Sequential Innovation, Patents, and Imitation

James Bessen[†] and Eric Maskin[*] [‡]

November 1999

Revised July 2002

[†] Research on Innovation and Visiting Scholar, MIT Sloan School of Management
jbessen@researchoninnovation.org

[*] Institute for Advanced Study and Princeton University maskin@ias.edu

[‡] We would like to thank Lee Branstetter, Daniel Chen, Iain Cockburn, Partha Dasgupta, Nancy Gallini, Alfonso Gambardella, Bronwyn Hall, Adam Jaffe, Lawrence Lessig, Jean Tirole, participants at many seminars and conferences, and especially Suzanne Scotchmer for helpful comments. We gratefully acknowledge research support from the NSF.

Abstract

How could industries such as software, semiconductors, and computers have been so innovative despite historically weak patent protection? We argue that if innovation is both "sequential" (each invention builds on its predecessor) and "complementary" (a diversity of innovators raises the chances of discovery), a firm's profit may actually be enhanced by competition, and a patent system may interfere with such competition and with innovation. A natural experiment in the software industry and the positive relationship between innovation and firm entry provide support for our model.

1. Introduction

The standard economic rationale for patents is to protect innovators from imitation and thereby give them the incentive to incur the cost of innovation. Conventional wisdom holds that, unless would-be competitors are constrained from imitating an innovation, the inventor may not reap enough profit to cover that cost. Thus, even if the social benefit of invention exceeds the cost, the potential innovator without patent protection may decide against innovating altogether[1].

Yet interestingly, some of the most innovative industries of the last forty years—software, computers, and semi-conductors—have historically had weak patent protection and have experienced rapid imitation of their products[2]. Defenders of patents may counter that, had stronger intellectual property rights been available, these industries would have been even more dynamic. But we will argue that theory and evidence suggest otherwise.

---

[1] This is not the only justification for patents. Indeed, we will emphasize a different, although related rationale in Section 2. But, together with the spillover benefit that derives from the patent system's disclosure requirements, it constitutes the traditional justification.

[2] Software was routinely excluded from patent protection in the U.S. until a series of court decisions in the mid-1980's. Semiconductor and computer patent enforcement was quite uneven until the organization of the Federal Circuit Court in 1982. Both areas contend with substantial problems of prior art [Aharonian (1992)], and some experts argue that up to 90% of semiconductor patents are not truly novel and therefore invalid [Taylor and Silbertson (1973)]. These problems make consistent enforcement difficult. Surveys of managers in semiconductors and computers consistently report that patents only weakly protect innovation. Levin et al., (1987) found that patents were rated weak at protecting the returns to innovation, far behind the protection gained from lead time and learning-curve advantages. Patents in electronics industries were estimate to increase initiation costs by only 7% [Mansfield, Schwartz, and Wagner (1981)] or 7-15% [Levin et al., (1987)]. Taylor and Silberston (1973) found that little R&D was undertaken to exploit patent rights. As one might expect, diffusion and imitation are rampant in these industries. Tilton (1971) estimated the time from initial discovery to commercial imitation in Japanese semiconductors to be just over one year in the 1960's.

In fact, the software industry in the United States was subjected to a revealing natural experiment in the 1980's. Through a sequence of court decisions, patent protection for computer programs was significantly strengthened. We will suggest that, far from unleashing a flurry of new innovative activity, these stronger rights ushered in a period in which R&D spending leveled off, if not declined, in the most patent-intensive industries and firms.[3]

We maintain, furthermore, that there was nothing paradoxical about this outcome. For industries like software or computers, there is good reason to believe that imitation *promotes* innovation and that strong patents (long-lived patents of broad scope) *inhibit* it. Society might well be served if such industries had only limited intellectual property protection. Moreover, many firms might genuinely *welcome* competition and the prospect of being imitated[4].

This is because these are industries in which innovation is both *sequential* and *complementary*. By "sequential," we mean that each successive invention builds on the preceding one, in the way that the Lotus 1-2-3 spreadsheet built on VisiCalc, and Microsoft's Excel built on Lotus. And by "complementary," we mean that each potential innovator takes a different research line and thereby enhances the overall probability that a particular goal is reached within a given time. Undoubtedly, the

---

[3] As Sakakibara and Branstetter (2001) show, a similar phenomenon occurred in Japan. Starting in the late 1980's, the Japanese patent system was significantly strengthened. However, Sakakibara and Branstetter argue that there was no concomitant increase in R&D or innovation.

[4] When IBM announced its first personal computer in 1981, Apple Computer, then the industry leader, responded with full-page newspaper ads headed, "Welcome, IBM. Seriously." Indeed, there is a high-tech cliché contending that competition "expands the market."

many different approaches taken to voice-recognition software hastened the availability of commercially viable packages.

Imitation of a discovery may be socially desirable in a world of sequential and complementary innovation because it helps the imitator to develop further inventions. And because the imitator may have valuable ideas not available to the original discoverer, the overall pace of innovation may thereby be enhanced.

From this standpoint, patents are undesirable because they block further innovation. Of course, patent defenders have a counterargument: if a patent threatens to interfere with valuable innovative activity, the patent holder is likely to have the incentive to license it (thereby allowing innovation to occur). After all, if the additional innovation is worthwhile, he could capture its value by an appropriately chosen licensing fee/royalty, thereby increasing his own profit (or so the argument goes). The problem with this argument is that the appropriate fee or royalty may well depend on information about the licensee's costs and potential revenue to which the patent holder is not privy. Thus, there is a significant chance that the fee will be set too high discouraging licensing and thus invention, which in turn hurts society[5].

[5] There is a large existing literature on patent licensing, and some of this work considers, as we do, the issue of licensing to one's own competitors, including Katz and Shapiro (1985, 1986), Gallini (1984), and Gallini and Winter (1985). In these papers, however, the social loss from failure to license tends to derive from higher costs (because of decreasing returns to scale in monopoly production) and higher consumer prices, rather than from reduced innovation.

But whether or not patent protection is available, a firm may well be better off if other firms imitate and compete against it. Although imitation reduces the firm's current profit, it raises the probability of further innovations, which improve the firm's future profit—either because the firm, in turn, imitates these later inventions (in the absence of patent protection) or because it can set a higher licensing fee (when patents are possible).

In short, when innovation is sequential and complementary, standard reasoning about patents and imitation may get turned on its head. Imitation becomes a *spur* to innovation, whereas strong patents become an *impediment*.

Sequential innovation has also been studied by Scotchmer (1991, 1996), Scotchmer and Green (1990), Green and Scotchmer (1996), Chang (1995) for the case of a single sequential innovation. Hunt (1995, O'Donoghue (1998), and O'Donoghue, Scotchmer and Thisse (1998) study a single technology with an infinite sequence of quality improvements. We differ from this literature primarily in our model of technological competition. In our analysis, the technologies that different firms discover through R&D at any given stage differ from one another.[6] That is, imitators do not necessarily produce direct "knock offs," but rather differentiated products. This sort of differentiation is widely observed and is, of course, the subject of its own literature. But here, different technological paths permit innovative

[6] This feature also figures prominently in the model in Dasgupta and Maskin (1987).

complementarities. Imitation then increases the “bio-diversity” of the technology, improving prospects for future innovation.

We proceed as follows. In Section 2 we review the static model that underlies the traditional justification for patents. We emphasize the point that, besides helping to ensure that innovating firms cover their costs, patents can also encourage inventive activity on the part of firms that would otherwise be more inclined to merely imitate. Then in Section 3 we turn to a dynamic model and delineate the circumstances in which patents inhibit innovation and firms are made better off when imitated. Finally, in Section 4 we discuss some evidence that this dynamic model applies to high tech industries. In an appendix, we present a historical case study—that of spreadsheets—which we believe fits our dynamic model quite well.

2. The Static Model

We consider an industry consisting of two (*ex ante* symmetric) firms[7]. Each firm can undertake R&D to discover and develop an innovation with expected (social) value v.[8] The cost of this R&D is likely to depend in part on the firm’s level of talent. In high tech industries, firms are highly diverse; e.g., small talented firms often develop products that are very costly for larger firms. We assume that the

---

[7] Limiting the model to two firms is a matter only of expositional convenience; all our results extend to three or more firms.

[8] There is no additional social value that accrues if both firms discover the innovation.

talent level and corresponding cost of R&D are private information: the cost is $\bar{c}$ with probability q and $\underline{c}$ with probability 1- q, where $\bar{c} > \underline{c} > 0$.[9]

We also assume that, if a firm is not copied, then it can capture the full social value of its innovation[10]. However, the other firm, unless prevented by patent protection, can develop an imitation (trade secrecy provides no protection), either by investing in R&D itself (at cost $\underline{c}$ or $\bar{c}$, depending on its type) or by costlessly producing a "knock off" imitation.[11] And if it does so, each firm earns a duopoly profit less than v but greater than zero. In the following analysis, the important ramifications of this last assumption are that: (a) imitation reduces the first innovator's profit, and (b) imitators earn positive profit[12] . Without loss of generality, we assume each firm obtains $\frac{1}{2}v$.[13]

---

[9] Private information about costs may also arise because firms have disparate experience with different technologies.

[10] This, of course, is a strong assumption. However, the incentive failures and monopoly inefficiencies that arise when it is not imposed are already well understood. The assumption is a simple way to abstract from these familiar distortions.

[11] In reality, even knock-off imitations typically involve substantial costs. However, invoking this assumption *strengthens* the case for patents. This will make the argument in Section 3 on the shortcomings of the patent system all the more compelling.

[12] As we will see, (b) implies that, particularly in our dynamic model, imitators may have the incentive to undertake R&D.

[13] Alternatively, we could suppose that some monopoly profit is dissipated through competition. This would change the formulas obtained below, but would not affect the qualitative conclusions as long as duopoly profits remained greater than zero. Our assumption that the firms make equal profits (i.e., that the innovator has no first mover advantage) also biases the argument below in favor of patents.

If a firm undertakes R&D, then it develops an innovation with probability p.[14] Assume that

$$pv - \bar{c} > 0, \tag{1}$$

i.e., the net social benefit of having one firm undertake R&D is positive, even if the firm's R&D costs are high. Suppose for now that

$$\frac{1}{2}pv - \bar{c} < 0. \tag{2}$$

The combination of (1) and (2) constitutes the classic incentive failure that the patent system is meant to address. The left-hand side of (2) represents a high-cost firm's expected profit from undertaking R&D in the absence of patent protection. Thus, when a firm's costs are high, there will be no R&D investment without patents despite the fact that (1) implies that such investment would be socially beneficial.[15] A patent proscribes imitation and therefore guarantees an innovator the full net social return on R&D expenditure, $pv - c$. Hence, (1) tells us that, because society gains from R&D investment, so will the innovating firm itself.

But even in a setting where (2) does *not* hold—so that R&D remains profitable despite imitation—patents may well serve a useful purpose. This is because they can encourage several firms to go after the same innovation. Typically, different firms have different ideas about how to achieve a

---

[14] Our framework in this section is static, but if it were viewed as the reduced form of a dynamic setting, then p could alternatively be interpreted as the discount factor corresponding to the time lag to innovation.

particular technological goal. Therefore, increasing the number of firms in pursuit of this goal raises the probability that *someone* will succeed; this is what we called "complementarity" in the introduction.

We model complementarity by assuming that if both firms undertake R&D, each firm's chance of success is statistically independent of the other's[16], i.e., the total probability of successful innovation is

(3) $$1-(1-p)^2 = 2p - p^2 .$$

Suppose that the marginal social benefit of having a second firm undertake R&D is also positive. From (3), this can be expressed algebraically (assuming both firms have high costs) as:

$$\begin{aligned}&\left[\left(2p - p^2\right)v - 2\bar{c}\right] - \left[pv - \bar{c}\right] \\ &\qquad = \left(p - p^2\right)v - \bar{c} > 0,\end{aligned}$$

and rewritten as

(4) $$v/\bar{c} > \frac{1}{p - p^2}.$$

In the absence of patent protection, a high-cost firm's expected profit if both firms undertake R&D is

---

[15] Incentive failure, of course, can also occur for low cost firms.

