\nu_i)} \quad \text{for } \nu_i > 0,$$
$$\text{and } \Gamma(\nu) = \int_0^\infty t^{\nu-1}e^{-t}dt. \tag{11}$$

$$\Pr(y_i|\mathbf{x}_i) = \frac{\Gamma(y_i+\nu_i)}{y_i!\Gamma(\nu_i)}\left(\frac{\nu_i}{\nu_i+\mu_i}\right)^{\nu_i}\left(\frac{\mu_i}{\nu_i+\mu_i}\right)^{y_i}. \tag{12}$$

While the first term of (9) fully specifies the mean of the negative binomial, various parameterizations of ν_i remain possible. Cameron and Trivedi (1986) propose the negative binomial II parameterization (or Negbin II model) when the variance/mean ratio of the observed data is linear in the mean. By contrast, the Negbin I holds the variance/mean ratio constant. The Negbin II specification is also much more robust to distributional misspecification than other parameterizations (Cameron and Trivedi 1986). I verified the applicability of the Negbin II by regressing predicted counts on the quantity (residuals^2/predicted). The coefficient and intercept were positive thus supporting the Negbin II parameterization (Cameron and Trivedi 1986). Also consistent with Cameron and Trivedi's argument (1986), Negbin II models demonstrated much more significant log likelihoods than Negbin I models.

Equation (13) specifies the Negbin II parameterization of the conditional variance. α_i is the inverse of ν_i and is parameterized as an exponential function, similar to the mean specification but with potentially different variables. Since $\text{var}(\delta_i) = 1/\nu_i = \alpha_i$, greater variance of the error term for a given estimated mean will result in an increase of α_i. I operationalize uncertainty by estimating the effects of the causal variables on the dispersion parameter α. Variables that decrease α will decrease the variability and hence the uncertainty of inventive outcomes.

$$\text{Var}(y_i|\mathbf{x}) = \mu\left(1+\frac{\mu_i}{\nu_i}\right) = \mu_i + \alpha\mu_i^2. \tag{13}$$

STATA estimates (12) by the method of maximum likelihood. This technique estimates both the mean μ_i and variance $1/\nu_i$ from (12), and hence produces only a single log likelihood.

Results

Table 4 presents estimates for the negative binomial dispersion models of citation counts. Model 1 estimates a baseline model of controls only and Model 2 adds substantive variables to the baseline model. Piecewise models indicated a nonmonotonic effect of component familiarity on the dispersion. To accommodate this, Models 3 and 4 include both first- and second-order terms for component familiarity. The substantive variables greatly increase the explanatory power of Models 2 and 3, as measured by twice the difference in log likelihoods and compared to a chi-squared distribution with degrees of freedom equal to the added number of variables.[4] To test robustness, Model 4 estimates substantive variables only. Coefficient magnitudes and significances vary but the substantive results remain unchanged and significant. To further check for robustness, I split the data set in half, randomly, and by month. Signs remained unchanged although not always significant. The parameter estimates are not standardized and should be interpreted as the predicted multiplier effect on the mean citation count and dispersion parameter (they should be exponentiated as in (8) and (13)).

Hypothesis one proposes that usage of familiar components will enable inventors to apply learning from previous efforts. Such learning improves inventors' abilities to select the best components and recombine them more successfully. Component familiarity has a positive and highly significant coefficient estimate. A patent receives 147% more citations at its maximum value of 2.28 as opposed to its lowest value of 0.[5] To interpret less extreme changes in the variable, a one standard deviation increase in component familiarity results in an 11.3% increase in expected citations.[6]

While the results support Hypothesis 1, they do not support Hypothesis 2, that component learning decreases the uncertainty of invention. Given that the first-order term by itself became insignificant without controls (model not shown), I estimated piecewise models to check for a nonlinear effect. These models indicated a nonmonotonic effect, with the maximum negative effect at the 60th percentile of the data. Compared to a baseline model of no effect of component familiarity on the variance, component familiarity decreases the variance by 9.6% at the 60th percentile.[7] Past that point, the effect becomes less negative and eventually becomes positive at the 96th percentile. It appears that increasing familiarity with components has an initially negative effect on the variability of invention. Eventually this changes, however, such that the use of very familiar components has a positive effect on variability.

In parallel to Hypothesis 1, Hypothesis 3 argues that reuse of familiar combinations lets inventors apply learning from previous efforts. Such learning enables refinement and improvement of previous inventions, such that future inventions are more useful. Combination familiarity demonstrates a positive, strong, and highly significant effect on the mean. A patent receives 332% more citations at combination familiarity's highest value of 11.25 as opposed to its lowest value of 0. A one standard deviation increase of combination familiarity results in a 10.3% increase in expected citations.

