RAND Journal of Economics
Vol. 21, No. 1, Spring 1990

Novelty and disclosure in patent law

Suzanne Scotchmer*

and

Jerry Green**

The stringency of the novelty requirement in patent law affects the pace of innovation because it affects the amount of technical information that is disclosed among firms. It also affects ex ante *profitability of research. We compare weak and strong novelty requirements from the standpoint of social efficiency. We ask how our answer depends on the rule that determines which firm gets a patent when two firms have patents pending on the same technology. The possible rules are "first-to-invent," which applies in the U.S., and "first-to-file," which applies everywhere else.*

1. Introduction

■ The value of a patent is determined by the competitive advantage it confers and by the period of time during which this advantage exists. This interval is rarely the full 17 year *de jure* life of the patent, particularly in industries with rapidly evolving technologies, where the effective life of the patent is the time until it is superseded by a superior technology.[1] The legal requirements of "novelty" and "nonobviousness" defined in Sections 102 and 103 of the patent code determine how broad the claims of a patent can be, and therefore determine how different a subsequent innovation must be in order not to infringe.[2] They therefore determine the value of patent protection, the incentives for research, and how much technical information is shared among firms through patenting.

The requirements of novelty and nonobviousness are hard to interpret. They are judicially determined standards, administered by the patent office and litigated in Federal courts. Inevitably, there is room for discretion in determining whether a new technology is

*University of California, Berkeley.
**Harvard University.
The authors thank Joseph Farrell, Kenneth Judd, Michael Katz, Nancy Gallini, Richard Gilbert, David Newbery, Garth Saloner, and Carl Shapiro for useful discussion. We thank Jay Choi for pointing out many (we hope all) of our mistakes. Research was supported by the National Science Foundation, grant SES 88 09107, and the Hoover Institution.

[1] Mansfield (1984) reports that, in a survey of R&D firms that he undertook, about 60 percent of patented products were successfully imitated within four years of patenting.

[2] These are separate requirements in the United States Code, but in practice they are often difficult to distinguish. Products of nature fail nonobviousness, although there are nuances regarding what a product nature is. For example, a purified natural substance, whose useful properties depend on purity, might not fail nonobviousness.

different enough not to infringe a previous patent.[3] One of the two principal policy questions addressed in this article is how stringent this requirement should be.

The reason to grant patent protection that has been emphasized in the literature is that it creates incentives to do research. The social goal of protecting profit is served by a strong novelty requirement, which we interpret to mean that small derivative improvements will infringe a prior patent. Thus, the patent is likely to have a long effective life before a sufficiently different technology supplants it.

A second reason to grant patent protection is to accelerate aggregate innovation through disclosure of inventions.[4] The disclosure requirement for patentability, delineated in Section 112 of the patent code, 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. . . ." If each small technological advance were disclosed, as would be encouraged by a weak novelty requirement, the shared technical knowledge would help other innovators in their own research, reduce redundancy, and hasten the time to subsequent innovation. The social goal of disclosure is served by a weak novelty requirement.

A consideration in balancing these two arguments is that firms might not patent or market every small technical advance, even if the novelty requirement is weak. Disclosing technical information confers a positive externality on a firm's competitors, which the firm might want to avoid.[5] Firms might therefore suppress small technical advances, and a weak novelty requirement might not have the desired effect of encouraging disclosure. Firms' strategic disinclination to patent or otherwise disclose small improvements could mitigate the erosion of *ex ante* profit that might otherwise follow from a weak novelty requirement, and the possibility of competition between close substitutes.

To study the pace of aggregate innovation under weak and strong novelty requirements, we need a dynamic model in which the decisions of what to patent and whether to enter or exit a race are made with foresight and in the knowledge that the subsequent decisions of other players will be similarly rational. The *ex ante* profitability of research and the strategic decision about whether to market or patent an innovation cannot be separated from each other. A firm might suppress an innovation that does not infringe a prior patent in order to retain a competitive advantage in the race. Such a decision can be anticipated when the firm decides to enter the race at the beginning or to continue later. This article develops a simple dynamic model that allows us to focus on how the novelty requirement affects the pace of innovation.

---

[3] The importance of, and confusion regarding, the novelty requirement was illustrated by cases in the 1980s concerning genetically engineered products and microorganisms. Several pharmaceutical companies litigated the question of whether a protein produced with recombinant DNA infringed a patent on the same protein produced synthetically. (See Ruby Baum, *Chemical and Engineering News*, July 20, 1987, p. 11+.) The same issues arose in the development of computer software, which is covered by copyright law rather than patent law. Computer companies sometimes build software systems that use the architecture of previous systems, e.g., spread sheets and menu-driven text editors. Should the initial architect's property rights extend to all subsequent software that is similar in its architecture?

[4] Most of the economics literature on patenting has assumed that the pace of innovations is determined only by each firm's rate of investment, which may be inefficiently high due to racing, and by the number of firms that enter the race.

[5] Externalities are fundamental to commercial research, just as they are fundamental to academic research. In the computer software example, the externality is that a subsequent software designer does not need to reinvent the architecture. In bioengineering, the technology for inserting foreign genes into bacteria (the Cohen-Boyer patent issued to the University of California at San Francisco and Stanford University in 1980) underlies many subsequent innovations. For example, in 1988 Genentech received a patent on a method of triggering human genes to express human proteins in bacteria (the Itakura-Riggs patent), once the genes have been inserted. The latter technology could not have been invented so soon or so cheaply without disclosure of the prior technology, as the prior technology would have had to have been invented first.