[16] More realistically, the techniques available to each firm might be correlated to some degree, leading to correlation between the two firm's chances of success. We assume no correlation only for convenience.

(5) $$\frac{1}{2}\left(2p - p^2\right)v - \bar{c}\,.$$

(To understand this, note that $2p - p^2$ is the probability that at least one firm successfully innovates. If innovation does occur, each firm enjoys a payoff of only $\frac{1}{2}v$ because of imitation.)

If instead this firm refrains from R&D, its expected profit from imitating the other firm is

(6) $$\frac{1}{2}pv\,.$$

But even if (2) is violated, so that a single firm is willing to innovate despite the prospect of being imitated—and despite the fact that (4) holds—(5) and (6) imply that the second high-cost firm will *refrain* from R&D unless

$$\left[\frac{1}{2}\left(2p - p^2\right)v - c\right] - \frac{1}{2}pv = \frac{1}{2}\left(p - p^2\right)v - \bar{c} > 0,$$

which can be rewritten as

(7) $$v/\bar{c} > \frac{2}{p - p^2}.$$

With the possibility of patent protection, by contrast, each of two innovating firms with high costs expects a profit of (5) (here we assume that even if both firms make the discovery, only one obtains the patent[17]; hence, from symmetry, each firm has a probability $\frac{1}{2}\left(2p - p^2\right)$ of obtaining a patent). Thus, both firms will undertake R&D provided that

$$\frac{1}{2}\left(2p - p^2\right)v - \bar{c} > 0,$$

which can be rewritten as

$$v/\bar{c} > \frac{1}{p - \frac{1}{2}p^2}. \tag{8}$$

Note that if (4) holds, then so does (8). Thus, if it is socially desirable for a second high-cost firm to invest, patent protection will induce it to do so, whereas without such protection, it might merely imitate. Patents accomplish more, therefore, than merely protecting innovators from imitation; they encourage would-be imitators to innovate themselves. Indeed, they create a risk of *over*-investment in R&D: (8) could hold even if (4) did not.[18] For low-cost firms, the innovation incentives are even stronger.

We summarize these results with:

---

[17] This gets at the idea that patents have *breadth*, and so a patent holder can block the implementation of other firms' discoveries that are similar, but not identical, to his own.

[18] The possibility that patents can give rise to excessive spending on R&D is well known from the patent-race literature; see Dasgupta and Stiglitz (1980) and Loury (1979).

Proposition 1: *In the static model, the level of R&D investment in a regime without patents is less than or equal to the social optimum. In particular, if (4) holds, but (7) does not, then there will be underinvestment (if (7) holds, then investment will be optimal). By contrast, the level of R&D investment with patent protection is greater then or equal to the social optimum. Specifically, if (4) holds the investment will be optimal (if (4) fails but (8) holds, there will be overinvestment).*

Observe that the possible over-investment in R&D induced by patents could, in principle, be avoided if there were no complementarities of research across firms. Specifically, one could imagine awarding a firm an *ex ante* patent, e.g., the right to research and develop a vaccine against a particular disease.[19] Such protection would, of course, serve to deter additional firms from attempting to develop the innovation in question. But this would be efficient, provided that these other firms would not enhance the probability or speed of development, i.e., provided that they conferred no complementarity.

Notice that patent licensing serves no purpose in this static model. Without licensing, a patent holder obtains payoff of v. Were the patent to be licensed to the other firm, the firms would split v, but the patent holder could presumably extract the licensee's share in the form of a fee or royalty. Hence, nothing is gained (although nothing is lost) by licensing.[20]

---

[19] Wright (1983) and Shavell and Ypersele (2001) explore similar schemes.

[20] This finding would change if, as in much of the patent-licensing literature, we dropped the assumption that no profit is dissipated by competition (see footnote 13). It would also change if the firms developed complementary innovations that could advantageously be cross-licensed; see Fershtman and Kamien (1992).

This simple static model thus captures basic results of patent-race models such as Loury (1979) and Dasgupta and Stiglitz (1980). It also illustrates aspects of static models involving spillover complementarities, such as Spence (1984), who emphasizes the socially redundant R&D that can occur under patents. The principle that patents promote efficient or over-investment in R&D requires considerable modification, however, when dynamic considerations are introduced.

3. Dynamic Model

Let us now enrich the model to accommodate a *sequence* of potential innovations each of which builds on its immediate predecessor. Formally, consider an infinite sequence of potential innovations (indexed by $t = 1,2,...$), each of which has value v. To avoid the complications that arise when a new invention renders old products obsolete (a phenomenon called the *replacement effect*), we suppose that v constitutes *incremental* value (i.e., an invention is simply an improvement that enhances the value of the first innovation).[21]

We suppose that a patent on the first innovation is sufficiently broad to block all subsequent inventions. However, in the absence of patents, a firm is free to imitate any innovation.

As before, there are two firms and each firm's R&D cost per innovation is either $\bar{c}$ (with probability q) or $\underline{c}$ (with probability $1-q$). If just one firm undertakes R&D, the probability that innovation t is discovered, conditional on innovation $t-1$ having already been discovered, is p (if innovation $t-1$ has not yet been discovered, then there is *no* chance that innovation t will be found). The corresponding probability if both firms undertake R&D is $1-(1-p)^2$.

[21] Hunt (1995), O'Donoghue (1998), and O'Donoghue, Scotchmer, and Thisse (1998) have analyzed the replacement effect in sequential innovation. Such an effect could be incorporated in our model but would shed no additional light on the issues we emphasize. Our assumption also seems more consistent with technologically differentiated products. Notice that in our model (in contrast to that of Scotchmer and Green 1990) a firm has no incentive to keep an invention off the market while developing a sequential improvement.

. We assume that, if just one firm undertakes R&D and is immune from imitation, its expected payoff, when $\bar{c}$ is its R&D cost, is positive, i.e.,

(9) $$-\bar{c} + pv + p(-\bar{c}v + pv) + \ldots$$

$$= \frac{-\bar{c} + pv}{1-p} > 0.$$

Notice that the left-hand side of (9) also constitutes the expected social benefit from having one firm undertake R&D when its cost is $\bar{c}$. To simplify the exposition, we assume that the first firm has high costs.

Suppose that if the second firm's cost is also high, the marginal social benefit from having it undertake R&D is positive, i.e.,

(10) $$\bar{S} = \left[\frac{-2\bar{c} + \left(2p - p^2\right)v}{(1-p)^2}\right] - \left[\frac{-\bar{c} + pv}{1-p}\right]$$

$$= \frac{pv - \bar{c}\left(1+p\right)}{(1-p)^2} > 0,$$

which can be rewritten as

(11) $$\frac{v}{\bar{c}} > \frac{1+p}{p}.$$

We first show that, in the absence of patents, the incentive to innovate is greater in this dynamic model than in the static framework. As in the static model, if one firm conducts R&D, then, without patents, the other firm gets half the gross expected profit by simply imitating the first firm's innovation (and conducting no R&D itself). Hence, the imitating firm's expected payoff is

$$\frac{pv}{2(1-p)}. \tag{12}$$

However, if the second firm also undertakes R&D, then its payoff (when its R&D cost is $\overline{c}$) is

$$\frac{-\overline{c}+\frac{1}{2}\left(2p-p^2\right)v}{(1-p)^2}. \tag{13}$$

Hence, the second firm will undertake R&D and refrain from merely imitating if (13) exceeds (12), i.e., provided that

$$\frac{pv-2\overline{c}}{2(1-p)^2}>0,$$

which can be rewritten as

(14) $$\frac{v}{\bar{c}} > \frac{2}{p}.$$

Notice that the right-hand side of (14) is smaller than that of (7). Hence, a second firm is more likely to engage in R&D in this dynamic setting than in the static model.[22] This is because the second firm has a greater incentive to conduct R&D in the dynamic model: its R&D investment raises the probability, not only of the next innovation, but of subsequent innovations, and this is advantageous to the firm, even if subsequently it merely imitates the first firm. Notice too that the difference between the right-hand sides of (14) and (11) is smaller than the difference between those of (7) and (4). Hence, if we think of $v/\bar{c}$ as being drawn from a probability distribution that is roughly uniform, a second firm is more likely to engage in R&D, conditional on this R&D being socially optimal, in the dynamic then the static model. In other words, in a world without patents, there is a better chance (assuming rough uniformity) that socially desirable R&D will be undertaken in a sequential than in a static setting. Indeed, when p is near 1—the case we are most interested in, since it places the greatest weight on the dynamic aspect of the model—the probability that the no-patent regime generates the social optimum in the

---

[22] Indeed, the comparative statics move in opposite directions in the two models. For p near 1, an increase in p makes two firms engaging in R&D more likely in the dynamic model but less likely in the static model (there is a corresponding discrepancy when comparing the socially optimal levels of R&D in the two models).

dynamic model (given that it is optimal to have two firms undertaking R&D) is practically 1 (since the difference between the right-hand sides of (14) and (11), $\frac{2}{p}$ and $\frac{1+p}{p}$ is small).

Next consider what happens if the first innovation is patentable.[23] A patent confers a hold-up right over subsequent inventions because these later discoveries make essential use of the patented innovation (they can be thought of as "significant embellishments").

Suppose that a firm discovers the first innovation and obtains a patent on it. An issue that the patent-holder faces that did not arise in the static model is whether or not to offer the other firm a license to use the invention. As we have seen, in the static model, licensing serves no purpose because, by assumption, there are no further discoveries to be made. In the dynamic model by contrast, licensing generates a potential benefit, viz., it may induce the licensee to participate in the search for further inventions, thereby raising the probability of their discovery. And, by charging a suitable licensing fee, the patent holder can attempt to capture some of this potential benefit for itself.

The complication, however, is that, not knowing the other firm's R&D cost, the patent holder cannot ensure that its fee will be set to extract all surplus. Indeed, given its incomplete information, its

[23] Our exposition proceeds as though subsequent innovations are not patentable. However, the analysis would not change significantly if *every* innovation were patentable (as long as a firm needed a license for innovation $t-1$ in order to discover innovation t).

optimal strategy may well be to set a fee that, with positive probability, discourages the other firm from obtaining a license at all.[24]

To see this, assume that bargaining over the fee takes the form of a take-it-or-leave-it offer by the patent holder. This assumption provides a simple way to capture the inefficiency arising from bargaining.[25] Because R&D costs take only the values $\underline{c}$ and $\overline{c}$, the patent holder has essentially three options: (i) to set a fee so that the other firm will want to buy a license whether it has cost $\underline{c}$ or $\overline{c}$; (ii) to set a fee so that only if the other firm's cost is $\underline{c}$ will it want a license; and (iii) to set a fee so high that the other firm never wants a license. But if (11) holds, option (i) dominates (iii): under (i), the patent holder can set a fee that extracts the entire marginal surplus $\overline{S}$ that derives from having the second firm conduct R&D when its R&D cost is $\overline{c}$. $\overline{S}$ is equal to the left-hand side of (10), which is positive because (11) holds (the patent holder's fee will also extract $\overline{S}$ if the second firm's cost is $\underline{c}$, but in that case, $\overline{S}$ does not constitute the full marginal surplus). By contrast, this surplus is forgone under (iii).

Thus the patent holder's optimal strategy comes down to a choice between (i) and (ii). Under (i), the patent holder will extract $\overline{S}$ regardless of the other firm's R&D costs and the latter will always buy a license. Under (ii), by contrast, the patent holder can extract all the marginal surplus $\underline{S}$ deriving

---

[24] Bessen (2002) explores *ex ante* patent licensing more generally and presents empirical evidence that such licensing does not occur often in the computer and semiconductor industries.