In contrast to the contrary results of component familiarity on the dispersion, combination familiarity demonstrates predicted and highly significant

[4] R-squared measures are inappropriate for maximum likelihood estimates and suffer from various problems (Cameron and Windmeijer 1996). Cameron and Windmeijer (1996) propose an imperfect (by their argument) measure for a negative binomial model, but it only considers the explanatory power of variables on the mean. Cameron (in personal communication) indicated that an acceptable R-squared measure for measuring the explanatory power of variables on the dispersion parameter has yet to be proposed.

[5] At its maximum value of 2.28, and from the coefficient estimate of 0.3958 in Model 3, the effect of component familiarity is $e^{(0.3958*2.28)} = 2.47$.

[6] $e^{(0.3958*0.27)} - 1 = 0.1128$.

[7] The maximum negative effect of component familiarity on the dispersion (from Model 3) occurs at 0.5992. At the mean citation count of 3.80 the variance is therefore

$$\mathrm{Var}(y_i|\mathbf{x}) = \mu_i + \alpha\mu_i^2 = 3.80e^{0.3958*0.5992} + e^{(-0.4139*0.5992+0.3454*0.5992^2)}(3.80e^{0.3958*0.5992})^2 = 25.32$$

At the mean citation count of 3.80 and no effect of combination familiarity on the dispersion at 0.5992, the variance is $\mathrm{Var}(y_i|\mathbf{x}) = \mu_i + \alpha\mu_i^2 = 3.80e^{0.3958*0.5992} + e^0(3.80e^{0.3958*0.5992})^2 = 28.02$. Hence the maximum negative effect is $(28.02 - 25.32)/28.02 = 9.6\%$ less.

Table 4 **Negative Binomial Models of Citation Counts (All U.S. Patents, May/June of 1990)**

Variable/Models	Model 1	Model 2	Model 3	Model 4
Effects on the mean				
Mean technology control	0.8830***	0.7790***	0.7800***	
	(0.0198)	(0.0215)	(0.0216)	
Number of prior art citations	0.0181***	0.0184***	0.0185***	
	(0.0012)	(0.0012)	(0.0012)	
Single-class dummy control	−0.1742***	−0.1568***	−0.1605***	
	(0.0349)	(0.0437)	(0.0438)	
Number of subclasses control	0.0271***	0.0183***	0.0185***	
	(0.0044)	(0.0049)	(0.0049)	
Number of subclasses squared	−0.0002+	−0.0001	−0.0001	
	(0.0001)	(0.0001)	(0.0001)	
Newest subclass control	−0.0003***	−0.0009***	−0.0009***	
	(0.0001)	(0.0001)	(0.0001)	
Number of classes control	0.0461***	0.0425***	0.0413***	
	(0.0097)	(0.0098)	(0.0098)	
Component familiarity		**0.3807***	**0.3958***	**0.7195***
		(**0.0430**)	(**0.0441**)	(**0.0353**)
Combination familiarity		**0.1329***	**0.1301***	**0.2301***
		(**0.0223**)	(**0.0225**)	(**0.0218**)
Cumulative combination usage		−**0.0384**	−**0.0363**	−**0.1523***
		(**0.0131**)	(**0.0132**)	(**0.0107**)
Constant	−0.0899*	−0.1090**	−0.1134**	0.9115***
	(0.0363)	(0.0369)	(0.0372)	(0.0230)
Effects on dispersion parameter				
Variance technology control	−0.0154**	−0.0187**	−0.0177**	
	(0.0054)	(0.0060)	(0.0060)	
Activity control	−0.0011	−0.0007	−0.0007	
	(0.0020)	(0.0020)	(0.0020)	
Single-class dummy control	0.1001	0.2426***	0.2111***	
	(0.0656)	(0.0725)	(0.0731)	
Number of subclasses control	−0.0170*	−0.0164*	−0.0178*	
	(0.0075)	(0.0081)	(0.0081)	
Number of subclasses squared	0.0003+	0.0003+	0.0003+	
	(0.0002)	(0.0002)	(0.0002)	
Newest subclass control	−0.0003+	−0.0003	−0.0002	
	(0.0002)	(0.0002)	(0.0002)	
Number of classes control	−0.0390*	−0.0505**	−0.0420*	
	(0.0189)	(0.0191)	(0.0193)	
Component familiarity		**0.1683***	−**0.4139***	−**0.5331**
		(**0.0749**)	(**0.2048**)	(**0.1798**)
Component familiarityˆ2			**0.3454**	**0.3121**
			(**0.1135**)	(**0.1059**)
Combination familiarity		−**0.1033***	−**0.1011***	−**0.0506***
		(**0.0249**)	(**0.0248**)	(**0.0200**)
Constant	0.0545	−0.0177	0.1609*	0.1644*
	(0.0560)	(0.0578)	(0.0819)	(0.0683)
Log-likelihood	−41095.26	−41001.92	−40997.26	−42022.00

