Standing on the Shoulders of Giants: Cumulative Research and the Patent Law

Suzanne Scotchmer


Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR’s Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

*The Journal of Economic Perspectives* is published by American Economic Association. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/aea.html.


JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2001 JSTOR
Standing on the Shoulders of Giants: Cumulative Research and the Patent Law

Suzanne Scotchmer

Sir Isaac Newton himself acknowledged, “If I have seen far, it is by standing on the shoulders of giants.” Most innovators stand on the shoulders of giants, and never more so than in the current evolution of high technologies, where almost all technical progress builds on a foundation provided by earlier innovators. For example, most molecular biologists use the basic technique for inserting genes into bacteria that was pioneered by Herbert Boyer and Stanley Cohen in the early 1970s, and many use a technique for causing bacteria to express human proteins that was pioneered at Genentech. In pharmaceuticals, many drugs like insulin, antibiotics, and anti-clotting drugs have been progressively improved as later innovators bettered previous technologies. Computer text editors are similar to one another, as are computer spreadsheets, in large part because innovators have inspired each other. An early example of cumulative research was Eli Whitney’s cotton gin, which was quickly modified and improved by other innovators who seriously curtailed his profit.¹

¹Eli Whitney was very generous in disclosing details of his gin to other innovators, even beyond what was required by patent law. Other innovators patented improvements, but after much litigation the new patents were held to infringe Whitney’s underlying patent. Whitney and his partner did not recover sufficient damage awards to compensate them for their litigation and the time that it took to enforce the patent. For an extensive discussion, see Jeanette Mirsky and Allan Nevins, The World of Eli Whitney (Macmillan Co., 1952). Klemperer (1990) also discusses the cotton gin.

*Suzanne Scotchmer is Professor of Public Policy, University of California, Berkeley, California.*
Most economics literature on patenting and patent races has looked at innovations in isolation, without focusing on the externalities or spillovers that early innovators confer on later innovators. But the cumulative nature of research poses problems for the optimal design of patent law that are not addressed by that perspective. The challenge is to reward early innovators fully for the technological foundation they provide to later innovators, but to reward later innovators adequately for their improvements and new products as well. This paper investigates the use of patent protection and cooperative agreements among firms to protect incentives for cumulative research.

The Available Tools

The breadth of patent protection is a key consideration in the incentives to innovate. Patent applicants protect themselves against competition from derivative products by claiming broad protection. Patent law would provide no protection at all if it did not protect against trivial changes, like color or size. The allowable breadth of claims is determined by patent examiners and the judiciary. If broad protection is granted, then a derivative or second generation product will likely infringe the prior patent, so a license on the original patent is required to market it. If patent protection is narrow, then many derivative products and applications can be patented and marketed without infringement.\(^2\) We might be tempted to conclude that broad protection encourages firms to find fundamental technologies but discourages them from seeking out second-generation applications and derivative products. However, these two conclusions may be inconsistent, since proper incentives to find fundamental technologies may require that the first patent holder earn profit from the second generation products that follow. There will be no such profit if no second-generation products follow.

Patent protection would be an unnecessary policy tool if the government had the same information about the costs and benefits of individual research projects as firms have. In that case, the government could simply select the research projects that would be socially efficient and commission research from the lowest cost firms. However, the government will generally have less information than firms,\(^3\) and I will therefore assume that the length and breadth of patent protection and other aspects of the government’s policy toward R&D cannot depend on firms’ private information about their expected costs.\(^4\)

---

\(^2\)U.S. patent protection is broader than that in most other countries, particularly Japan, partly due to the “doctrine of equivalents,” which can broaden protection beyond the claims in the patent according to similarity of function.

\(^3\)Wright (1983) discusses the private information of firms as the main justification for patent protection, rather than using prizes or contracts as incentive instruments.