[25] As long as the patent holder has positive bargaining power, other bargaining processes will also generate inefficiencies. Note that a license based on *ex post* observed R&D expenditures will be subject to inefficiencies arising from moral hazard.

from the second firm when the latter's cost is $\underline{c}$. Note that $\underline{S}$ exceeds $\overline{S}$ $\left(\text{specifically, } \underline{S} = \overline{S} + \frac{\overline{c} - \underline{c}}{(1-p)^2}\right)$, but corresponds to a fee so big that it will be rejected if the second firm's cost is $\overline{c}$. Hence if the patent holder chooses (ii), it will sell a license and receive $\underline{S}$ with probability $1 - q$, and it will refrain from licensing with probability q. Now, if its own cost is high, the patent holder will favor option (ii) over (i) when

$$q\left(\frac{pv - \overline{c}}{1-p}\right) + (1-q)\left(\frac{(2p - p^2)v - \underline{c} - \overline{c}}{(1-p)^2}\right)$$

$$\geq \frac{(2p - p^2)v - 2\overline{c}}{(1-p)^2},$$

which can be rewritten as

(15) $$1 - \underline{c}/\overline{c} \geq q\left(p\, v/\overline{c} - p - \underline{c}/\overline{c}\right).$$ [26]

Thus, when (14) and (15) hold, a regime in which innovations cannot be patented induces more R&D and innovation than in one in which patents are possible. Furthermore, in this case, society is better off without patents (indeed, the firms themselves are better off *ex ante*, since they capture the

---

[26] Notice that (15) will hold provided that q is sufficiently small.

social benefit). To see this, observe that, without patents, joint expected profit—which is the same as total expected welfare—is given by

$$(16) \qquad \frac{-2c + \left(2p - p^2\right)v}{(1-p)^2},$$

where $c = (1-q)\underline{c} + (1+q)\bar{c}$. Now, (14) implies that (11) holds, and under (11) formula (16) constitutes the maximum possible joint profit. But, as we have seen, a firm with cost $\bar{c}$ will not be licensed by a patent holder when (15) holds. Thus with patents, joint profit will be strictly less than (16). With patents, some socially beneficial licensing does not occur, interfering with innovation that would occur in the absence of patents.

We summarize these findings as follows:

Proposition 2: *Conditional on its being optimal for two firms to undertake R&D, the level of R&D investment without patents is more likely to be socially optimal in the dynamic than in the static model (provided that $v/\bar{c}$ is distributed roughly uniformly). Indeed, in the dynamic model, R&D investment without patents will be socially optimal if (14) holds. By contrast, in this same model, R&D investment in a regime with patent protection is strictly lower than the social optimum if (14) and (15) hold, implying that both society and the firms themselves are better off ex ante without patents.*

In comparing the static and dynamic models, we have been focusing so far on the issue of whether or not *patents* are desirable (to society and to the firms). But there is another important difference between the two models, turning on the question of whether or not an innovating firm benefits from *competition* and *being imitated*.

In the static model, a firm undertaking R&D clearly loses from competition and imitation. If it is the only firm pursuing innovation, its expected payoff (assuming that its R&D costs are high) is $pv - \bar{c}$. Against a competitor, however, the firm's expected profit is only $\frac{1}{2}\left(2p - p^2\right)v - \bar{c}$ (this assumes that the competitor engages in R&D, i.e., that (7) holds; if not, profit is even lower: $\frac{1}{2}pv - \bar{c}$).

By contrast, in the dynamic model, a firm with high R&D costs has an expected profit of

$$\frac{pv - \bar{c}}{1 - p} \tag{17}$$

if it has no rivals. But, if (14) holds, its profit is

$$\frac{\frac{1}{2}(2p - p^2)v - \bar{c}}{(1 - p)^2} \tag{18}$$

when it has a competitor (and there is no patent protection). Now, given (14), (18) exceeds (17), and so in this case both the dictum that "competition expands the market" and Apple's welcoming greeting to I.B.M. (see footnote 4) are fully justified.[27]

To summarize, we have:

Proposition 3: *A firm never benefits from competition and being imitated in the static model. By contrast, it gains from facing a competitor that is free to imitate it in the dynamic model (provided that (14) holds, implying that the competitor will choose to undertake R&D).*

4. Empirical evidence of dynamic innovation

We present two types of empirical evidence suggesting that the dynamic rather than the static model applies to high-tech industries. First, we exhibit data showing that innovation in these industries is both sequential and complementary in the sense we have discussed. Second, we show that the natural experiment in software protection is difficult to reconcile with the static model.

---

[27] This logic shows that a firm may welcome competition in the absence of patent protection; a similar argument shows that the same is true when patents are possible.

Evidence of Sequential and Complementary Innovation

There is, of course, much anecdotal evidence about the sequential and complementary nature of innovation in high-tech industries. New and better versions of microprocessor chips and PC operating systems appear regularly every three years or so. New entrants, rather than incumbent firms, have made some of the most important contributions in the software and semiconductor industries.

Data from a study of major new products by Michael Gort and Steven Klepper (1982) permit a more systematic evaluation of these hypotheses. For 23 major products, Gort and Klepper collected data on innovations (major and minor) and on firm entry over time. On average, each of these products experienced 19 subsequent improving innovations. It is possible that additional minor innovations went unreported. In any case, these data indicate the sequential nature of innovation.

The hypothesis of complementarity is somewhat more difficult to demonstrate. According to our hypothesis, different firms pursue somewhat different technologies and this improves the joint prospect of finding a successful innovation. All else equal then, a greater number of technologies pursued in an industry should generate a higher rate of innovation. Often these new technologies are introduced to an industry by a new firm. Indeed, a firm may *choose* to enter an industry precisely because it has a potentially valuable complementary technology.

However, the ability of firms (new or old) to introduce new technologies changes over the course of the product life cycle. From the Gort and Klepper study, we know that, initially, a single firm typically enjoys a monopoly, often supported by an initial set of patents. When these patents expire (or entry is allowed for other reasons), other firms freely enter and new technologies can be introduced more readily. Over time, additional innovations are made to improve the product. As some firms build large portfolios of in-force patents on these improvements, entry once again becomes more difficult. Combined with maturing demand and learning effects, new firms stop entering and less successful firms exit, resulting eventually in a stable industry.

Thus, when firms can enter, they can also introduce new technologies. On the other hand, when entry is restricted, the pool of firms that can introduce new technologies is also restricted. Hence, the rate of firm entry can proxy for the ability of firms to introduce complementary technologies. If our hypothesis is correct, the rate of innovation should correlate with the rate of firm entry. Note that this is true *only* as long as entry does not also diminish innovation incentives too much.

The evidence squarely supports this hypothesis. Gort and Klepper divide their data into five phases of each product's life-cycle, the phases defined by net entry behavior. The first phase is the monopoly stage (or near-monopoly in a few cases); the second exhibits positive net entry, in the third, entrants roughly balance exiters, the fourth has negative net entry, and the fifth exhibits rough stability again. For each phase they report the annual rate of entry and of innovation. Innovation rates are further divided into rates for major innovations and for minor innovations.

Table 1 and Figure 1 show the means (weighted by duration) of the annual rates of innovation for each phase for both major innovations and total innovations. As can be seen, neither the initial monopoly phase nor the final phases—both periods when entry is most constrained—have particularly high rates of innovation. The highest rates of innovation appear instead during the second and third phases, during and immediately following the period of greatest firm entry.

This result can be explored more formally with a Poisson model of the innovation count data. For the *i*th product during a phase of duration Δt_i, we assume that the hazard for an innovation, λ_i, is an exponential function of the net entry rate of firms, n_i:

$$\lambda_i = \Delta t_i \cdot e^{\alpha + \beta n_i}.$$

The probability that the number of innovations during this period is y is

$$P(Y = y) = \frac{e^{-\lambda_i} \cdot \lambda_i^y}{y!}.$$

It is possible that changes in both the rate of innovations and the rate of firm entry result from exogenous changes in technological opportunities. That is, firms may choose to enter when opportunities to innovate are greater. In this case, the independent variable would be correlated with the error term. To correct for this, we perform an instrumental variables estimation.

We begin, however, with a straightforward maximum likelihood estimation of this simple model displayed in column 1 of Table 2, both for all innovations (top) and for only major innovations (bottom). The results show a significant positive relationship between entry and innovation.

The Poisson regression model assumes that the variance of y equals the mean. But this will not be the case if there are stochastic errors in addition to the Poisson sampling error (see Hausman, Hall and Griliches, 1984, Cameron and Trivedi, 1986). To allow for possible "over-dispersion," we also performed the negative binomial regression described in Cameron and Trivedi. These regressions are shown in column 2. The coefficients are quite similar, positive and highly significant. A likelihood ratio test indicates that over-dispersion does exist, hence the negative binomial model is preferred.

For comparison we also perform a nonlinear least squares estimation in column 3. This method is consistent, but not efficient for our model, and so we estimate standard errors using White's (1980) heteroscedastic-consistent method. Results are similar, but in the estimation on major innovations the coefficient for entry is significant only at the 5% level.

As mentioned, it is possible that a third factor such as "technological opportunities" could be positively correlated with both the rate of entry and the rate of innovation. Perhaps periods of greater opportunity might generate both more innovation and greater entry of firms seeking to capitalize on this opportunity. In this case, the correlation between rate of entry and innovation might be overstated or spurious.

To correct for this possible endogeneity in the independent variable, we instrument the rate of firm entry with two variables. Desirable instruments should be correlated with the rate of entry, but uncorrelated with changes in technological opportunity. The first instrument is the average number of firms in the industry over the entire product life.[28] Although technological opportunity may influence the *gross* rate of entry, the exit process is independent, and the equilibrium number of firms over all phases would seem to be determined by market size and structure independently of opportunity. The second instrument is a simple dummy flag that takes the value of 1 during the initial monopoly phase and 0 otherwise. The initial monopoly period, if it exists, is presumed to result from an original set of strong, broad patents and should thus be independent of technological opportunity as well.[29]

Estimates using these instruments are shown in Column 4. Again, coefficients are similar and the coefficient on firm entry is significantly positive. A Hausman specification test does not support the hypothesis of endogeneity.[30] Thus the relationship between innovation and entry appears to be independent of technological opportunities.

Note, moreover, that this result would not obtain if entry destroyed innovation incentives. In the static model of intellectual property, innovation incentives depend on the patent holder's ability to extract monopoly rents. The magnitude of these rents depends on product market conditions. Rents will

---

[28] These supplementary data were graciously provided by Steven Klepper.

[29] If initial monopolies do not arise from patent protection, then the static model would be irrelevant in any case. Note further that since we are instrumenting a nonlinear least squares estimation, the instruments apply to the pseudo-regressors of a linearized model, not to the rate of entry directly. To correspond to the form of the pseudo-regressors, the instruments were multiplied by the duration of the phase. Also, terms were included using the square of the average number of firms, the phase duration and a constant.

[30] A Hausman specification test could not reject the null hypothesis that the simple nonlinear least squares estimator was consistent. The instrumental variables estimation was repeated using only the second instrument, the initial patent flag, plus phase duration and a constant. Results were positive with an even higher coefficient. Generally similar although sometimes less significant results were also obtained including industry dummies and performing a fixed effects analysis conditioning on the product sums [Hausman, Hall and Griliches, 1984].

be greatest when the patent holder enjoys a complete monopoly; rents will generally be inferior when other firms can enter the product market, even if they produce only imperfect substitutes.[31] Thus a high rate of firm entry is often taken as *prima facie* evidence of insufficient appropriability. Clearly this is not a valid inference.

Thus the Gort and Klepper data provide evidence supporting our key assumptions that innovation is both sequential and complementary. And this analysis also suggests that the static model is inconsistent with the evidence on entry. The recent experience of the software industry provides additional evidence on this score.