($n = 17{,}264$ all models, standard errors in parentheses, $+p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$)

results on the dispersion. The negative effect supports Hypothesis 4, that inventive uncertainty decreases with the refinement of previously used combinations. At the mean citation count of 3.80 and the mean value of combination familiarity at 0.28, the variance is 2.2% less than it would have been without the effect of combination familiarity on the dispersion. At the maximum value of combination familiarity, the difference grows to 64%.

Finally, Hypothesis 5 predicts that cumulative use of a particular combination will eventually exhaust recombinant potential. The results support this hypothesis, as indicated by the negative and significant estimate on cumulative combination usage. When cumulative combination count reaches its maximum value of 26.76, a patent receives only 37.9% of the citations it would have received at the variable's lowest value of 0. A one standard deviation increase of cumulative combination use results in a 5.6% decrease in expected citations.

Discussion

These results should be viewed cautiously for a variety of reasons. The typical reservations regarding the use of patent data certainly apply, most notably that patenting practices and effectiveness vary across industries (Levin et al. 1987). Furthermore, even though these data cover all patented technologies across a two-month time period, much inventive activity remains unpatented and, hence, unobserved. These issues cause at least three problems. First, the dataset misses unimportant inventions that failed to merit a patent. Use of citation data mitigates this problem, however, because it includes the bulk of relatively useless inventions that receive no or few citations. Second, the dataset may miss breakthroughs that firms chose not to patent, presumably for strategic reasons. However, unless there is systematic bias in those firms' use of particular components or combinations, these results should remain valid. Third, technological communities vary in their propensity to patent. Again, while these models controlled for much variance across technologies, they did not introduce explicit industrial controls. But given the focus on recombinant search across technological communities, it would be difficult to compile similarly vast observations across so many communities. The dataset therefore trades detail and depth in exchange for breadth and possibility of observation.

In addition to differences in the dependent variable of citations across communities, the accuracy of the independent variables also varies across communities. Subclasses and combinations of subclasses represent only a proxy for inventors' components and architectures. While the subclasses of digital hardware patents correspond very closely to my engineering experience, other subclasses may not. For example, financial patents tend to be classified in fewer subclasses and, hence, may not reflect a process of recombinant search. They may also be better characterized as knowledge or algorithms and not technology. Such patents, however, became popular only after the observation period.

In addition to concerns about data and variables, these results also remain open to alternate interpretations. Most importantly, these models cannot definitively separate learning and familiarity, technological exhaustion, and life-cycle effects. For example: an exogenous variable (such as the inventive myopia described by Henderson (1995)) could drive the use of particular components and combinations, such that increased citations might simply reflect the popular usage of those components and combinations. As a result, the positive signs in the mean for component and combination familiarity and negative sign for cumulative use might reflect only the normal progression of technological life cycles, instead of learning and exhaustion. Even with the multiple controls for differences in citation patterns across technologies, types of combinations, and retrospective identification of seminal subclasses, the models cannot convincingly reject the alternative argument that the results merely reflect the normal life-cycle progression.

These reservations do not apply, however, to the uncertainty hypotheses. Indeed, all of the explanatory variables in the mean could simply be interpreted as control variables for the dispersion estimates. The combination result supports the intuitive argument that uncertainty decreases with refinement and highlights the importance of early architectural refinements as a source of destabilizing technologies for organizations and industries (Henderson and Clark

1990). In contrast, the nonmonotonic results for components does not support the simple linear relationship predicted in Hypothesis 2. It remains consistent, however, with the argument that technological breakthroughs derive from new combinations of well-used components (Usher 1954, Nelson and Winter 1982, Sahal 1985). For example, Utterback (1996, p. xxvii) argues that "radical innovations often are seen to be based on the synthesis of well-known technical information or components." These results also indicate that inventors take more inventive risk with extremely familiar components. These results support Utterback's argument and motivate further research. For example, are breakthroughs most likely when inventors combine very familiar components in new combinations? If this were the case, then breakthroughs would be most likely to emerge from social contexts that brought together inventors with deep experience in previously disparate fields. Such contexts would also be more likely to be the technological source of potential future product life cycles and trajectories.