\(^4\)A policy of reimbursing costs to the successful innovator would not be adequate, since a project that was a “good bet” at the beginning might nevertheless fail.
Given that the length and breadth of patent protection cannot depend on the expected costs of an R&D project, the only way to ensure that firms undertake every research project that is efficient is to let the firms collect as revenue all the social value they create. Otherwise, some projects that are socially desirable will not be undertaken. If an innovation is a reduction in the cost of producing a good, then the social value is the saved costs. If the innovation is an improvement to a product, the social value is the difference in consumers’ willingnesses to pay for the improved and unimproved products. When research firms collect all the social value as profit, households still benefit, but in their capacity as shareholders rather than as consumers.

But there are at least two problems with allowing research firms to collect all the social surplus as profit (or as much as possible). First, strong patent protection leads to socially inefficient monopoly pricing. Second, firms in a patent race may overinvest in research if the patent is worth more than the (minimum) cost of achieving it (Louy, 1979). This problem is related to the problem of the commons: An increase in one firm’s rate of investment transfers some probability of becoming the patentholder from other firms to itself. Because of this transfer, all firms might overinvest. These points are well-recognized in the R&D literature.

When an initial innovation facilitates later ones, as is the case with basic research, another issue arises. Part of the first innovation’s social value is the boost it gives to later innovators, which can take at least three forms. If the second generation could not be developed without the first, then the social value of the first innovation includes the incremental social surplus provided by second generation products. If the first innovation merely reduces the cost of achieving the second innovation, then the cost reduction is part of the social surplus provided by the first innovation. And if the first innovation accelerates development of the second, but at the same cost, then its social value includes the value of getting the second innovation sooner.

Because of these externalities provided to later innovators, developing the first innovation may be efficient even if its expected cost exceeds its value as a stand-alone product. First innovators will have correct incentives to invest only if they receive some of the social surplus provided by second generation products. But at the same time, enough profit must be left for the second innovators so that they will invest if investing is efficient. This essay asks how close patent incentives can come to accomplishing that goal.

A premise in much of what follows is that firms other than the first innovator should participate in development of second generation products. Since the first innovator might not have expertise in all applications, more second generation products are likely to arise if more researchers have incen-

But Gandal and Scotchmer (1989) show that prior agreements among firms that would otherwise race can overcome the incentive for overinvestment, even if their research costs are private information. When such prior agreements are allowed, the firms will invest at the efficient rates if and only if the private value of the patent is equal to the social value.
tive to consider them. In this view, contrary to the premise of much of the patent race literature, creativity is largely serendipitous. Not every R&D firm sees the same opportunities for new products.

However, outside research firms can integrate with initial patent holders in at least two ways: the firms can form cooperative ventures to research and develop new products, and they can form licensing agreements after products have been developed and patents have been awarded. I will call these two types of contracts prior agreements and licenses, respectively. Prior agreements permit firms to share the costs, as well as the proceeds, of research. Licenses are negotiated after research costs are sunk and patents have been awarded. Both types of agreements can increase profit by improving efficiency and possibly by reducing product market competition. Although many authors have discussed cooperation in research, they have not focussed, as I will, on how the breadth of patent protection and cooperation among research firms work together in protecting incentives to innovate. In this view of how incentives to innovate are protected, a key role of patent protection is that it sets bargaining positions for the prior agreements and licenses that will form, and therefore determines the division of profit in these contracts.

**Patent Protection and Licensing**

A system of property rights that might seem natural would be to protect the first innovator so broadly that licensing is required from all second generation innovators who use the initial technology, whether in research or in production. But such broad protection can lead to deficient incentives to develop second generation products. When the licensing agreement is negotiated after a patent has been granted, research costs have already been sunk. The bargaining surplus to be split between the first and second innovators at that time is the incremental market value of the second product, not net of research costs. A second innovator who cannot market the next generation product without a license has a very weak bargaining position. If the second innovator does not get all the surplus being bargained over, he will earn only a fraction of the new product’s market value and presumably only a fraction of its social value, and this fraction may be less than the cost of developing it. Hence the incentive for an outside firm to develop second generation products can be too weak. Under such broad patent protection, the incentive for the first innovator to develop a second generation product will be stronger than for an