The Natural Economic Experiment in Software

The semiconductor, computer and software industries have historically experienced high levels of innovation despite weak patent protection. This suggests that the dynamic model is applicable, but, by itself, this evidence is not conclusive. Although these industries have been innovative without strong patent protection, perhaps they would have been *far more innovative* with strong protection; perhaps these industries offer many technological possibilities, but only the most highly profitable possibilities are realized under weak patent protection.

---

[31] For example, consider the case where an innovator holds a patent on an improvement to a base product. The rents on the improvement will be greatest when the innovating firm has a monopoly on the base product as well. In general, the innovating firm will not realize the same rents from the improvement if entrants can freely offer an unimproved base product as a substitute—the unimproved version of the product can be offered at a lower price, tending to dissipate some of the rents.

Fortunately, this alternative explanation can be tested. The patent courts subjected the software industry to a natural economic experiment during the 1980's.[32] Before this time, patent protection for innovations was very limited; instead, innovations were protected by copyright. This meant practically that direct copying of a software product was prohibited, but that copying the ideas and concepts embodied in software was not. Market entry therefore required significant investment in development, but entry could not be barred.

A series of court decisions in the early 1980’s had the effect of extending patent protection to many software ideas. Consequently the number of patents issued annually covering software grew exponentially from the mid-80's to about 7,000 in 1995 (see Figure 4). Within the software industry, this has sometimes been described as a case of "fixing what ain't broke." Advocates counter, arguing along the lines of the static model, that increased patent protection should increase software innovativeness (USPTO, 1994).

If the static model is correct, then the extension of patent protection should have produced a sharp increase in R&D spending among those firms and industries applying for patents. This should have subsequently been followed by an increase in productivity growth. The changes should be measurable and large after controlling for other, possibly offsetting changes.

---

[32]Some other natural experiments involving the extension of patent protection are Scherer and Weisburst (1995) and Challu (1995), and Sakakibara and Branstetter (2001).

According to the static model, R&D should increase with patent protection because firms can profitably pursue R&D projects that yield smaller returns, that is, projects with lower values of $v/\bar{c}$. This can be seen as follows. As noted in section 2, projects with low values of $v/\bar{c}$ may be unprofitable without patent protection, but, become feasible with patents. Assume a stationary distribution of R&D opportunities ranked by $v/\bar{c}$ such that $F(v/\bar{c})$ is the cumulative R&D spending required to invest in all opportunities with a return less than $v/\bar{c}$. For simplicity assume F is concave. For the case without patents, designate the entry threshold value of $v/\bar{c}$ for one firm as T_1^N and the threshold for two firms as T_2^N (from section 2, we have $T_1^N = \frac{2}{p}$ and $T_2^N = \frac{2}{p-p^2}$). Then all opportunities with returns between x and $x+dx$ such that $T_1^N \leq x < T_2^N$ will consume in total $dF(x)$ R&D dollars and will generate expected value of $pxdF(x)$. Opportunities where $T_2^N \leq x$ will require $2dF(x)$ R&D dollars (two firms investing) and will generate expected value of $p(2-p)xdF(x)$. Then the average value of $v/\bar{c}$ for the industry is the ratio of total value v to total R&D investment,

$$A_N = \frac{p\int_{T_1^N}^{T_2^N} xdF(x) \; + \; p(2-p)\int_{T_2^N}^{\infty} xdF(x)}{1\int_{T_1^N}^{T_2^N} dF(x) \; + \; 2\int_{T_2^N}^{\infty} dF(x)} = \frac{p\int_{T_1^N}^{\infty} xdF(x) \; + \; p(1-p)\int_{T_2^N}^{\infty} xdF(x)}{\int_{T_1^N}^{\infty} dF(x) \; + \; \int_{T_2^N}^{\infty} dF(x)}.$$

With patents, the corresponding thresholds are $T_1^P = \frac{1}{p}$ and $T_2^P = 1/\left(p - \frac{1}{2}p^2\right)$, and the corresponding average value of $v/\bar{c}$ is A_P. Now, we have $T_1^P < T_2^P < T_1^N < T_2^N$. Using this, it is straightforward to show that $A_P < A_N$.

In other words, the *average* value of $v/\bar{c}$ should decrease for industries with the extension of patent protection. This logic can be readily extended to cases with more than two firms. Further, allowing each firm to have an equal chance of being an early-mover for any R&D opportunity means that the average value of $v/\bar{c}$ should also decrease for firms, or alternatively, the average value of $\bar{c}/v$ should increase.

For empirical analysis, it is useful to note two aspects of this predicted change. First, since productivity is increasing in these industries (see below), the net social value v will increase at least as fast as output. Therefore, an increase in $\bar{c}/v$ implies an increase in the ratio of R&D spending to output. In other words, the extension of patents should cause an increase in *relative* (to output) R&D spending. Relative R&D spending is a more useful measure than absolute R&D spending, given the changing composition of industries as firms acquire, divest, startup and discontinue product lines and industries grow.

Second, v represents a discounted stream of future values. Typically, the increase in value (and the associated increase in output) associated with an innovation will follow the expenditure of R&D only after considerable delay. For this reason, we should expect the ratio of R&D to output to increase quite rapidly upon the extension of patent protection and subsequently level off.

To summarize, if the static model holds, relative R&D spending should have increased sharply, followed by productivity. We examine these changes among three different samples of firms:

1.) The top 10 U.S. software patentees in 1995, accounting for 35% of the software patents issued to U.S. companies in that year,

2.) The industry groupings for computer hardware and programming services in the NSF R&D survey (company R&D funds for SIC 357 and part 737 and 871) [NSF, 1996, 1997], and,

3.) The grouping of computer, telecommunications and electronic components (SIC 357, 365-7) in the NBER R&D Masterfile [Hall, 1988], a listing of publicly traded U.S. firms.[33]

For the first and last samples, the R&D and sales measures are global. For the NSF sample, the R&D measures are domestic only and we measure R&D intensity using the NSF figures for sales for SIC 357 and 737.[34]

---

[33]Data for the top 10 firms was obtained from annual reports, 10-Ks and the NBER Masterfile. The series for AT&T was based on consolidated figures including NCR, the computer company which was purchased by AT&T in 1991. For this reason we use only the top 9 firms prior to 1991, although the difference is not significant. Both the NSF samples and the NBER R&D Masterfile are firm-based surveys where all of the R&D and sales of a firm are assigned to the SIC category of the firm's major product line. Thus our measures are diluted by non-software R&D and non-software output. Nevertheless, as long as software development constitutes a substantial portion of R&D, then we should expect to see a significant increase in R&D intensity.

We initially explore R&D spending relative to sales (R&D intensity) rather than output. The trend in these measures is shown in Figure 3. The late 80's display a leveling off and possibly a reversal of an upward trend in research intensity over the previous decade. There does not appear to be so much as a 10% increase in R&D intensity among the firms and industries obtaining software patents.[35] Real R&D intensity is displayed in Figure 5 and shows a clear decline in the late 80's.

There could be two sorts of offsetting changes: 1.) Technological opportunities may have simultaneously fallen abruptly, and, 2.) The cost of performing R&D could have simultaneously risen sharply. A decline in technological opportunity seems at odds with the continued rapid growth and rapid innovation in these industries. Hall (1993) performs an econometric analysis on the same NBER dataset and finds that the output elasticity of R&D did not fall during the 1980's, but instead rose.[36]

Hall also presents evidence that the general costs of performing R&D did not rise sharply. If R&D costs had increased overall, offsetting an erstwhile increase in R&D spending, then the R&D intensity of other industries should have *fallen*. Figure 4 presents ratios of R&D intensity of software-

---

[34] Note that beginning in 1985 FASB required that a portion of software development expense should be capitalized, hence reported R&D includes directly expensed items plus the amortization expense of capitalized software. The introduction of this change may have had a slight distortionary effect on reported R&D, tending to delay a portion of expenditures. The effect of this accounting change was to spread the impact of any sharp changes in R&D spending over two or three years. This effect was temporary, significant largely for pure software firms and of relatively brief duration (software is typically amortized over three years or less). This was not a substantial factor for the 10 largest software patentees and, based on this, would not seem to be a major factor for industry measures either.

[35] The NSF series becomes erratic after 1992 as the result of sample changes and as some firms were re-classified into different industries.

[36] The increase was concentrated among smaller public firms as large firms apparently lost productivity switching from mainframe technology to microcomputers. But overall technological opportunity did not decline.

related industries to the R&D intensity of the entire manufacturing sector. As can be seen, the relative R&D intensity of software-related industries fell over this time period. Thus, not only did these industries fail to show a large increase in relative R&D spending, but they lagged behind the rest of the manufacturing sector over this period.

It is possible, however, that R&D spending relative to sales may understate R&D relative to output because of price effects. That is, as firms gain monopoly power with patents, prices may rise, inflating the sales figure in the denominator. To consider this possibility, Figure 5 displays the ratio of real R&D to output where R&D has been deflated using the NBER R&D deflator and sales have been deflated by a shipments-weighted index derived from the NBER Productivity Database for the industries involved. As can be seen, R&D relative to output exhibits a significant decline during the late 1980's. Perhaps prices have been mis-measured for the computer industry. However, it seems unlikely that any measurement error could be so large as to mask major price increases. Hence this evidence is hard to reconcile with the static model.

Hall has suggested (1993) that competition may have hit the large mainframe firms in the industry especially hard as new firms entered the computer industry in the early 1980's. Consequently the response of the large firms (and by implication industry averages) might not be representative of firms in the industry as a whole. To consider this possibility, we examined two sub-samples from the software-related firms in the NBER R&D Masterfile: a balanced panel of 49 small firms and an unbalanced panel of new public firms.[37] Figure 6 shows the R&D intensity of these panels compared to the performance of the top 9 software patentees. As can be seen, R&D spending diverged between these groups during the early 1980's, consistent with Hall's interpretation, but these groups did not increase relative R&D spending either during the late 1980's in response to software patents.

Thus, the extension of patent protection to software did not generate a relative increase in R&D spending as predicted by the static model; instead, R&D spending seems to have remained roughly steady or to have declined. Not surprisingly, these industries did not demonstrate increased productivity growth as a result of the patent bonanza, as seen in Figure 7. Although multi-factor productivity may have fallen for reasons related to the transition from mainframes to microcomputers, there is no evidence of any underlying productivity increase commensurate with the increase in patents.

Given that a variety of possible offsetting factors do not appear to explain the stagnation of R&D spending, two other explanations remain. First, the software ideas may have been highly appropriable by other means, such as first mover advantages or learning effects, both before and after the extension of patent protection. In other words, patents (and the static model) might simply be irrelevant to the software industry. But this explanation has a problem: why then do large firms spend

[37] The small firms are all those existing in 1980 and 1990 with fewer than 1,000 employees in 1980. The new firms are defined as firms that first enter the NBER R&D Masterfile after 1973 and have fewer than 5,000 employees their first recorded year. Conversations with Compustat confirmed that this procedure was likely to screen out most spin-offs, re-organizations and listing changes. New firms were dropped from the panel after 8 years.

millions of dollars each year obtaining hundreds of software patents? Clearly, at the very least, some sort of holdup problem is involved that counters the simple static model.

The second explanation is that patents themselves may generate negative effects that offset the greater innovation incentives for the initial innovator. This might be the reduction in the value of patents through replacement as modeled by Hunt (1995) and O'Donoghue (1998). Or it might be the loss of complementarity described in our dynamic model.