Conclusion

This paper developed and tested an explanation for the sources of purely technological uncertainty. It argued that the source of technological novelty and uncertainty lies within the combination of new components and new configurations of previously combined components. Inventors' experimentation with new components and combinations leads to less success on average, but it also increases the variability that can lead to breakthroughs. Empirical results supported the arguments with the exception that the use of more familiar components has a nonmonotonic and eventually positive effect on the uncertainty of invention. In contrast to the nonmonotonic effect of component familiarity, the refinement of previously used combinations has a negative and monotonic effect on uncertainty.

Coupled with recent and complementary research regarding market influences (Adner and Levinthal 2000) and formal economic models (Klepper 1996), this work helps us understand the causal forces that underlie the widely observed regularities of the product life cycle. The data and methods presented here also provide a strong basis for further investigations into the sources of invention and technological uncertainty. Most importantly, negative binomial count and dispersion models enable researchers to analyze the first and second moments of patent citation data. Such models will enable us to move beyond basic counts in analysis of patent data and quantitatively analyze the outliers of the highly skewed distributions of inventive trials. Such tools can enable more formal analysis of breakthrough inventions, heretofore, "the domain of economic historians" (Scherer and Harhoff 2000). For example, the classic controversy about the sources of technological breakthroughs, whether they emerge from smaller, entrepreneurial, and "outside" firms (Schumpeter 1939, Marquis 1969, Klein 1977), or large, industrial incumbents (Schumpeter 1942), can be reconsidered with these data and methods. The approach presented here also has application beyond the study of technology and patents, mainly that we should think of varying variance as an opportunity instead of a nuisance.

Future work should also investigate the relationships between invention as a recombinant search process and other literatures such as integrality, coupling, and modularity (Ulrich 1995), complex systems (Kauffman 1993), modular operators (Baldwin and Clark 1999, Goldenberg et al. 1999), the evolution of modular design choices (Simon 1996), and the market implications of such issues (Christensen and Verlinden 2000). For example, interdependence between components should increase inventive uncertainty and modularity should decrease it. Kauffman's (1993) models of search over interdependent landscapes also imply a positive but nonmonotonic effect on the mean. Conceptualizing technological invention as search over interdependent landscapes implies a complexity catastrophe in technological evolution, as inventors face greater numbers of components and greater interdependence between them (Fleming and Sorenson 2001). Inventors are not blind search agents, however, and their search strategies will differ greatly and presumably be more effective than genetic recombination (Gavetti and Levinthal 2000, Rivkin 2000).

The work has strategic implications as well. Organizations that seek technological breakthroughs should experiment with new combinations, possibly with old components. They do so, however, at the risk of an increased number of failures. If invention is indeed a partially blind search process, such failures are to some extent unavoidable. Simple portfolio strategies are unfortunately not the complete answer, because of the extreme skew of inventive distributions. Although the overall variance of such distributions decreases with an increasing number of trials, firms—and indeed, as Scherer and Harhoff (2000) demonstrate—entire economies cannot anticipate completely stable returns. There might exist organizational mechanisms that encourage recombinant exploration, while limiting the downside of increased failures. For example, firms that can screen or test their nascent inventions more effectively will benefit more and should increase the variability of their inventive trials. The recent progress in drug design as a result of automated screening processes provides a salient example (Amato 1999). Finally, science should enlighten and shorten technological search over difficult landscapes. Science can either motivate recombinant search across particular technological landscapes, or aid in search across landscapes that inventors have discovered empirically.

Acknowledgments
The author would like to acknowledge the helpful comments of Gautam Ahuja, Steve Barley, Bill Barnett, Phil Bromiley, Laurie Calhoun, Mike Hannan and his seminar students, James March, Candace Fleming, Olav Sorenson, Toby Stuart, and three very helpful reviewers. He would also like to acknowledge the support of William Simpson and the Faculty Research Computing Center at the Harvard Business School, the Department of Industrial Engineering at Stanford University, and Corey Billington, Phil Johnson, and Ellen King and the library staff of Hewlett Packard Company. Errors and omissions remain my responsibility.

References
Abernathy, W., J. Utterback 1978. Patterns of industrial innovation. *Tech. Rev.* (June) 40–47.