---

6Prior cooperation in research has been treated leniently by antitrust law and the authorities. For example, the National Cooperative Research Act (1984) established that joint ventures are not per se illegal, but will be treated according to a “rule of reason.” The Act also reduced damages in civil suits from treble to single, provided the firms follow the proper notification procedure. Perspectives on this act are discussed in the symposium on “Collaboration, Innovation and Antitrust” in the Summer 1990 issue of this journal.

7See Green and Scotchmer (1990) for an elaboration of this idea.
outside firm (provided the first innovator has expertise to develop the new product, and thinks of it), since the first innovator will earn the entire incremental profit.

As well as offering deficient incentives for second innovators, broad patent protection might inefficiently inflate incentives for the first innovation. In licensing agreements, the first innovator will earn a share of the market value of each infringing later product. If the first innovation reduces the cost of achieving later innovations, but is not the only possible vehicle to achieve them, the first innovator’s share should not exceed the cost reduction. If it does, the first innovator will be overrewarded.

In what follows I explore two solutions to these defective incentives. The remainder of this section investigates what happens if the first innovator’s patent protection is narrowed so that a different enough second generation product does not infringe and thus can be marketed without a license from the first innovator. In the following section, I investigate prior agreements in which second innovators can “sell” their ideas to the first innovator or integrate with the first innovator. Neither solution is perfect, as we shall see.

The inadequacies of narrowing patent protection are most easily exposed if we first suppose that first and second generation products do not compete in the market, although second generation products build on the first generation technology; for example, many new pharmaceuticals that are therapies for different illnesses all build on a few basic techniques of bio-engineering. Second innovators cannot have excessive incentive to invest, since they cannot earn more than consumers’ willingness to pay in the markets they serve. Licensing from the first innovator would transfer away some of the second innovators’ revenue and hence reduce their incentive to invest. To provide efficient incentives to the second innovator, society should protect the first innovation so narrowly that a new product never infringes and therefore second innovators never have to license. But such a scheme does not sufficiently reward the first innovator, since the first innovator does not profit from the cost reduction conferred on the second innovators.

The first innovator’s incentive to invest becomes still weaker under narrow patent protection if the second generation product is a substitute for the first. Competition between the two patent holders would erode their joint profit, transferring some of the social surplus of the combined innovations to consumers. As an example, suppose that the second generation product is a superior version of a drug, and that the two patent-holders compete on price. Then the second generation product will survive in the market and its price will equal the difference in consumers’ willingness to pay for the two drugs plus the marginal cost of producing the drug. In this outcome, the second innovator earns as profit exactly the incremental social value of the newer drug, while the first innovator’s profit falls to zero.

Such profit erosion could be mitigated if the antitrust authorities permitted collusive licensing among patent holders who would otherwise compete. For
example, licensing with per-unit royalties can lead to collusive outcomes, since the royalty raises the licensees’ private production cost and therefore keeps the equilibrium price high. In ordinary antitrust law, collusion through licensing would violate the spirit of the Sherman Act and subsequent legislation. But where incentives to innovate are at stake and where later technology builds on an earlier technology, such collusion allows the first innovator to profit from the externality conferred on later innovators. Of course, firms would be tempted to exploit any leniency by the antitrust authorities in contexts where incentives to innovate are not at stake. This problem should not be minimized.

There is something quite general economists can say about the combined effects of patent law with licensing: No such policy can achieve fully efficient incentives, even if society permits collusive licensing between patent holders who would otherwise compete and the firms jointly collect all the social surplus as profit. This is essentially because of “double marginalization.” To give the second innovator an incentive to invest whenever social benefits exceed R&D costs, the second innovator must earn the entire social surplus of his innovation. But to compensate the first innovator for the externality or spillover she provides, she too must earn part of this surplus. It is impossible to give the surplus to both parties.