If these latter explanations hold, we would expect an ultimate reduction in R&D. However, this transition might be quite gradual for two reasons. First, any patent holder faces a large body of well-established prior art and possibly competing claims. In such an environment, a patent portfolio capable of fencing off an area of research can be built up only gradually. In fact, there has been relatively little software patent litigation so far and the companies with large portfolios are only just beginning to pursue software patent claims (Business Week, 1997).[38]

Second, some of the most innovative firms may be reluctant to aggressively pursue patent claims. As we have seen above, although static firms will be better off with patent protection, dynamic firms may actually be better off without it. Thus the most innovative firms might seek to maintain industry norms of cooperation rather than to aggressively exert all patent rights. In this case, these norms will

[38] Also, given the incompleteness of patent portfolios, much of this activity is directed not toward exclusive control of a market, but toward extracting royalties. Nevertheless, excessive royalties may limit complementary activity at the margin.

deteriorate slowly, and so problems of exclusive development may appear slowly. In fact, support for cooperative norms appears strong among many innovative software companies—senior executives from companies such as Microsoft, Sun and Oracle have expressed a general reluctance to pursue patent litigation and view their patenting activity as primarily defensive [PC Magazine, 1997, USPTO, 1994].

Indeed, pure software companies as a whole have not applied for many patents. A naive view might expect software patents to be obtained predominately by firms in the computer programming and data processing industry (SIC 737). In fact, the largest software patentees are in the computer hardware and telecommunications industries—industries which sell software products and also incorporate software in hardware products. The top 10 U. S. firms obtaining software patents in 1995 are listed in Table 3. The top ranked pure software firm in 1995 was Microsoft (rank 24) with 39 software patents.[39]

Of course, these industries have remained innovative and productivity growth is still positive. This does not, however, contradict the dynamic model; rather, the negative effects of the patent extension may not be felt for some time as industry cooperative norms continue and as litigation remains limited. The bill for this experiment has not yet come due.

---

[39]The software patent series used in this analysis were developed by Greg Aharonian of the Internet Patent News Service. The criteria for software patents include not only the USTPO patent class, but also detailed examination of the specification, claims and abstract.

Conclusion

Intellectual property appears to be an area in which results that seem secure in a static model are overturned in a dynamic setting. Imitation invariably inhibits innovation in a static world; in a dynamic world, imitators can provide benefit to both the original innovator and to society as a whole. Patents foster innovation incentives in a static world; in a dynamic world, firms may have plenty of incentive to innovate without patents and patents may constrict complementary innovation.

This suggests a cautionary note regarding intellectual property protection. The reflexive view that "stronger is always better" is incorrect; rather a balanced approach is required. The ideal patent policy limits "knock-off" imitation, but allows developers who make similar, but potentially valuable complementary contributions. In this sense, copyright protection for software programs (which has gone through its own evolution over the last decade) may have achieved a better balance than patent protection. In particular, industry participants complain that software patents have been too broad and too obvious, leading to holdup problems [USTPO]. Also in this regard, patent systems that limit patent breadth, such as the Japanese system before the late 1980's, may offer a better balance. Thus our model suggests another, different rationale for narrow patent breadth than the recent economic literature on this subject.

References

AHARONIAN, G. 1995. *Patent News Service.*

ARROW, K. 1962. “Economic Welfare and the Allocation of Resources for Inventions,” *The Rate and Direction of Inventive Activity,* Princeton Univ.

BARTELSMAN, E. AND GRAY, W., 1996. "The NBER Manufacturing Productivity Database", *NBER Technical Working Paper,* no. 205.

BESSEN, JAMES. 2002. “Hold-up and Patent Licensing of Cumulative Innovations with Private Information,” ROI Working Paper.

CAMERON, A. AND TRIVEDI, P. 1986. “Econometric models based on count data: comparisons and applications of some estimators and tests,” *Journal of Applied Econometrics***,** v. 1**,** p. 29.

CHALLU, P. 1995. “Effects of the monopolistic patenting of medicine in Italy since 1978,” *International Journal of Technology Management*, v. 10**,** p. 237.

CHANG, H. 1995. “Patent scope, antitrust policy and cumulative innovation,” *RAND Journal of Economics,* v. 26, no.1, p. 34.

DASGUPTA, P. AND MASKIN, E. 1987. “The simple economics of research portfolios,” *The Economic Journal,* v. 97**,** p. 581.

DASGUPTA, P. AND STIGLITZ, J. 1980. “Uncertainty, Industrial Structure, and the Speed of R&D,” *Bell Journal of Economics,* 11, p. 1.

FERSHTMAN, C. AND KAMIEN, M. 1992. “Cross licensing of complementary technologies,” *International Journal of Industrial Organization,* v. 10**,** p. 329.

GALLINI, N. 1984. "Deterrence through Market Sharing: A strategic incentive for licensing," *American Economic Review,* v. 74, p. 931-41.

GALLINI, N. AND WINTER, R. 1985. "Licensing in the Theory of Innovation," *RAND Journal of Economics,* v. 16, p. 237-252.

GREEN, J. AND SCOTCHMER, S. 1995. "On the division of profit in sequential innovation," *RAND Journal of Economics,* v. 26, no.1, p. 20.

GORT, M. AND KLEPPER, S. 1982. "Time paths in the diffusion of product innovations," *The Economic Journal,* v. 92, p. 630.

GRINDLEY, P. AND TEECE, D. 1997. "Managing intellectual capital: licensing and cross-licensing in semiconductors and electronics," *California Management Review,* v. 39, no. 2, p. 8.

HALL, B. 1988. "The manufacturing sector master file: 1959-1987," *NBER Working Paper Series,* No. 3366, p. .

HALL, B. 1993. "Industrial research during the 1980s: Did the rate of return fall?," *Brookings Papers: Microeconomics 2,* , p. 289.

HAUSMAN, J., HALL, B., AND GRILICHES, Z. 1984. "Econometric models for count data with an application to the patents-R&D relationship," *Econometrica,* v. 52, no. 4, p. 909.

HUNT, R. 1995. "Nonobviousness and the Incentive to Innovate," *Federal Reserve Bank of Philadelphia Working Paper,* 99-3.

KATZ, M. AND SHAPIRO, C. 1985. "On the Licensing of Innovation," *RAND Journal of Economics,* v. 16, p. 504-20.

KATZ, M. AND SHAPIRO, C. 1986. "How To License Intangible Property," *Quarterly Journal of Economics,* v. 101, p. 567-89.

LEE, T. AND WILDE, L. 1980. "Market Structure and Innovation: A Reformulation," *Quarterly Journal of Economic,* 94, p. 429.

LEVIN, R. 1982. "The Semiconductor Industry," *Government and Technical Progress: A cross-industry analysis,* Nelson, R. ed., Pergamon, NY.

LEVIN, R. AND REISS, P. 1988. "Cost-reducing and demand-creating R&D with Spillovers," *RAND Journal of Economics,* v. 19, no. 4, p. 538.

LEVIN, R., KLEVORICK, A., NELSON, R. AND WINTER, S. 1987. "Appropriating the Returns from Industrial Research and Development," *Brookings Papers on Economic Activity [3], ,* p. 783.

LOURY, G. 1979. "Market Structure and Innovation," *Quarterly Journal of Economic,* 93, p. 395.

MANSFIELD, E., SCHWARTZ, M. AND WAGNER, S. 1981. "Imitation Costs and Patents: An empirical study," *Economic Journal,* v. 91, p. 907.

MYHRVOLD, N. 1997. "Interview," *PC Magazine,* Sept. 9, 97, p. 30.

NATIONAL SCIENCE FOUNDATION. Annual. *Research and Development in Industry*.

NATIONAL SCIENCE FOUNDATION. Annual. *Science and Engineering Indicators*.

NORDHAUS, W. 1969. *Invention, Growth and Welfare: A Theoretical Treatment of Technological Change,* MIT, Cambridge.

O'DONOGHUE, T. 1998. "A Patentability Requirement for Sequential Innovation," *RAND Journal of Economics,* v. 29, p. 654-79.

O'DONOGHUE, T., SCOTCHMER, S. AND THISSE, J. 1998. "Patent Breadth, Patent Life and the Pace of Technological Progress," *Journal of Economics and Management Strategy,* **v. 7,** p. 1-32.

OFFICE OF TECHNOLOGY ASSESSMENT 1992. *Finding a Balance: computer software, intellectual property and the challenge of technological change*, Congress of the US.

SAGER, I. 1997. "Big Blue is out to collar software scofflaws," *Business Week*, March 17, 1997.

Sakakibara, M. and Branstetter, L. (2001). "Do Stronger Patents Induce More Innovation? Evidence from the 1998 Japenese Patent Law Reforms,"*Rand Journal of Economy*, v.32, p. 77-100.

SCHERER, F. 1980. *Industrial Market Structure and Economic Performance*, **,** Rand McNally.

SCHERER, F. AND WEISBURST, S. 1995. "Economic effects of strengthening pharmaceutical patent protection in Italy," *International Review of Industrial Property and Copyright Law,* p. 1009.

SCOTCHMER, S. 1991. "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law," *Journal of Economic Perspectives,* v. 5, p. 29-41.

SCOTCHMER, S. 1996. "Protecting early innovators: should second-generation products be patentable?," *RAND Journal of Economics,* v. 27, no. 2, p. 322.

SCOTCHMER, S AND GREEN, J. 1990. "Novelty and Disclosure in Patent Law," *RAND Journal of Economics,* v. 21, p. 131-46.

SHAVELL, STEVEN AND TANGUY VAN YPERSELE. 2001. "Rewards vs. Intellectual Property Rights," *Journal of Law and Economics*, v. XLIV, pp. 525-548.

SPENCE, M. 1984. "Cost reduction, competition, and industry performance," *Econometrica,* **v.** 52, no. 1, p. 101.

TAYLOR, C. T. AND SILBERSTON, Z. A. 1973. *The economic impact of the patent system: a study of the British experience,* Cambridge Univ. p. 294.

TILTON, J. 1971. *International Diffusion of Technology: The Case of Semiconductors,* **,** Brookings Inst.

USPTO 1994. "Hearings on Software Patent Protection," *United States Patent and Trademark Office*, January-February 1994.

WEBBINK, D. 1977. "The semiconductor industry : a survey of structure, conduct, and performance : staff report to the Federal Trade Commission," *Federal Trade Commission.*

WHITE, H. 1980. "A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedacticity," *Econometrica***, v. 48,** p. 817.

Table 1 . Weighted means of annual rates of innovation by phase.

	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5
Rate of all innovations	0.39	0.57	0.62	0.36	0.43
Rate of major innovations	0.19	0.29	0.22	0.18	0.22
Rate of net firm entry	0.22	5.05	-0.07	-4.97	0.16

Source: Gort and Klepper, 1982

Table 2. Regressions on innovation counts.

Regression	1	2	3	4
Regression method	Poisson	Negative Binomial	Nonlinear least squares	Nonlinear least squares instrumental variables
Dependent variable	Total number of innovations			
Coefficient of firm net entry rate	.046* (.007)	.034* (.003)	.052* (.014)	.062* (.009)
Constant	-.820* (.052)	-.694* (.105)	-.934* (.366)	-.943* (.120)
Θ in NEGBIN II	--	1.373* (.216)	--	--
R^2	.43	.37	.44	.43
Dependent variable	Number of major innovations			
Coefficient of firm net entry rate	.041* (.011)	.035* (.003)	.045 (.022)	.046* (.015)
Constant	-1.547* (.074)	-1.444* (.105)	-1.668* (.261)	-1.666* (.132)
Θ in NEGBIN II	--	1.560* (.216)	--	--
R^2	.35	.30	.36	.36

*Significant at the 1% level.
Asymptotic standard errors in parentheses, using White heteroscedasticity consistent standard errors for nonlinear regressions. Regressions cover 418 total innovations, 200 of these rated as major innovations, during 77 product phases over 887 product-years. Data are from Gort and Klepper [1982]. Instruments include a flag indicating initial monopoly phase, the average number of firms over the entire product life cycle, the square of the average number of firms, all multiplied by the duration of the phase, phase duration, and a constant. A Hausman specification test between the third and fourth columns does not reject the null hypothesis that the third column is consistent (P = .346 for all innovations and P = .917 for major innovations). NLLS-IV using only the monopoly flag times phase duration, duration and a constant as instruments generates significant and even larger coefficients on net entry.