Section 2 discusses the externality that disclosure of one firm's progress confers on other firms. With disclosure, other firms do not have to duplicate progress; they can immediately build upon the progress made elsewhere. In Sections 3 and 4 we present an extensive form game that models three decisions made by firms as the race unfolds. First, there is the initial decision to begin the race, which depends on whether R&D is expected to be profitable. Second, if the novelty requirement is weak so that a small technical advance does not infringe a previous patent, the inventor of a small advance must decide whether to disclose it by marketing or patenting it. There is a tradeoff between the profit of marketing the small advance and the value of maintaining a competitive advantage in technical knowledge for later stages of the race.[6] The third decision is that the lagging firm may drop out of the race when the first innovator has not disclosed its technological advance, since it is unlikely that the lagging firm will catch up.[7] An important reason why an early innovator might not patent the first advance is that it might be able to force a shakeout in the industry by sending a credible signal that it has innovated, but not patented.[8]

A firm's strategic decisions about whether to disclose or suppress an innovation and whether to stay in a race depend on a legal issue surrounding the conditions of priority under which a patent can be granted. Suppose a firm has suppressed a patentable innovation that does not infringe a previous patent, and then a second firm duplicates the technology and tries to patent it. Now the first innovator no longer has a motive to suppress, and a dispute may arise as both firms seek patents for the same innovation. There are two dispute resolution rules, called "first-to-file" and "first-to-invent," that may determine which firm is granted the patent. The first-to-file rule, which applies in all countries except the U.S., means that the patent issues to the first applicant independently of priority in discovery. The first-to-invent rule, which applies in the U.S., means that the patent will issue to the first inventor, provided the date of first invention can be documented.[9] In Section 5 we compare the efficiency of these two dispute resolution rules.

Two principal conclusions follow from the model we present. First, each firm's strategic disinclination to disclose a small improvement can avoid the *ex ante* profit erosion that might otherwise result from a weak novelty requirement. For no parameter values in our model is *ex ante* profit negative with the weak novelty requirement when it would be positive with the strong novelty requirement. The only sense in which the strong novelty requirement can be socially preferred to the weak one is that, for some parameter values, it rectifies incorrect incentives for a firm to drop out of the race when it is technologically behind. This incentive will be too weak, relative to efficiency, when the dispute resolution rule is first-to-file and will be too strong, relative to efficiency, when the dispute resolution rule is first-to-invent.

Second, the first-to-invent rule discourages disclosure, relative to first-to-file. With first-to-invent, a first innovator does not need to patent in order to keep a claim on the market.

[6] Our explanation for why an innovator may fail to disclose (i.e., to protect its competitive advantage for winning a later race) differs from that offered by Horstmann, MacDonald and Slivinski (1985), who observe that an innovator may want to avoid imitation in the product market itself.

[7] Grossman and Shapiro (1986) study a race similar to ours and also observe that firms may drop out.

[8] In order for the signal to be credible, a firm that has not innovated must find the signal too costly to send. For example, the signal might be an invitation to venture capitalists. Fraudulent announcements that attract venture capital make firms liable to criminal prosecution. Shareholders sued Genentech in 1988 for misrepresenting its patent position and market prospects regarding TPA.

[9] The anomalous American rule is currently being reconsidered by U.S. negotiators in talks organized by the World Intellectual Property Organization. (*Wall Street Journal*, June 15, 1988.) In 1987, a dispute between the Japanese firm Sankyo and the American firm Merck regarding an anticholesterol drug was settled differently in Japan and the U.S. because of this difference in rules. While Sankyo's application predated Merck's in both countries, the patent issued to Merck in the U.S. because Merck could document prior invention. A further description of this case is given by A. Yoshikawa (1987).

If the competitor catches up and attempts to patent, the first innovator will successfully counterpatent, unlike in the first-to-file system where the counterpatent would be unsuccessful.[10] And there might be an advantage in not patenting initially in that no information is disclosed.

First-to-file engenders more disclosure than first-to-invent, but it also creates excessive incentives for firms to stay in the race. In contrast, the incentives to remain in the race under first-to-invent are weaker than those under first-to-file; indeed, first-to-invent can sometimes encourage firms to drop out of the race when it is socially efficient, whereas first-to-file would have induced them to stay in.

2. The social value of disclosure

■ We consider a model that isolates the information externality and the effect of the novelty requirement on the pace of innovation. For simplicity, we assume that there are only two firms racing, that there is a maximum rate of investment fixed by each firm's personnel and technology, and that firms are equally able. Each firm has a flow variable cost, c, per unit time, of doing research, so that firm i's flow cost of doing research is c if it invests full time. In the game we specify below, a firm can always increase its expected payoff by increasing the fraction of time that it invests, provided its expected payoff is positive. Therefore, in equilibrium, each firm will invest at rate zero or one. Since efficiency also demands that each firm invest at rate zero or one, we can restrict the rates of investment to be zero or one without loss of generality.[11]

We study the simplest case, in which there is a base-level of technological knowledge, beyond which there are two possible innovations. We assume that the first innovation adds one unit of social value to the value of the base-level. The second possible innovation adds an additional unit of value. These are flow values that go on forever. The marginal cost of producing each product is zero, and the size of the market is one.

We say that the novelty requirement is "strong" if the more advanced technology, which has added social value of two, can be patented and marketed without infringing a patent on a base technology, but the smaller innovation would infringe and cannot be patented. We say the novelty requirement is "weak" if the smaller improvement can be patented and does not infringe the base technology.