When both first and second generation products are developed, the division of profit between the two innovators depends on the breadth of patent protection. To see this, assume that there is a random component to the outcome of a research project, so that when a research firm invests in a second generation product, it does not know whether its product will infringe the prior patent. The breadth of the prior patent determines the probability that the second generation product will infringe. If the second product turns out to infringe, the second innovator must license and this will force him to share the profit of the improvement with the first innovator. The second innovator is in a better position if its product turns out not to infringe, since the second innovator can profitably compete with the prior patent holder in the market. Thus, if breadth of the first patent could be interpreted to depend on the expected costs and benefits of a second generation product, we could ensure that the second innovator’s expected profit would be zero. If not, some second generation products will be stymied even though they would contribute positively to joint profit and to social welfare, and the second innovators who invest will typically make positive profit.

To summarize, the “natural” system of property rights—requiring every later innovator to license any underlying technology—will on average give deficient incentives for outside firms to develop second generation products.

---

8 If royalties are permitted, then licensing is similar to permitting the initial patent-holder to buy up the patents on later products that use the initial technology.
9 Green and Scotchmer (1990) argue this in a model where the second innovation would be impossible without the first innovation. It is also true in the less drastic case when the first innovation merely reduces the cost of achieving the second innovation.
This is because the second product infringes and therefore the second innovator must transfer some of the innovation’s revenue to the first innovator by licensing. If the first innovator can be relied upon to develop all second generation products, this would not matter. Second, no system of narrower patent protection and licensing can give the right incentives to both the first innovator and other firms that develop improvements, even if collusive licensing among noninfringing products were allowed. The latter result depends on my premise that the breadth of an underlying patent cannot be separately tailored to the costs and benefits of each second generation product.

In the next section, I ask to what extent these inadequacies of patent protection and licensing can be overcome with prior agreements reached before some or all the patents have been obtained. Incentives with licensing are defective mainly because firms negotiate after all costs have been sunk and patents have been issued. A prior agreement integrates the potential second innovator into the firm of the first innovator before investing in the second innovation. Such prior agreements can indeed guarantee efficient investment in second generation products, but cannot perfectly solve the incentive problem unless the negotiation is before all costs are sunk, including the costs of the first innovator, or unless the first innovator has all the bargaining power.

Prior Agreements

Prior agreements among research firms are often called research joint ventures. Joint ventures presumably form to increase the joint profit of the members, but they do not necessarily increase social welfare, since the cooperation is among firms only and does not include consumers. Joint ventures increase profit both by providing incentives to the members to invest more efficiently, and by finding ways to transfer social surplus from consumers to firms. The greater efficiency might result from exploiting economies of scale (Katz, 1986), from sharing technological know-how (Bhattacharya, Glazer and Sappington, 1988), or from undoing the inefficiencies of a patent race (Gandal and Scotchmer, 1989).

One solution to the incentive problem would be to integrate all possible innovators into one firm before even the first innovator has invested. Then, provided the integrated firm gets most of the social surplus from the joint innovations, it should invest (close to) efficiently. Although research firms do not know with certainty what projects they will think of after the first generation technology has been developed, they have expectations about the possible benefits and costs of such projects. Provided all researchers have similar expectations, an agreement negotiated before the first investment could ensure

---

10 The Coase theorem would conclude that bargains increase the joint welfare of all the parties. Thus, if consumers and firms could jointly cooperate, prior agreements would inevitably benefit both groups. But since some of the parties are excluded—namely consumers—there is no guarantee that prior agreements increase total welfare.
that the first innovation is undertaken if and only if efficient, where efficiency is defined relative to the prior judgments about costs and benefits. But the more serendipitous is the discovery of second generation products, and the more difficult it is to include all potential second innovators, the less feasible such an agreement seems. I therefore consider the more limited prospects when integration occurs after the first innovation. The difficulties in transferring profit to the first innovator are clearest if we assume that the innovators can jointly collect all the social surplus as profit, provided they do not compete in the market.