Table 3. Top 10 Software Patentees, 1995

Firm	Software Patents Issued 1995	Total Utility Patents Issued 1995	R&D Spending 1994 (millions)
International Business Machines	503	1383	$3,382
AT&T	185	638	$3,110
Motorola	157	1012	$1,860
Xerox (including Fuji Xerox)	121	551	$895
Hewlett Packard	89	470	$2,027
Digital Equipment	80	189	$1,301
General Electric	59	758	$1,176
Apple Computer	57	129	$564
Ford Motor Co.	53	334	$5,214
Eastman Kodak	49	772	$859

Sources: PATNEWS, USPTO, annual reports

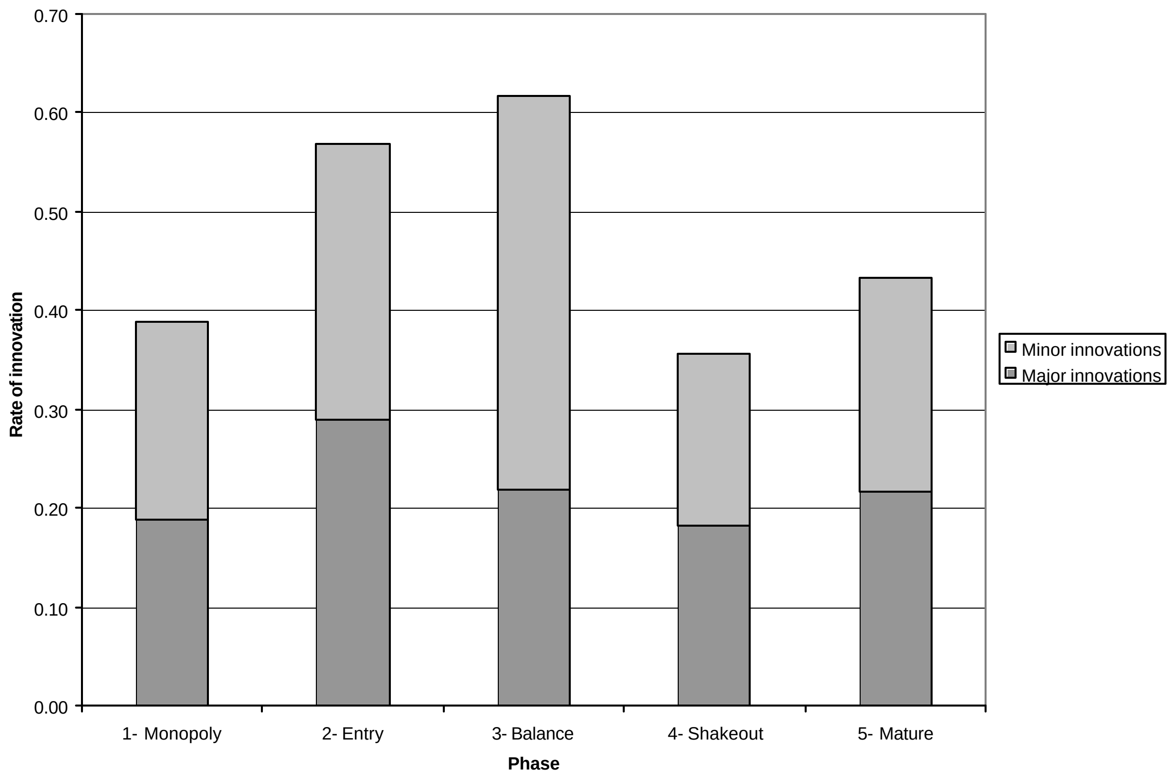
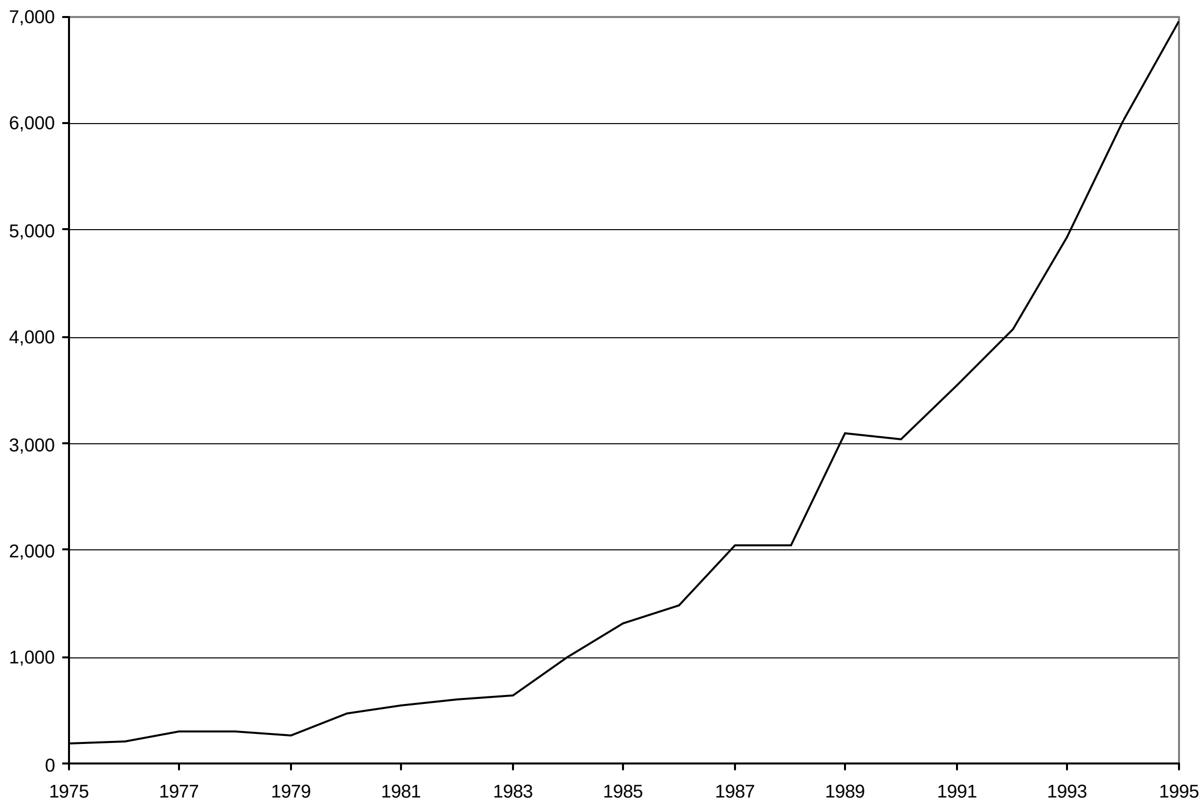


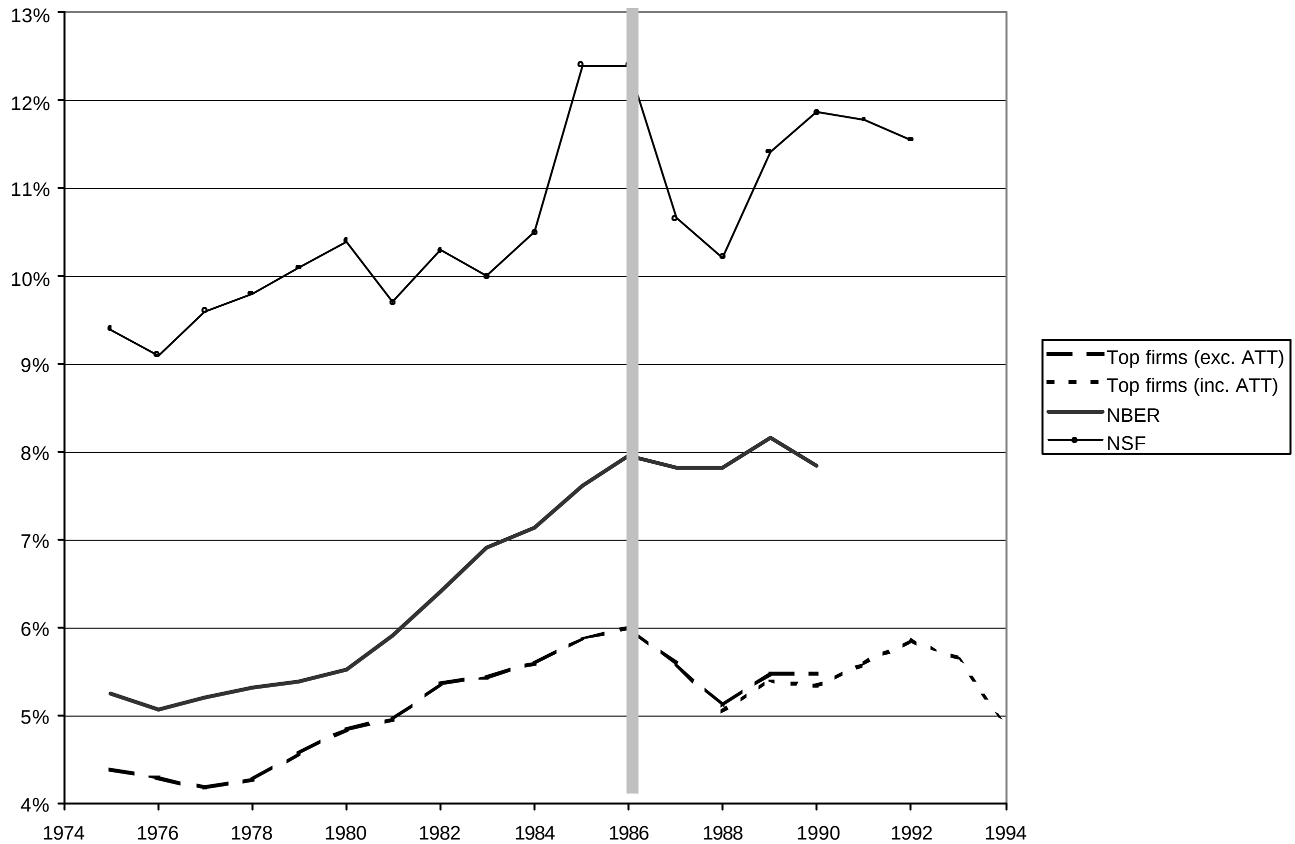
Figure 1. Innovation rates during product life-cycle phases

Figure 2 . U. S. Software Patents Issued



Source: Internet PATNEWS service.