We shall use the Poisson discovery process (or exponential distribution of waiting time) to model R&D.[12] We assume that the smaller innovation requires one Poisson hit and the more advanced technology requires two Poisson hits. But if one firm achieves the first innovation (the first "hit") and discloses it to the other firm through patenting or marketing, then the other firm requires only one additional hit in order to have the more advanced technology. Thus, it is cheaper to achieve the more advanced technology if something has been learned from the previous technology. The first innovator can reduce the competitor's expected cost of achieving the more advanced technology by disclosing the first technology.[13]

[10] If the invention had previously been made public, the inventor would have a year in which to file the patent application in the U.S. Elsewhere, there is no grace period. Prior publication bars patentability. But the prior-publication restriction applies to both inventors, irrespective of who publicized the invention. With prior publication more than a year before the first application, no patent could issue.

[11] If a research firm could invest at a rate greater than one, it would invest at an infinite rate, and all R&D would be compressed into an instant. This is equivalent to assuming that the interest rate is zero.

[12] This is for simplicity. The Poisson discovery process has been used extensively to study patent races in which discovery has a random component. See, for example, Grossman and Shapiro (1986).

[13] One possible interpretation of this model is that the more advanced technology is comprised of the less advanced technology plus an "add-on." Here we prefer to interpret the decreased cost of the second step as a consequence of a pure information externality, rather than interpreting the second step as an add-on. Even though the most improved technology requires two technological advances, the composite technology is distinct from the previous technology.

Thus, disclosure of the first innovation is socially valuable because it accelerates discovery of the second innovation and reduces the aggregate cost by shortening the time during which firms invest. The distribution of waiting time until the first innovation has mean $1/(2\lambda)$, where λ is each firm's Poisson hit rate. The expected time until the second innovation differs according to whether the first technology is disclosed. If the first technology is disclosed, both firms race symmetrically for the next innovation, and the expected time to the next innovation is again $1/(2\lambda)$. If the first innovation is not disclosed, the competitor needs two Poisson hits to achieve the final improved product. Instead of $1/(2\lambda)$, the expected time until the final innovation is the expected value of min (w_A, w_B), where w_A is the length of time until the first innovator A makes a second hit, and w_B is the length of time until the competitor B makes two hits. Since w_A has a Poisson distribution with hit rate λ, and w_B has a gamma distribution with parameters $(2, \lambda)$, where 2 represents the number of hits required and λ represents the hit rate, it can be verified that the expected value of min (w_A, w_B) is $3/(4\lambda)$. Thus, the expected time until the final patent is achieved is $1/\lambda$ when the first innovation is disclosed, as is possible with the weak requirement, and $5/(4\lambda)$ with the strong requirement.

It is easy to show that if $c < (\lambda/r)$, the first best requires that both firms invest full time and that the first innovation be disclosed to the other firm.

While the value of disclosure suggests that a weak novelty requirement is best, this intuition might be invalid for two reasons that we shall explore in the equilibria discussed below. First, with the weak novelty requirement, the first innovator may fail to disclose. This is because disclosing the information reduces the probability that the first innovator will achieve the final patent. Since the final patent has a longer effective life than the interim patent, the innovator may want to preserve the initial innovation as a competitive advantage, rather than earn interim profits and simultaneously make the competition more potent. As a consequence, there may be no effective difference between the weak and strong requirements.

Second, *ex ante* profit might be larger with the strong novelty requirement than with the weak one. The weak novelty requirement leads to competition between close substitutes—either competition between the first innovation and the base technology, or between the more advanced second innovation and the first innovation. With the strong novelty requirement, there can only be competition between the most advanced technology and the base-level technology, and therefore the innovator is assured of a higher profit flow. The strong novelty requirement might therefore be socially better than the weak requirement, as it might induce entry into the R&D race when the weak requirement would not.

3. Patenting and disclosure in first to file

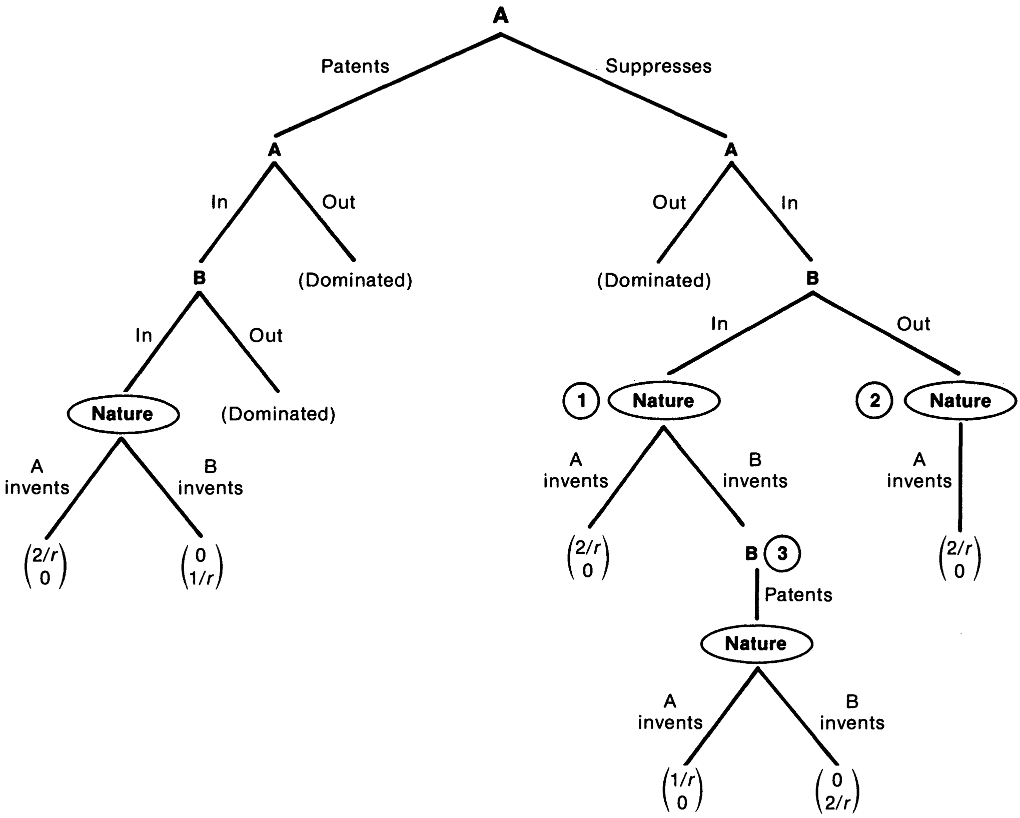
■ We first characterize equilibria of the dynamic games induced by the first-to-file patent rule under a weak novelty requirement and a strong novelty requirement, respectively. Having characterized the equilibria for the weak and strong novelty requirements, we compare the social surplusses that these rules generate.