After the first patent has issued, a potential second innovator could approach the first patent-holder with an idea for an improvement or new product, and suggest that they share both the costs and proceeds of research. Such an agreement can increase joint profit by increasing investment in profitable second generation products and by preventing market competition among firms that would otherwise own competing patents. If patent protection is broad, without this prior agreement the second innovator could have deficient incentive to invest, as explained above. With a prior agreement, the initial patent-holder can agree to share both the costs and the proceeds of the second innovation, and will do so whenever benefits exceed costs.

Prior agreements are a social improvement over licensing because they can improve incentives to invest in second generation products, whatever the breadth of patent protection. With licensing, the breadth of patent protection serves two purposes: It determines investment in second generation products and determines how the firms’ joint expected profits will be divided. With prior agreements, the breadth of patent protection serves one purpose instead of two: The two innovators have an incentive to invest efficiently in second generation products whatever the breadth of patent protection. The breadth of protection determines only the bargaining positions, hence the division of profit.\(^\text{11}\)

Whether a prior agreement can provide efficient incentives for the first investment as well as the second depends on two factors: how much social surplus must be transferred from the second innovator to the first (how big the externality is), and the second innovator’s bargaining power. A second innovator who has a strong bargaining position will earn positive profit in a prior agreement, thus limiting how much social surplus the first innovator can collect. The second innovator’s bargaining position is strongest if there is a high probability the second innovation would not infringe and if the second generation product is itself patentable. The second innovator will also have a strong bargaining position if no other firm is capable of developing the second generation product.

\(^\text{11}\)On the other hand, approaching the first innovator with the idea for the second innovation might give away the idea of the potential second innovator, and thereby undermine its bargaining position. The law has remedies for this problem, but presumably they do not work perfectly.
To bid the second innovators’ profits to zero, it is not necessarily enough that each potential second innovator has many competitors. Suppose the second generation product is itself patentable and the first innovator agrees with one potential second innovator that that firm will develop the improved product and have an exclusive license on the initial patent. Despite this agreement, the parties will be forced to negotiate with any other firm that gets the improvement first since they could not duplicate the second invention without infringement. Anticipating this, potential second innovators might not bid away their profit in trying to make an agreement with the first innovator (Scotchmer, 1990).

Some authors who have written on research joint ventures have pointed out that firms’ incentives to cooperate at the research stage depend partly on whether the members can also collude in using the resulting patents (Ordover and Baumol, 1985; Katz, 1986; d’Aspremont and Jacquemin, 1988; Choi, 1989). My discussion has assumed they can, and that such collusion is an important way to protect incentives to innovate. It would be difficult to implement a rule that permitted the consolidation of property rights through a prior agreement only if the eventual patents would otherwise be infringing. When research outcomes are random, patent authorities cannot know whether an R&D project will result in new technology that infringes an earlier patent.

Conclusion

A main conclusion of this essay is that it is misleading to ask how broad patent protection should be without simultaneously asking whether research firms can integrate or otherwise cooperate. Similarly, it is misleading to ask how leniently society should treat cooperation among research firms without simultaneously asking how broad patent protection should be. Prior agreements have the advantage of leading to more efficient investment in second-generation products. But if the patent authorities disallow collusive licensing of patents that are noninfringing, they might also want to disallow prior agreements. A prior agreement merges the interests of the two firms whether or not the patents would be noninfringing, and could circumvent the antitrust authority’s desire to force competition. Of course, competition has a disadvantage where incentives to innovate are at stake: It erodes the joint profit of the firms, which undermines at least one firm’s incentive to invest in R&D when such investment is efficient.