Adner, R., D. Levinthal. 2000. Technology evolution and market heterogeneity: Implications for product and process innovation. Working paper, INSEAD, France.

Ahuja, G. 2000. Collaboration networks, structural holes, and innovation: A longitudinal study. *Admin. Sci. Quart.* **45** 425–455.

Albert, M. B., F. Narin, D. Avery, P. McAllister. 1991. Direct validation of citation counts as indicators of industrially important patents. *Res. Policy* **20** 251–259.

Allen, T. 1977. *Managing the Flow of Technology*. MIT Press, Cambridge, MA.

Altschuler, G. 1998. *The Innovation Algorithm*. Technical Innovation Center, Inc., Worcester, MA.

Amato, I. 1999. Industrializing the search for new drugs. *Fortune* (May 10)

Anderson, P., M. Tushman. 1990. Technological discontinuities and dominant designs: A cyclical model of technological change. *Admin. Sci. Quart.* **35** 604–633.

Argote, L., S. Beckman, D. Epple. 1990. The persistence and transfer of learning in industrial settings. *Management Sci.* **36**(2) 140–154.

Ayres, R. 1988. Barriers and breakthroughs: An “expanding frontiers” model of the technology-industry life-cycle. *Technovation*. **7** 87–115.

Baldwin, C., K. Clark. 1999. *Design Rules: The Power of Modularity*. MIT Press, Cambridge, MA.

Basalla, G. 1988. *The Evolution of Technology*. Cambridge University Press, Cambridge, MA.

Bijker, W. 1987. The social construction of bakelite: Toward a theory of invention. Bijker, Hughes, Pinch, eds. *The Social Construction of Technological Systems*. MIT Press, Cambridge, MA.

Cameron, A., P. Trivedi. 1986. Econometric models based on count data: Comparisons and applications of some estimators and tests. *J. Appl. Econometrics*. **1** 29–53.

———, F. Windmeijer. 1996. *R*-squared measures for count data regression models with applications to health-care utilization. *J. Bus. Econom. Statist.* **14**(2) 209–220.

Carr, F. 1995. *Patents Handbook*. McFarland, Jefferson, NC.

Christensen, C., M. Verlinden. 2000. Disruption, disintegration, and the dissipation of differentiability. Working paper 00-074, Harvard Business School, Boston, MA.

Clark, K. 1985. The interaction of design hierarchies and market concepts in technological evolution. *Res. Policy* **14** 235–251.

Cohen, W., D. Levinthal. 1990. Absorptive capacity: A new perspective on learning and innovation. *Admin. Sci. Quart.* **35** 128–152.

Darr, E., L. Argote, D. Epple. 1995. The acquisition, transfer, and depreciation of knowledge in service organizations: Productivity in franchises. *Management Sci.* **41**(11) 1750–1762.

Epple, D., L. Argote, R. Devadas. 1991. Organizational learning curves: A method for investigating intra-plant transfer of knowledge acquired through learning by doing. *Organ. Sci.* **2**(1) 58–70.

Fleming, L., O. Sorenson. 2001. Technology as a complex adaptive system: Evidence from patent data. *Res. Policy* Forthcoming.

Gavetti, G., D. Levinthal. 2000. Looking forward and backward: Cognitive and experiential search. *Admin. Sci. Quart.* **45** 113–137.

Gilfillan, S. 1935. *Inventing the Ship*. Follett Publishing Co., Chicago, IL.

Greene, W. 1993. *Econometric Analysis*. Prentice Hall Inc., Upper Saddle River, NJ.

Goldenberg, J., D. Mazursky, S. Solomon. 1999. Toward identifying the inventive templates of new products: A channeled ideation approach. *J. Marketing Res*., **36** 200–210.

Hall, B., A. Jaffe, M. Trajtenberg. 2000. Market value and patent citations: A first look. NBER Working paper 7741,

Hargadon, A., R. Sutton. 1997. Technology brokering and innovation in a product design firm. *Admin. Sci. Quart.* **42** 716–749.

Hausman, J., B. Hall, Z. Griliches. 1984. Econometric models for count data with an application to the patents-R&D relationship. *Econometrica* **52** 909–938.

Henderson, R. 1995. Of life cycles real and imaginary: The unexpectedly long old age of optical lithography. *Res. Policy* **24** 631–643.

———, K. Clark. 1990. Architectural innovation: The reconfiguration of existing product technologies and failure of established firms. *Admin. Sci. Quart.* **35** 9–30.

Jaffe, A., M. Trajtenberg, R. Henderson. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *Quart. J. Econom.* **108** 577–598.