For simplicity, we assume that once the innovations have been made, firms compete on price, and all consumers (of mass one) have the same willingness to pay for the product. Thus, if the first innovator competes with the base technology, it will earn a flow profit of one, and the owner of the base technology will be forced out of the market. This flow profit of one is the social value of the first innovation during the period of market incumbency. If the first innovator also achieves the second and final innovation, then its profit per unit time is two, which is again the cumulative social surplus created. If different firms market the first and second innovations, then the first innovation is forced out of the market and the superior product earns one unit of profit per unit time.

Generically there are only equilibria in pure strategies. That is, at each point in time

FIGURE 1

THE PATENT GAME FOR FIRST-TO-FILE



a firm invests at rate one or zero with probability either one or zero, and it is not equally profitable to patent or to suppress.

We first describe equilibria of the game with the weak novelty requirement. Assuming that firm A has made the first discovery, we are at the initial node of the subgame depicted in Figure 1. We assume, for simplicity, that A cannot market its discovery without disclosure, e.g., because the technology could be reverse engineered. Unless the innovation is patented, it will therefore not earn a flow of profit. The payoffs given at the bottom of the tree in Figure 1 are not the entire payoffs from having played. Each is the discounted value of future profits, valued at the final discovery date.[14] In order to determine the equilibrium decision whether to enter, we must compute the *ex ante* payoffs. Thus, we must discount the payoffs from the terminal nodes back to the beginning of the game from the (random) date of final discovery, add any profits from sale of the good during the interim stage, and subtract all costs of R&D until the firm ceases R&D.

In the game represented by Figure 1 we have assumed for simplicity that the firms make one-time decisions whether to invest between discoveries and patenting. But the following dominance arguments suggest that a rigorous formulation of our game, in which firms could come in and out of the race in different time periods between discoveries, would give the same equilibria.

[14] For ease of exposition, we assume that the patent lasts forever. The spirit of our results does not depend on this.

The only subgame equilibrium in the left branch of Figure 1, after A has patented, is for both firms to stay in continuously, as the following argument demonstrates. First we show that it is a dominant strategy for A to invest at each moment of time, which we model with the "infinitessimal" dt. We assume that $x \geq c$, where $x = \lambda/r$.[15] In the time period dt, A earns a flow profit of $1\,dt$, and has a probability of λdt of achieving the final patent, worth $(2/r)$. During the time period dt, firm A pays costs $c\,dt$. If B invests also, there is a probability $(1 - 2\lambda dt)$ of continuing, and there is a probability λdt that B innovates and terminates A's profit. Firm A should invest in time period dt if

$$1\,dt + (2x - c)dt + (1 - 2\lambda dt)P_A e^{-rdt} \geq 1\,dt + (1 - \lambda dt)P_A e^{-rdt},$$

where P_A is A's continuation value if neither firm innovates in time period dt. The probability that both firms innovate in period dt is $\lambda^2(dt)^2$, but terms that involve the square of dt will drop out of the inequality when we divide by dt and then let dt go to zero. Therefore we have excluded it. The right-hand side of the inequality is A's expected future profit if A does not invest in period dt and B does. In that case, the probability of continuing is $(1 - \lambda dt)$. The inequality becomes $(2x - c) \geq xrP_A$ when we divide by dt and then let dt go to zero. Now suppose that B does not invest in period dt. Then the relevant inequality is $1\,dt + (2x - c)dt + (1 - \lambda dt)P_A e^{-rdt} \geq 1\,dt + P_A e^{-rdt}$, which reduces to the same inequality $(2x - c) \geq xrP_A$.

The strategies that maximize the continuation value to A are for A to invest continuously after dt and for B not to invest. Then A earns a flow value of one forever, worth $(1/r)$, and earns an additional $(1/r)$ from the time of the next innovation. The value of the additional $(1/r)$, discounted to the present, is $(1/r)e^{-rz}$, where z is the date of the next innovation. Taking the expected value, where z is exponentially distributed with parameter λ, we get an additional value of $(x/(r(x + 1))$. Until the next innovation, A must invest at a flow cost of c. The present value of the cost in time period dt is $ce^{-rt}dt$. This flow will terminate at a date z which has exponential distribution with parameter λ, and the discounted expected costs are therefore $c/(r(x + 1))$. Thus, $P_A \leq (1/r)(1 + ((x - c)/(x + 1)))$, and the inequality is satisfied for the upper bound on P_A. It is a dominant strategy for A to invest in time period dt.