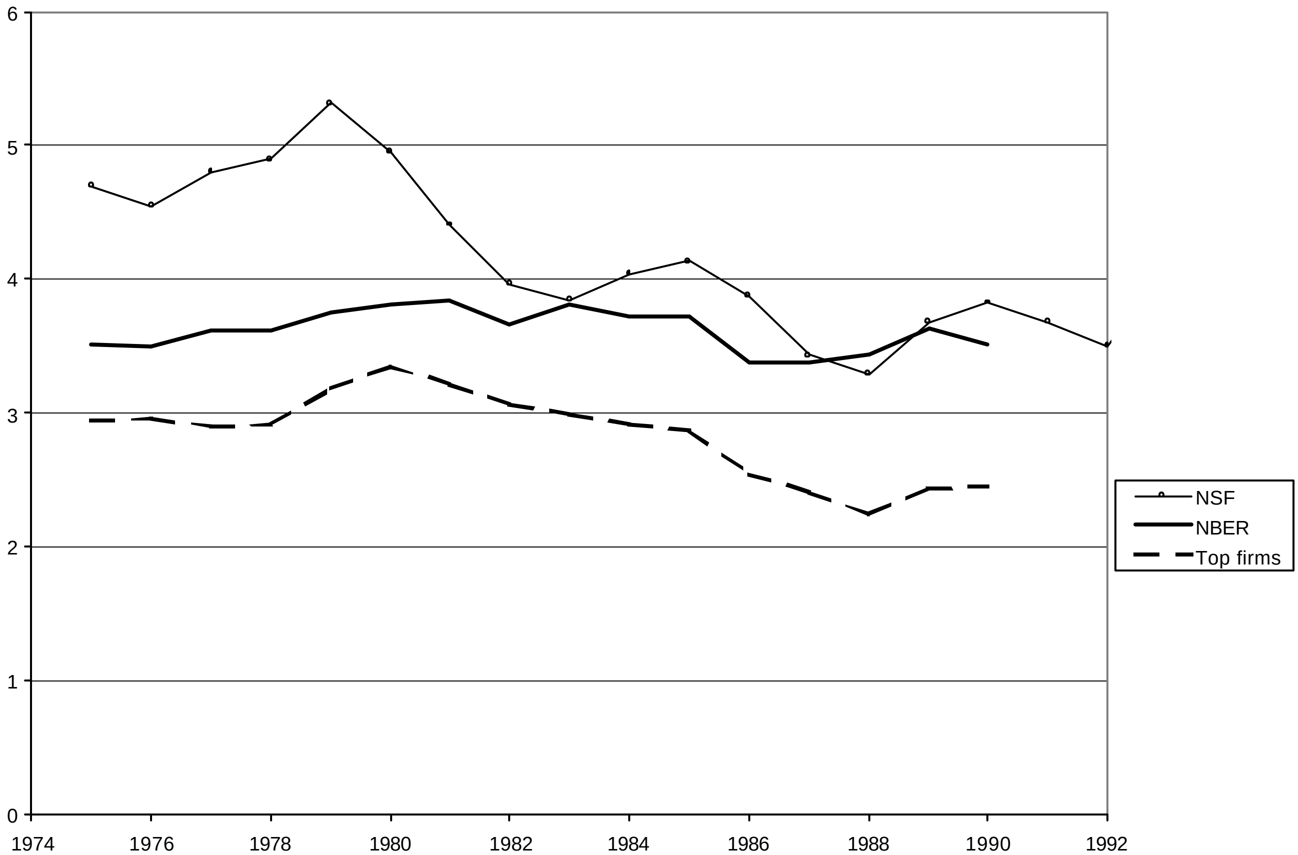
Figure 3. R&D Intensity for Software-related industries and firms



Sources: NSF *Research and Development in Industry, Science and Engineering Indicators*
NBER R&D Masterfile, Annual reports.

NBER series includes SIC 357, 365, 366, 367. NSF series includes SIC 357, and after 1986 part 737 and part 871. NSF series includes sample changes and hence is not directly comparable from year to year. Top firms come from Patent News Service rankings of software patents issued in 1995.

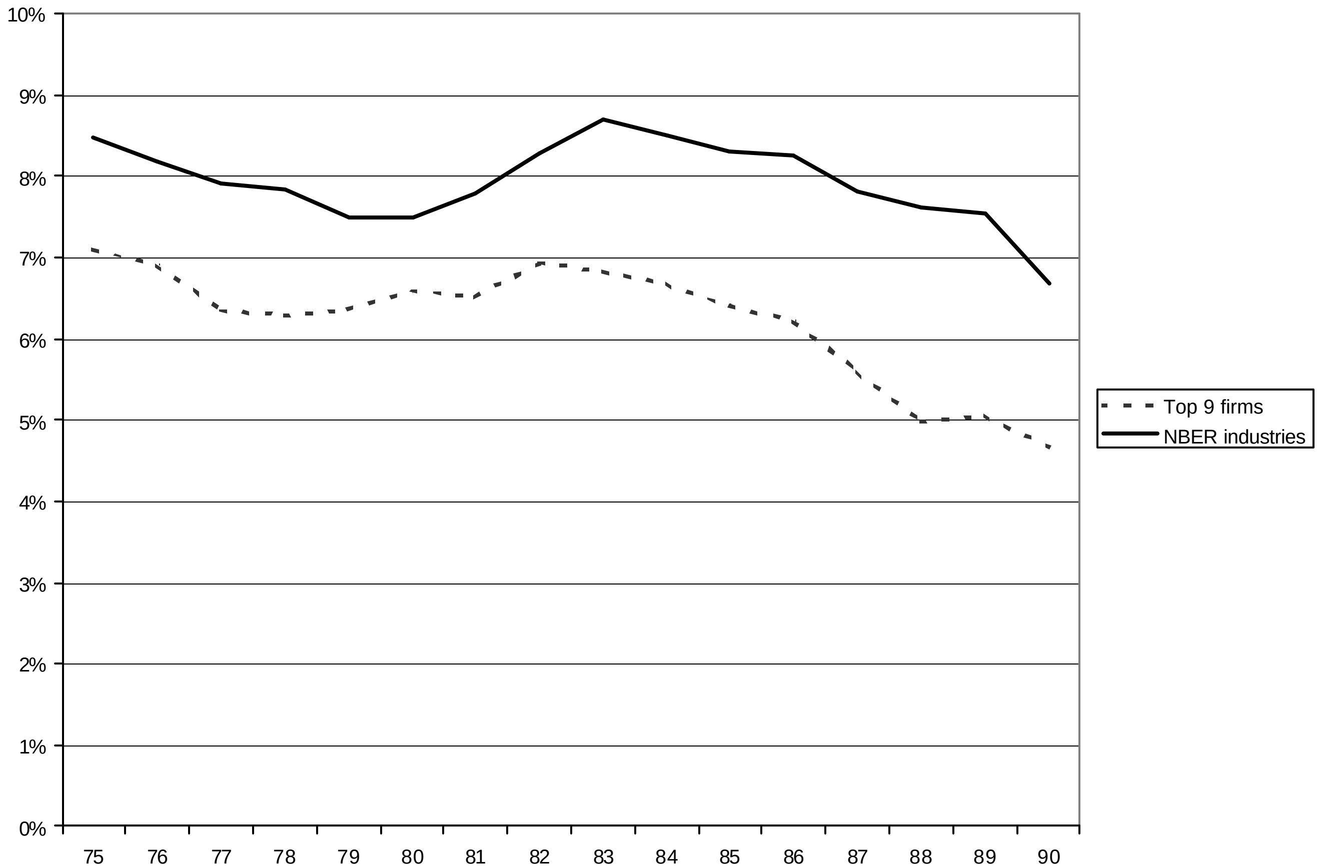
Figure 4. Relative R&D Intensity



Sources: NSF *Research and Development in Industry, Science and Engineering Indicators*
NBER R&D Masterfile, Annual reports.

Relative intensity is the ratio of R&D spending to output divided by that ratio of R&D spending to output for the entire manufacturing sector. NBER series includes SIC 357, 365, 366, 367.

NSF series includes SIC 357, and after 1986 part 737 and part 871.
Top firms come from Patent News Service rankings of software patents issued in 1995 relative to the NBER series for all manufacturing.



Sources: NBER R&D Masterfile, Annual reports.

NBER series includes SIC 357, 365, 366, 367. R&D is deflated using NBER R&D deflator. Sales are deflated using a shipments-weighted mean for these industries from the NBER Productivity Database.

Figure 5. Real R&D / Real Output (Deflated R&D Intensity)

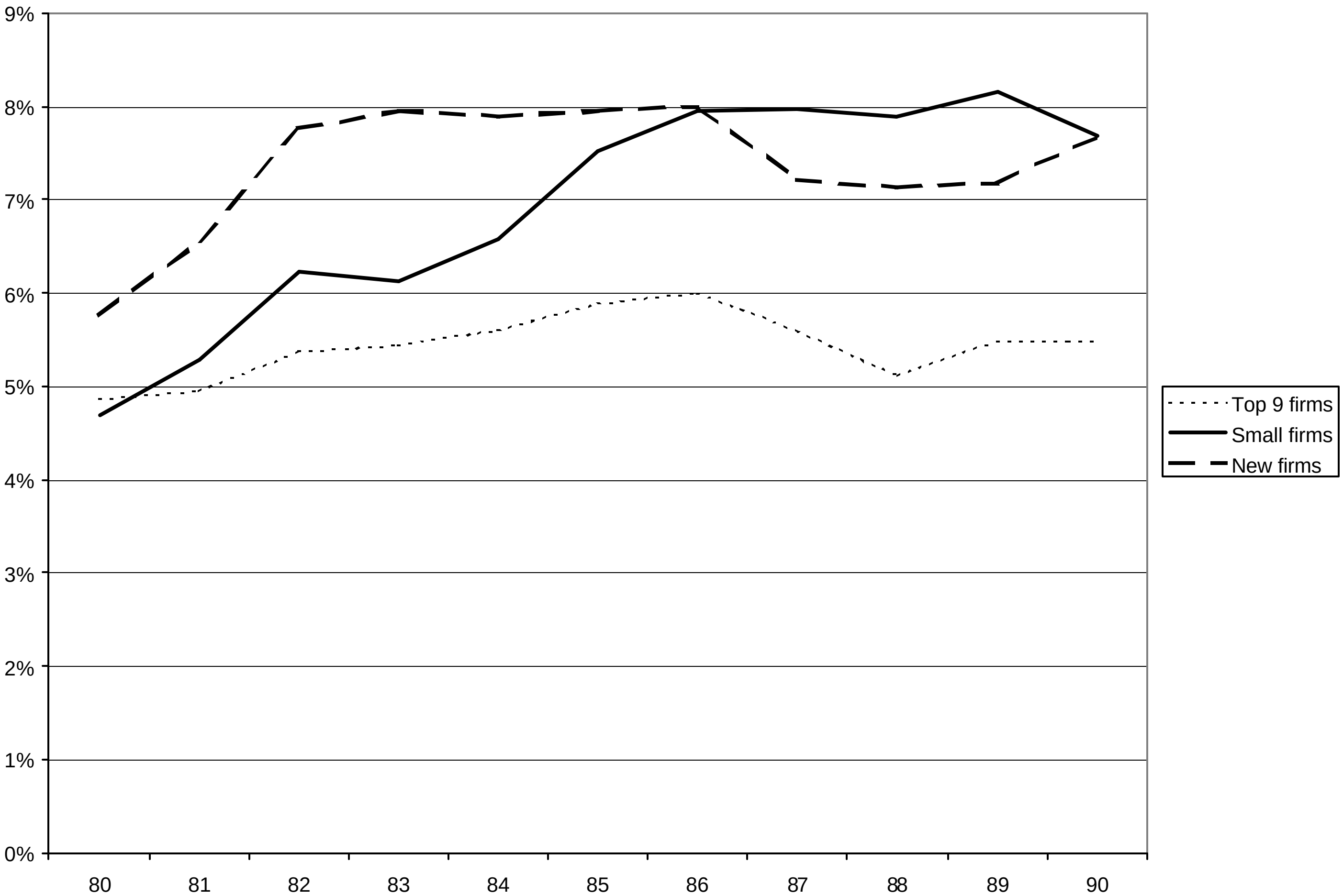
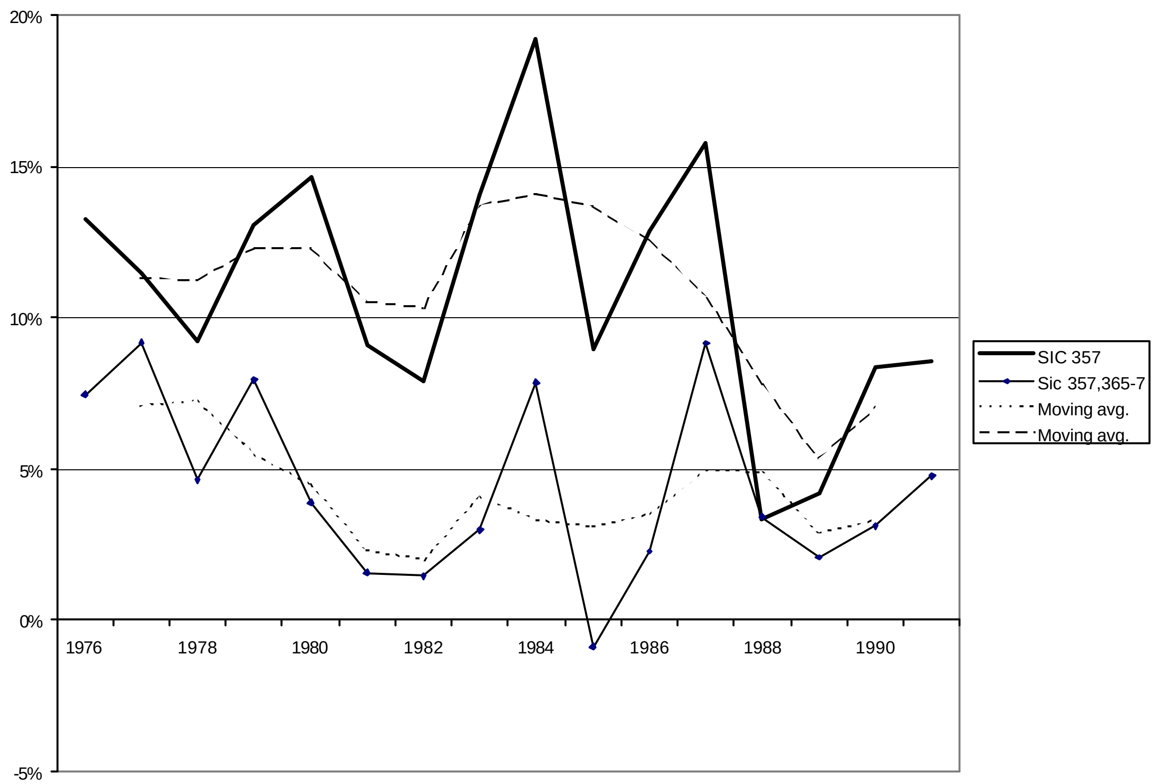