Jorgensen, B. 1997. *The Theory of Dispersion Models*. Chapman and Hall, London, U.K.

Kauffman, S. 1993. *The Origins of Order*. Oxford University Press, New York.

Kim, D., B. Kogut. 1996. Technological platforms and diversification. *Organ. Sci*., **7** 283–301.

King, G. 1989. Event count models for international relations: Generalizations and applications. *Internat. Stud. Quart.* **33** 123–147.

Klein, B. 1977. *Dynamic Economics.* Harvard University Press, Cambridge, MA.

Klepper, S. 1996. Entry, exit, and innovation over the product lifecycle. *Amer. Econom. Rev.* **86**(3) 562–583.

———, 1997. Industry life cycles. *Indust. Corporate Change* **6**(1) 145–181.

Kogut, B., U. Zander. 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. *Organ. Sci.* **3** 383–397.

Kroeber, A. 1948. *Anthropology*. Harcourt Brace Jovanovich, Inc., New York.

Levin, R., A. Klevorick, R. Nelson, S. Winter. 1987. Appropriating the returns from industrial research and development: Comments and discussion. *Brookings Paper Econom. Activity* (3) 783–831.

Levinthal, D. 1998. The slow pace of rapid technoligical change: Gradualism and punctuation in technological change. *Indust. Corporate Change* **7**(2) 217–247.

Long, J. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Sage Publications, London, U.K.

March, J. 1991. Exploration and exploitation in organizational learning. *Organ. Sci.*, **2** 71–87.

———, H. Simon. 1958. *Organizations*. Blackwell Publishers, Cambridge, MA.

Marquis, D. 1969. The anatomy of successful innovations. *Innovation* **1** 35–48.

McCluskey, E. 1986. *Logic Design Principles with Emphasis on Testable Semi-Custom Circuits*. Prentice Hall, Englewood Cliffs, NJ.

Mead, L. Carver. 1980. *Introduction to VLSI Systems*. Addison-Wesley, Reading, MA.

MicroPatent. 1996. Product of micropatent USA, New Haven, CT.

Mokyr, J. 1990. *The Lever of Riches: Technological Creativity and Economic Progress*. Oxford University Press, New York.

Nelson, R., S. Winter. 1982. *An Evolutionary Theory of Economic Change*. Belknap Press, Cambridge, MA.

Rivkin, J. 2000. Imitation of complex strategies. *Management Sci.* **46** 824–844.

Romer, P. 1993. Ideas and things: The concept of production is being re-tooled. *The Economist* (Sept 11) 70–72.

Rosenberg, N. 1996. Uncertainty and technological change. R. Landau, R. Taylor, G. Wright, eds. *The Mosaic of Economic Growth*. Stanford University Press, Stanford, CA.

Ruttan, V. 1959. Usher and Schumpeter on invention, innovation, and technological change. *Quart. J. Econom.* (November)

Sahal, D. 1985. Technological guideposts and innovation avenues. *Res. Policy* **14** 61–82.

Scherer, F., D. Harhoff. 2000. Technology policy for a world of skew-distributed outcomes. *Res. Policy* **29** 559–566.

Schumpeter, J. 1939. *Business Cycles*. McGraw-Hill Book Company, Inc., New York.

———, 1942. *Capitalism, Socialism, and Democracy*. Harper and Row Publishers, New York.

Simon, H. 1996. *The Sciences of the Artificial*. MIT Press, Cambridge, MA.

Stuart, T., J. Podolny. 1996. Local search and the evolution of technological capabilities. *Strategic Management J.* (Summer)

Tratjenberg, M., R. Henderson, A. Jaffe. 1997. University versus corporate patents: A window on the basicness of invention. *Econom. Innovation New Tech.* **5** 19–50.

Tushman M., P. Anderson. 1986. Technological discontinuities and organizational environments; *Admin. Sci. Quart.* **31** 439–465.

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

Usher, A.P. 1954. *A History of Mechanical Invention*. Cambridge, MA.

Utterback, J. 1996. *Mastering the Dynamics of Innovation*. Harvard Business School Press, Boston, MA.

Vincenti, W. 1990. *What Engineers Know, and How They Know It*. Johns Hopkins University Press, Baltimore, MD.

Weitzman, M. 1996. Hybridizing growth theory. *Proc. Amer. Econom. Assoc.* (May) 207–212.

Accepted by Karl Ulrich; received October 27, 1998. This paper was with the author 8 months for 3 revisions.