A similar argument shows that it is also a dominant strategy for B to invest in time period dt if A has patented the first innovation, provided $x \geq c$.

In the right branch of the game tree depicted by Figure 1, A suppresses the first innovation. We have simplified the game tree by assuming that B will patent the first innovation at Node 3 if B catches up. B loses nothing by patenting, since A already has the information that is thereby disclosed. B will gain the interim flow profits on the first innovation and has no reason not to patent. Further, we have just shown that if the first innovation has been patented, both firms (where A's and B's roles are reversed) will stay in. To simplify, we have suppressed the firms' decision to stay in at Node 3.

Following the same line of proof as in the left branch of the tree, where the first innovator patents, it is a dominant strategy for A to invest in each time period dt, given that A has suppressed the first innovation and that $x - c \geq 0$. Thus the only in/out decision that depends on the parameter values (c, λ, r) or on (x, c) is whether B will stay in or drop out if A has suppressed the initial innovation. The other equilibrium strategies that depend on the parameters (x, c) are (i) whether A will patent or suppress the initial innovation, and (ii) whether firms will enter *ex ante* at the beginning of the race.

Table 1 shows A's and B's continuation payoffs in the left branch of the tree where A patents and both stay in. At each time period dt until the next innovation, A earns a flow

[15] Since expected profit, before entering at the beginning of the game, cannot be greater than expected social surplus, expected profit is negative if $x < c$. Thus firms will not enter initially.

TABLE 1 **Continuation Payoffs if A Patents the First Innovation: Both Firms Continue to Invest**

Payoff to A	Payoff to B
$\left[\frac{1}{r}\right]\left[1 - \frac{c}{(2x+1)}\right]$	$\frac{(x-c)}{r(2x+1)}$

revenue of one and incurs a flow cost of c. The present value of this revenue and cost is $(1-c)e^{-rt}dt$. This flow will terminate at a date z which has exponential distribution with parameter 2λ, and therefore has present expected value $(1-c)/(r(2x+1))$. At this same termination date, with probability one-half, A will win a patent on the advanced technology which has worth $(2/r)$. The expected value of $(1/2)(2/r)e^{-rz}$ is $2x/(r(2x+1))$ when z is exponentially distributed with parameter 2λ. Thus, the continuation value to A in the left branch of the tree is $(1/r)(1-(c/(r(2x+1))))$. On the other hand, B wins the final innovation with probability one-half. If B wins, the value of the final innovation to B is only $(1/r)$, since B must compete with A's patent on the first innovation. Reasoning analogously, B's continuation value in the left branch of the tree is $(x-c)/(r(2x+1))$.

We now characterize the parameter values for which B will invest after A has suppressed an initial innovation. Table 2 gives the continuation payoffs to firms A and B calculated at Nodes 1 and 2. At Node 2, B's continuation payoff is zero because B drops out. Using the above reasoning, A's continuation value is $(2x-c)/(r(x+1))$.

Now consider Node 1. The date of the next innovation is exponentially distributed with parameter 2λ. Reasoning as before, each firm's discounted expected costs until the next innovation are $c/(r(2x+1))$. There is probability one-half that A is the next innovator, in which case A receives a property right worth $(2/r)$. With probability one-half the firms reach node 3, and at that date their expected future payoffs are as given in Table 1 with the identities of the firms reversed. (B, rather than A, earns a flow of profit on the first innovation.) Each of these payoffs must be weighted by one-half and discounted from the date of the next discovery. Taking the expected discounted value gives the payoffs in Table 2.

The equilibria are summarized in Figure 2. The notation "Weak: Enter/Suppress/Out" means that, under the weak requirement, firms will enter initially, then the first innovator will suppress rather than patent, and finally the competitor will drop out while the first innovator continues to invest. In the description "Weak: Enter/Suppress/Out," we again ignore firm A's decision to stay in when A has suppressed, since that is a dominant strategy. Similarly, we have written only "Weak: Enter/Patent," since both firms always stay in after patenting. (Figure 2 also summarizes equilibria of the game with the strong novelty requirement, which we shall discuss below.)

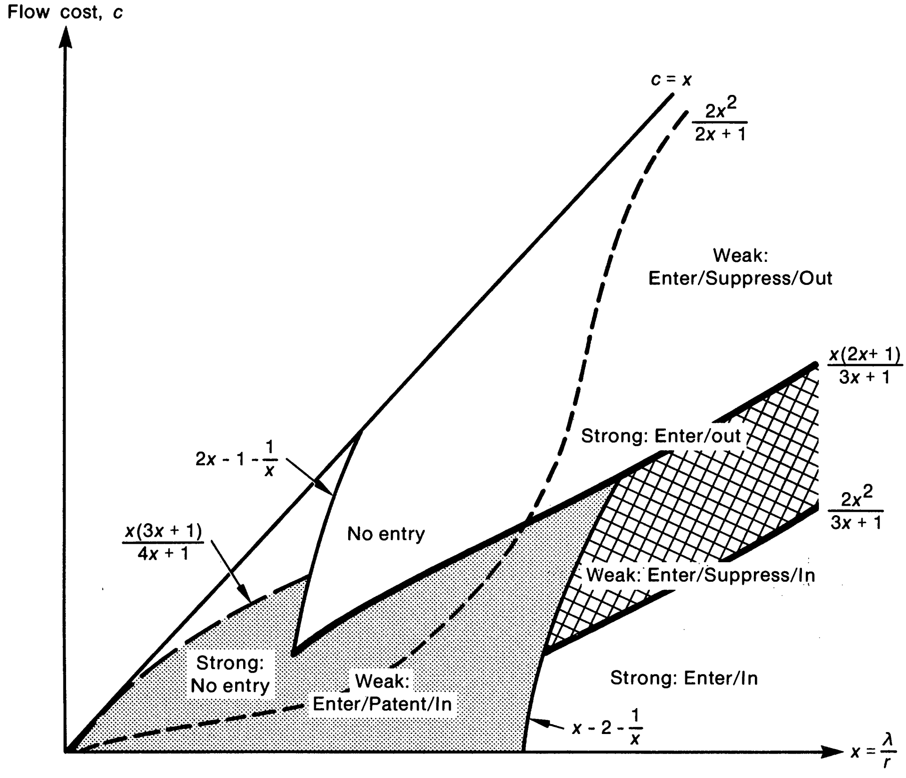
The equilibria of the subgame, and therefore the equilibrium decisions about whether