Figure 6. R&D Intensity for Small and New Firms

Source: NBER R&D Master File, Annual Reports

“Small firms” is a balanced panel of 49 firms from the software-related industries found in the NBER Master File in both 1980 and 1990 that had fewer than 1,000 employees in 1980. “New firms” are firms from the NBER Master File in software-related industries that first appear after 1973 and that had fewer than 5,000 employees their first recorded year. The unbalanced panel of new firms includes only the first eight years that a firm appears in the file. Figure 1. Total Factor Productivity Growth of Software-related Manufacturing Industries

Figure 7. Total Factor Productivity Growth of Software-related Manufacturing Industries



Source: NBER Manufacturing Productivity Database, Output-weighted aggregates of Bartelsman & Gray [1996] calculations for 4-digit industries. Moving averages are over three years.

Appendix

A Case Study: SPREADSHEETS

Historical Account

VisiCalc was the first electronic spreadsheet program, developed by Dan Bricklin and Bob Frankston and introduced in 1979. This software provided an interactive user interface for performing a wide variety of quantitative tasks without formal software programming. It was easy to use, inexpensive and ran on personal computers, first the Apple II and then others including the IBM PC in 1981. VisiCalc was credited with spurring sales of personal computers, becoming the first “killer application.”

Bricklin tried to sell or license the product to Apple Computer, but they were not interested. Instead, Bricklin and Frankston established Software Arts to develop their product, and they signed on with Personal Software Inc. (later Visicorp) to market the software.

VisiCalc was not patented. Bricklin and Frankston discussed patenting with their attorney who advised them they had little chance of getting a patent (this was two years before the Diamond v. Diehr decision, which threw out the main subject matter exclusion for software).

VisiCalc was imitated widely – by 1982 there were at least 18 spreadsheet products on the market. These products ran on a wide variety of personal computers and some were ported to larger computers. Nevertheless, VisiCalc remained the dominant product until 1983. None of the competing products were complete imitations; many had additional features. Some (such as PFS Plan) had fewer features, but were designed to be even easier to use.

In 1983 two products appeared that included several different types of software functions in one integrated suite. These were Context MBA and Lotus 1-2-3. Lotus, the more successful of the two, integrated a spreadsheet with a graphing program and a data management program. The data management functions allowed the user to create databases, this data could then be easily used in one or more spreadsheets. The graphs could be generated from the spreadsheets and printed with just a few keystrokes.

Although all of these features were available in separate products on the market, the efficient integration of these functions, especially in Lotus 1-2-3, provided much greater utility than could be realized with separate products. By the end of 1983, Lotus sales exceeded those of VisiCalc.

The developers of VisiCalc had difficulty producing a similar integrated product. The graphics routines were a critical component, and VisiCalc’s developers had little in-house experience with graphics programming. In fact, Mitch Kapor, one of the founders of Lotus, had worked for Personal Software (the company that distributed VisiCalc) and had developed separate programs that plotted graphs of VisiCalc files. VisiCalc ran into difficulties and Lotus purchased Bricklin and Frankston’s company in 1985.

Lotus dominated the spreadsheet market for a decade. Additional firms entered the market and, for the first time, some firms produced clones of Lotus 1-2-3: the programs did not have the identical underlying computer code, but they were designed to look and function just like the Lotus product. Lotus sued for copyright infringement and won. (A second case involved Borland's Quattro Pro, which did not look and function just like Lotus 1-2-3, but which used compatible menu sequences. After a protracted legal battle, Lotus lost this suit on appeal.)

In 1985, Microsoft introduced the Excel spreadsheet program for the Apple Macintosh. The Macintosh was the first personal computer to employ a graphical user interface (mouse, point-and-click, drag-and-drop, etc.). This interface made the program even easier to learn and use. Microsoft also used this new interface in its Word word processing program and PowerPoint, its slide presentation program. However, the Macintosh market was much smaller than the IBM-PC market.

In 1988, Microsoft introduced Excel to the much larger IBM-PC market, running on the first version of the Windows operating system. Then in 1991 it introduced Microsoft Office, an integrated package that included Excel, Word and PowerPoint. This product was designed to allow easy transfer of content between these programs. Meanwhile, Lotus had been slow to upgrade 1-2-3 with a graphical user interface and Lotus lagged on integrating its spreadsheet with other programs to the same degree. Sales of Microsoft's suite of products took off, and in 1993 Microsoft took the lead in the spreadsheet market.

This simplified account makes clear that important design elements included graphic displays and interfaces and integration of different functions. Also, the early players did not seem to appreciate the significance of these design elements sufficiently, suggesting the importance of complementarities in innovation. Kapor's experience with graphics technology gave Lotus an advantage in understanding the benefits of integrated graphics and also expertise in the design of an integrated graphics product. Microsoft's experience in the Macintosh market provided a greater appreciation of this interface and programming expertise with it.

Time Line

1979	VisiCalc	First electronic spreadsheet
1981-2	VisiCalc on IBM-PC	17+ competitors VisiCalc leads
1983	Lotus 1-2-3 & Context MBA	Integrated graphing and data management
1984		Lotus leads
1985	Microsoft Excel on Macintosh	Graphical User Interface
1988	Excel on IBM-PC	
1991	Microsoft Office	Integrated office functions
1993		Microsoft leads

Issues

Did the initial innovators have sufficient incentive?

Bricklin and Frankston, sold their company to Lotus and, apparantly, profited significantly despite some terrible business problems, including a falling out with the firm they chose to market their product. Kapor made out very well. IBM bought Lotus in 1995 for $3.5 billion (of course, it had other products then).

What if VisiCalc had had a patent?

If VisiCalc had had a patent but did not license it, it seems likely that the spreadsheet would not be as nearly well developed as it is today. (This view is widely held in the software industry.) VisiCalc did not have much appreciation of the value of integrated graphics (and certainly not graphical user interfaces).

The real question is whether VisiCalc would have provided an affordable license to Lotus. The license fee would presumably have been set to attempt to extract monopoly rents from Lotus. Note that at this time, profit margins were *very* high (Lotus's margin in 1983 was 51%) and consequently such royalties would also have to be very high (perhaps 50% or so). Kapor's business plan only estimated sales of $2-3 million for 1983 (it actually sold $53 million). A high royalty combined with the high development costs would almost certainly make Kapor's projection too unprofitable for venture capital investment. Hence Lotus could not have accepted a license that extracted such high rents. On the other hand, VisiCalc, not appreciating the potential complementary value of Lotus 1-2-3 (Kapor himself underestimated the value) would have had little reason to offer a lower royalty.

Thus although some licensing might have occurred, a VisiCalc patent might have prevented the crucial entry of new technology into the spreadsheet market. At the very least, this would have delayed the development of advanced spreadsheet programs; at worst, they may never have become so advanced.

Did competition expand the market?

Here are some very rough estimates of worldwide market size (variety of data sources):

	Units (1,000)	Retail Sales ($ million)	Avg. Retail Price ($)	Lotus share of units	Lotus share of $
1979	100	25	250	--	--
1983	1,000	300	300	20%	33%
1988	2,400	1,040	430	70%	70%
1991	3,800	1,340	350	60%	60%
1996	20,000	1,600	80	30%	15%

First, if nothing else, *price* competition dramatically expanded the unit sales. Microsoft's introduction of the Office suite was accompanied by a price war – suites of software were sold for the price of the individual products.

Second, the growth of the market cannot be separated from the growth of the PC market; it is impossible to separate cause from effect. That is, did more spreadsheets sell because there were more PCs or did more people buy PCs because they wanted better and cheaper spreadsheet programs? But undoubtedly, people are using spreadsheets today to do things they would not have thought about doing on them in 1984.

Third, even though Lotus "lost" the competition dramatically in the early 90's, I estimate its direct spreadsheet revenues in 1996 were about $120 million (slightly less than in 1984) and its unit shipments were a whopping 6 million. That is, it still had a very sizeable and potentially profitable business, although growth prospects were not great. In other words, Excel did not immediately replace Lotus 1-2-3.

Was competition sequential and complementary?

Competition in the spreadsheet market is an example of what has been called "feature wars." Companies competed to provide new and different applications for the software and to improve the quality and usability of the products. They achieved this by releasing new versions of software every two years or so that incorporated many new features. Companies were under intense pressure to successfully incorporate the most important new features in each release [Cusumano and Selby]. The major innovations mentioned above [integrated graphics, graphical user interface] were, in fact, effected through the addition of many detailed features. The extent of the feature growth can be seen in the size of the source code for Microsoft Excel: from Excel 3.0 (1990) to Excel 4.0 (1992) the source code increased 31% [Cusumano and Selby, p. 310].

This form of innovation was clearly sequential. It was also complementary: innovation in features improved product quality, providing greater value per customer, and also new applications, reaching more customers. Both increased the extent of the market. Reviews in computer publications provide evidence in the growth in product quality [Liebowitz and Margolis]. Also, new versions offered sufficiently enhanced value to induce a large portion of customers to upgrade. And the growth in the overall market was also large (see above). This growth was also due to the growing market for computers, but spreadsheets were a "killer application" that promoted the adoption of PCs.

The growth in features did not directly make previous versions obsolete, but instead the new versions delivered added value.

Sources

Value Line (Lotus)

http://www.spa.org/divisions/research/sales.asp

http://www.bricklin.com/visicalc.htm

http://www.kapor.com/homepages/mkapor/Bio299.html

http://www.lotus.com/home.nsf/welcome/corporate

Cusumano, Michael A. and Richard W. Selby. 1995. *Microsoft Secrets: How the world's most powerful software company creates technology, shapes markets and manages people.* New York: Simon and Schuster.

Liebowitz, Stan J. and Stephen E. Margolis, Winners, Losers and Microsoft: Competition and Antitrust in High Technology. Independent Institute. Oakland: 1999