There are no simple conclusions to draw about the optimal breadth of patents. It is not necessarily optimal to protect the first innovation so broadly that every derivative or second generation product infringes. If prior agreements are disallowed or ineffective for some reason, then broad patent protection could discourage the development of second generation products, as
explained above. And if the first innovator does not expect to profit by licensing to second generation innovators, broad protection could inhibit the first innovation as well, thus undermining the entire research line.

Broad protection might also be undesirable when prior agreements are allowed. To encourage researchers to invest in second generation products, the first innovator might have to make prior agreements with firms that have bargaining power. Their bargaining power derives from the fact that, without an agreement, they might have a credible threat not to invest, and from patentability of the second generation product. Because of their bargaining power, they may get a share of the bargaining surplus. Suppose patent protection is narrowed enough so that second innovators will invest without a prior agreement, but their profit is kept low. This would increase the first patent-holder’s profit, as may be necessary to compensate the first innovator for the externality or spillover conferred on second innovators.\textsuperscript{12}

Patent law is limited in its instruments: the main ones are the patent life and the breadth of protection.\textsuperscript{13} The private value of patent protection is linked to the social value of the technology through market demand, but is not linked to firms’ research costs. The optimal rule for the breadth of a patent can only use information that is available to patent examiners and courts.\textsuperscript{14} Thus, the patenting rule can depend on observable aspects of discovered technologies, but not on prior expectations regarding technological outcomes and costs of research. This restriction greatly reduces the effectiveness of patent law in protecting incentives.

Before investing in a second generation technology, the researcher must evaluate the probability that the new technology will not infringe the prior patent. This probability depends on the breadth of the prior patent and on the distribution of possible outcomes of the second investment. The probability of infringing the first patent is lower if the distribution of outcomes places greater weight on outcomes that lie outside the allowed claims of the first patent. A project with low probability of infringing will look more profitable to the second generation researcher than a project that places greater prior weight on outcomes that infringe the first patent. But both projects could have the same expected social value if the expected costs of one project were sufficiently higher than the other. The patent policy should equally encourage two projects

\textsuperscript{12}See Green and Scotchmer (1990) for an elaboration of this argument.
\textsuperscript{13}Of course, details like priority rules also matter. Everywhere except the United States, a disputed patent issues to the first applicant. In the United States, a disputed patent issues to the first inventor, regardless of when application occurs. See Scotchmer and Green (1990) for a discussion of the incentive effects of these two rules. The anomalous American rule is now being reconsidered.
\textsuperscript{14}Except for Wright (1983), authors have not focussed explicitly on what information is available to patent authorities. Gandal and Scotchmer (1989), Green and Scotchmer (1990), Klemperer (1990) and Scotchmer and Green (1990) assume that patent protection cannot depend on costs. Klemperer (1990) and Gilbert and Shapiro (1990) assume that the breadth of patent protection can depend on aspects of market demand. Green and Scotchmer (1990) assume that the breadth of patent protection can depend on the level of previous technical advance.
with the same social value, but that cannot be accomplished with a patent rule that depends only on technological outcomes.

A disadvantage of narrow patent protection that I have not yet discussed is that it might discourage first innovators from patenting and disclosing their technologies. Patent law requires disclosure for the same reason that innovators dislike it: it is the vehicle by which technical knowledge is passed from the patenting firm to its competitors.\textsuperscript{15} The first innovator would rather develop second generation products than let other firms develop them, since that would be more profitable. As a consequence, the first innovator has a strategic incentive not to patent the innovation. Instead, the first innovator could hold the product off the market until it develops the more valuable second generation products, or it could market the first product and rely for protection on the law of trade secrets, which does not require disclosure, but also does not protect against independent invention.

The first innovator’s incentive to patent the initial technology depends on: (i) the profitability of marketing the first technology prior to the development of second generation products; (ii) the extent of disclosure that patenting entails;\textsuperscript{16} (iii) the ease with which the technology could be reverse-engineered if marketed but not patented; and (iv) the breadth of patent protection. The incentive not to patent is especially strong when patent protection is narrow, since a second generation product is then more likely to damage the first innovator’s profit.