TABLE 2 **Continuation Payoffs if A Suppresses the First Innovation**

	Node 1 B stays in	Node 2 B drops out
Payoff to A	$\frac{x(5x+2)-c(3x+1)}{r(2x+1)^2}$	$\frac{2x-c}{r(x+1)}$
Payoff to B	$\frac{x}{r(2x+1)} - \frac{c(3x+1)}{r(2x+1)^2}$	0

FIGURE 2

EQUILIBRIA OF THE FIRST-TO-FILE GAME



to enter the race initially, will differ for different parameter values (x, c). In order to know whether B will stay in or drop out after A has invented but suppressed the first innovation, we must compare B's continuation payoffs in Table 2. We see from Table 2 that B stays in when A suppresses if and only if $[x(1 + 2x)]/[1 + 3x] \geq c$. Therefore we can divide the relevant part of the parameter space $(x \geq c)$ into two regions, according to whether this condition is met. These regions are separated by the bold line in Figure 2. (To reduce clutter in Figure 2, we have drawn the boundary lines only in regions where they apply. Therefore, we have terminated this line at the smallest x for which A will suppress.)

In order to know whether A will patent or suppress, we must compare A's payoff in Table 1 with the relevant payoff in Table 2. For parameters (x, c) for which B would stay in after A has suppressed an initial innovation, A will patent if $c \geq x - 2 - (1/x)$. If A can make B drop out by announcing, but not disclosing, the first innovation, A will patent if $c \geq 2x - 1 - (1/x)$. These regions are separated by the lighter lines in Figure 2.

Finally we must calculate the firms' *ex ante* profit, when each firm has probability one-half of being in position A or B after both firms enter. Firms will enter if *ex ante* profit is nonnegative. To find *ex ante* profit, the sum of both firms' continuation profits from the date of first innovation must be discounted back to the beginning of R&D, and the costs of R&D in that interval must be subtracted. There is a maximum cost parameter c for which *ex ante* expected profit is nonnegative. The maximum value of c differs over the three regions of the parameter space that determine the equilibrium of the postentry subgame, as shown in Table 3.

Above the line with long dashes in Figure 2 (in the region where B stays in) *ex ante* expected profit is negative. In the shaded region below the dashed line, *ex ante* expected

TABLE 3 **Upper Bound on c for Positive *Ex Ante* Profit**

A Patents	A Suppresses: B In	A Suppresses: B Out
$\frac{x(3x+1)}{(4x+1)}$	$\frac{7x^3+3x^2}{1+6x+10x^2}$	$\frac{2x^2}{(2x+1)}$

profit is positive. *Ex ante* profit is positive throughout the region where A suppresses and B drops out.

The region of Figure 2 labeled "No Entry" shows where *ex ante* profit would be negative if both firms entered. One might imagine that there could be asymmetric equilibria of the *ex ante* entry game, where one firm would profitably enter and the other would stay out. We now argue that this cannot happen. In the regions labeled "No Entry," initial entry by one firm alone would also be unprofitable.

Suppose that only one firm enters and does research until it makes a discovery. At this point it plays the role of firm A in our model. Even though no rival was racing for the initial innovation, one may enter after the first innovation if it is profitable to do so. This potential entrant is in the same situation as firm B. Firm A's expected costs until the first innovation are $c/[r(x+1)]$. Its continuation revenues at the first innovation differ according to whether we are in the left-hand side of the region of "No Entry," where the first innovation would be patented and firm B would then enter, or whether we are in the right-hand side of the region of "No Entry," where the first innovation would be suppressed and firm B would not enter. In the left-hand region, the continuation profits after the first innovation, given in Table 1, are $(1/r)(1-(c/(1+2x)))$. To find the profit of the single initial entrant, we must discount this continuation value to the beginning ($t=0$), from each possible date of discovery t, and subtract the expected flow costs of R&D that would be incurred until then. Hence, firm A's expected profit, when it enters alone, is equal to

$$[x-(c(3x+1)/(2x+1))]/[r(x+1)],$$

which is negative in the left-hand region of "No Entry," since the lower boundary of that region is defined by $c = x[(3x+1)/(4x+1)]$, which is larger than

$$x[(2x+1)/(3x+1)].$$

The calculation for the right-hand region of "No Entry" is similar, except that we must use the continuation profit in Table 2 Node 2. Firm A's expected profit, if it enters alone, is $[2x^2-(c(2x+1))]/[r(x+1)^2]$, which is also nonpositive in the right-hand side of the region labeled "No Entry." Therefore, there is no asymmetric equilibrium with one initial entrant.

Now we consider the strong novelty requirement, for which we have not drawn the game tree. A strong novelty requirement means that the initial discovery cannot be patented or marketed without infringing the previous technology. Therefore, the only decisions are whether firm B drops out when it learns of the first discovery, and whether firms enter initially. Equilibria of this game are also summarized in Figure 2, where, for example, the notation "Strong: Enter/Out" means that firms will enter initially, but then B will drop out after A announces the first innovation.

Firm B's decision whether to drop out is similar to its decision whether to drop out with the weak requirement if A has suppressed, but there is one key difference. With the weak requirement, B can earn unit profit until the second innovation if it catches up. Under the strong requirement, neither firm can sell anything until after the second discovery. In the cross-hatched region of Figure 2, B drops out with the strong requirement, but not with the weak requirement.

Conditional on B staying in after an initial unpatented discovery, B's expected continuation profit is slightly lower under the strong requirement than under the weak requirement when A suppresses. As can be seen in Figure 2, B will stay in after the initial unpatented discovery only if costs are relatively low. This is true for both novelty requirements. In this region, where costs are relatively low, *ex ante* profit is substantially higher than zero, and firms will enter initially under both novelty requirements. The fact that *ex ante* profit is slightly higher with the weak requirement does not make any difference to the entry decision in this region of the parameter space. If B will drop out after the initial unpatented discovery, *ex ante* expected profit is the same under the weak and strong novelty requirements. *Ex ante* profit is negative if $c > 2x^2/(3x + 1)$, shown by the line with short dashes in Figure 2.

The initial entry decision is different for the weak and strong novelty requirements only in the region where, under the weak requirement, the first innovator would patent. With the strong requirement, each firm's *ex ante* expected profit is lower because the first innovator will have to forego the profits it would have earned at the interim stage with the weak requirement. Thus, above the line with short dashes in the shaded region of Figure 2, no R&D will occur under the strong requirement, whereas R&D would have occurred under the weak requirement.

The equilibria of the games with the weak and strong requirements differ only in the shaded and the cross-hatched regions of Figure 2. In the shaded region, firms in equilibrium take the actions under the weak requirement that a social planner would prefer. Firms achieve the first best because all information is disclosed, and both firms participate, as is efficient if $x \leq c$. In the top portion of that region, the strong requirement would inhibit innovation altogether because *ex ante* profit is negative. Thus, in the shaded region, the weak novelty requirement is socially preferred to the strong one.

In the cross-hatched region, the equilibrium of the game with the strong requirement provides more social surplus than the equilibrium of the game with the weak requirement. The first innovation will be suppressed in both cases, but, with the weak requirement, B will stay in. B's incentive to stay in is too strong, relative to the social value of B continuing, because, due to the first-to-file rule, B might profit from duplicating A's initial discovery. We will show that this phenomenon is reversed with first-to-invent.

We can now shed light on our conjecture in Section 2 that the strong requirement might preclude competition between first and second inventors and thus enhance *ex ante* profit and increase entry. Indeed, for some parameter values, *ex ante* profit is larger under the strong requirement. But this occurs in a region (not shown in the diagram) where *ex ante* profit is already positive. Thus there is no case for the strong novelty requirement based on increased R&D incentives under the first-to-file rule.

This conclusion follows from several special features of the model, in particular, our assumptions that patenting involves full disclosure and that reverse engineering is possible. Suppose that disclosure were very incomplete. Then an innovator would always patent, since it would have little to lose by doing so. Patenting and marketing a small advance would undermine the previous innovator's profit without the reciprocal social benefit of reducing subsequent innovators' costs of progress. Similarly, if a marketed innovation could not be reverse engineered, firms would market small advances without patenting them. This would again undermine a previous innovator's profit. In both cases—incomplete disclosure or the impossibility of reverse engineering—the strong novelty requirement would sometimes lead to *ex ante* positive profit when the weak novelty requirement would not. For those parameter values, the strong novelty requirement is socially preferred.

4. Patenting and disclosure in first-to-invent

■ We now turn to the alternative legal rule that has applied in the United States, first-to-invent. The consequence for our game tree appears as a difference in the right branch of the tree in Figure 1. Suppose that the rule is first-to-invent and A chooses not to patent the first innovation. Then if B duplicates the first innovation and applies for a patent, A will also apply for a patent, and will be awarded the patent.[16] We can therefore assume that A patents directly, although in fact A will be forced by B to patent.[17] At node 3 in Figure 1, A, rather than B, will patent and the terminal payoffs below become $(2/r, 0)$ and $(0, 1/r)$. The interim flow profits of one that A earns are delayed relative to the left branch of the tree in which A patents at the outset. But delaying may nevertheless be valuable to A, since there is a ½ probability that A will also get the next invention and thereby achieve flow profits of 2, rather than being forced to patent the first improvement and then compete on an equal footing for the final improved technology.

With first-to-file, there were more parameter values for the weak requirement than for the strong requirement for which B would drop out after the first unpatented innovation. This was because interim profit was possible with the weak requirement, as was final profit. This conclusion is reversed with first-to-invent. With the weak requirement, B cannot get a flow rent of two at the terminal node, unlike with the strong requirement. Even if B wins the final improved technology, it cannot earn more flow profit than one, since A can market the first improved technology (whether or not A patents it). With the strong requirement, there is a chance that B will end up with a flow profit of two.