The problem of cumulative research is especially acute when the first technology has very little value on its own, but is a foundation for valuable second generation technologies. Even with licensing, the first innovator might not capture the full social value that it facilitates and may have deficient incentive to invest. This is presumably why governments fund basic research. The branches of government that fund research are not those that set patent policy, and the decision to support basic research might be interpreted as a recognition that patents and licensing are inadequate.

Governments have taken different views of whether publicly sponsored research should also be patentable. The U.S. government permits and even encourages patenting of results from government sponsored research; for example, the Boyer-Cohen patent. In contrast, the British government forbade the Cambridge Molecular Biology Lab from patenting monoclonal antibodies.

\textsuperscript{15}The disclosure requirement in section 112 of the patent law states that “the specification shall contain a written description . . . in such full, clear, concise and exact terms as to enable any person skilled in the art . . . to make and use the same . . .” Scotchmer and Green (1990) show that a leading firm might not want to patent a patentable technology even if this means holding it off the market until the next innovation in order to avoid reverse engineering.

\textsuperscript{16}Disclosure of some technologies, such as chemical compositions, teaches competitors much that is valuable. For example, after the basic material of superconductors was disclosed, many other researchers developed aspects of it (\textit{The New York Times}, January 2, 1989, p. 34). Disclosure of other technologies, like bio-engineered proteins, gives away much less that is useful to competitors, and therefore innovators will not fear disclosure as much.
in the mid-1970s. Permitting patents on government sponsored research rewards successful innovators twice, once through government funding and again through the patents.

I have focussed in this essay on how to divide joint profit among innovators when one innovator’s technology builds on another’s. I have not focussed on the length of patent protection as a policy tool. My simplifying assumption that the innovators can jointly get all the social surplus as profit if they do not compete in the market essentially means that the patent continues forever. When owners of noninfringing patents cannot collude through licensing, a noninfringing second generation product will undermine the profitability of the first patent, and this may happen before the end of the patent life. Mansfield (1984) surveyed research firms and found that the effective lives of most patents are much shorter than the stipulated 17 years. In the extreme case where a second product is an improved version of the first, the effective length of the first patent may simply end when the second product is introduced. The effective patent life is determined directly by its breadth, since the breadth determines how long it takes until this happens. On the other hand, when second generation products serve different markets, as when consumers have different tastes and a second generation product is a variant that serves a different set of consumers, length and breadth can be chosen independently, although the monopoly power conferred by a patent may be eroded over the patent’s life as similar but noninfringing products are invented.

Gilbert and Shapiro (1990) and Klemperer (1990) have discussed breadth of patent protection in the context of single innovations without focussing on the problem of dividing profit among sequential innovators. In their conceptualization, breadth and length are substitute ways to provide a fixed profit to a single innovator, and they ask what combination of length and breadth minimizes the cost of monopoly distortions. They do not ask how large the fixed profit should be and therefore do not focus on how to preserve incentives to innovate. The prescription for patent breadth that comes out of that perspective may well conflict with the prescription for patent breadth that comes from considering incentives to innovate when research is cumulative.

It appears that patent policy is a very blunt instrument trying to solve a very delicate problem. Its bluntness derives largely from the narrowness of what patent breadth can depend on, namely the realized values of the technologies. As a consequence, the prospects for fine-tuning the patent system seem limited, which may be an argument for more public sponsorship of basic research.

Prepared for the Symposium on Intellectual Property Law, funded by the RAND Corporation and the John P. Olin Foundation, which took place in Washington D.C. on October 24, 1989. I thank Stan Besen, Joe Farrell, Jerry Green, Leo Raskind, Steve Salop, Carl Shapiro, Eugene Smolensky, Joe Stiglitz and Timothy Taylor for useful comments. I thank Lucette Decorde for her able research assistance and the NSF, Grant SES 89 09503, for financial support.
References