Using calculations similar to those in Section 3, we discover that B will drop out after the first unpatented innovation if $c \geq x^2/(3x + 1)$ when the novelty requirement is weak, and if $c \geq 2x^2/(3x + 1)$ when the novelty requirement is strong. These are shown as bold lines in Figure 3. Since A retains property rights in the first innovation whether or not it is initially patented, A is less inclined to patent. A will patent when $c \geq 2x - 1 - (1/x)$, whether or not B will drop out. Thus, in Figure 3 we have only one patenting line, rather than two as in Figure 2.

The *ex ante* profit calculations for first-to-invent are the same as for first-to-file, conditional on whether B stays in or drops out. If B stays in, the interim profit accrues to A when B duplicates the first innovation, rather than to B, but the two firms' total profit is the same. The lines that separate regions of positive and negative *ex ante* profit are shown as the dashed and dotted lines in Figure 3. *Ex ante* profit is negative above the dashed line when the novelty requirement is weak and the first innovator would patent. *Ex ante* profit is negative above the dotted line when the first discovery is suppressed and the competitor drops out. This may occur either with the weak or strong requirement.

In the shaded area in Figure 3 labeled "Weak: Enter/Patent," the weak novelty requirement is unambiguously better than the strong one, since the first innovation is disclosed and then B stays in. With the strong novelty requirement, B would drop out after the first innovation in part of this region, and, in another part, the firms would not enter at all. (To avoid clutter, we have not drawn the lines that bound these regions.)

In the areas of the diagram with no shading, the equilibria of the strong and weak game give the same social value. In the areas marked "No Entry," firms would not enter *ex ante* under either novelty requirement. In the lower such area, this is true only because, in perfect equilibrium, the first innovator will not patent, even if possible. If the firms could commit *ex ante* to patent the first innovation, then the weak novelty requirement would be preferred to the strong requirement in this region.

In the areas of Figure 3 with shaded lines, the equilibria of the games with the weak

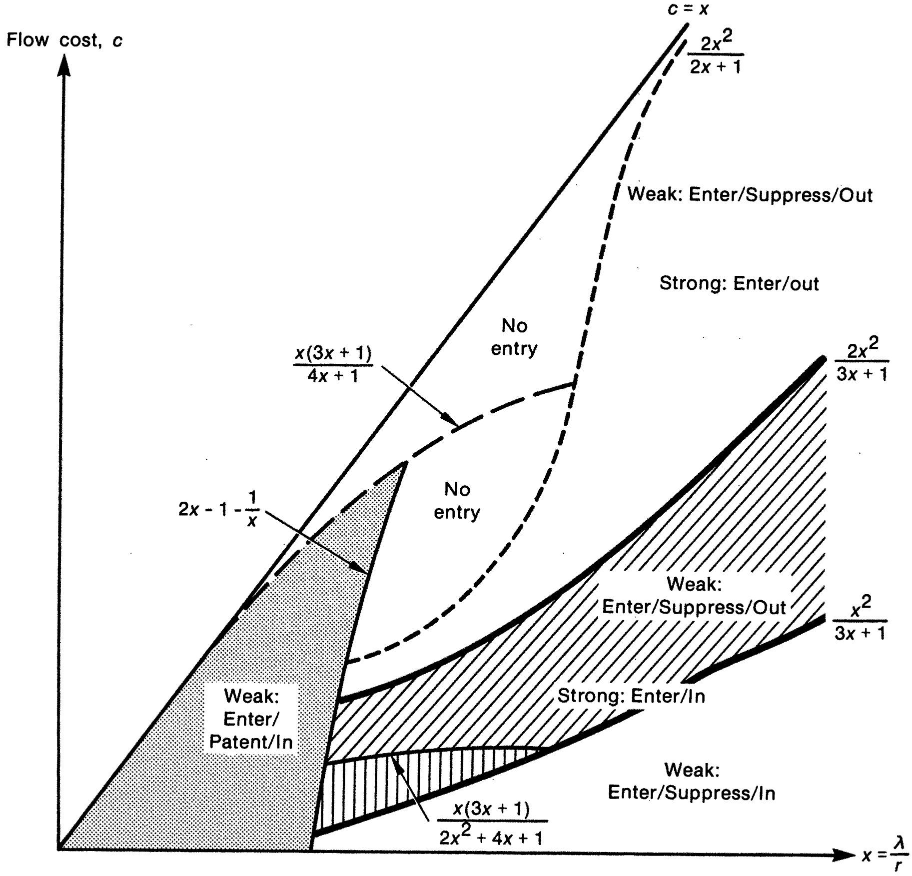
---

[16] This is only possible if A applies before a patent issues to B. Current lags suggest this assumption is well justified.

[17] B has no incentive to obscure the fact that it has invented, since it cannot earn flow profits of 2 at the end, even if A does not patent. If B patents the final invention and the novelty requirement is weak, A can still market the first, inferior invention without infringement. This is what it means for the novelty requirement to be weak. By announcing that the innovation may have other benefits to B, such as enticing venture capital, we assume that B makes its invention known.

FIGURE 3

EQUILIBRIA OF THE FIRST-TO-INVENT GAME



and strong requirements are similar in that the first innovation would be suppressed. The games differ, however, in whether B would stay in after A innovates. With the weak requirement, B would drop out, since even if B achieves the final step before A, B's profit would be eroded by competition with A because A can market the first innovation.

The social value of the competitor staying in the race is that the final improvement is achieved faster. The cost is that the competitor is one step behind and will have to duplicate the knowledge embedded in the previous technology in order to make progress. If $c \leq x(3x+1)/(2x^2+4x+1)$, the benefits of speeding up the research dominate the costs of duplication, and the strong novelty requirement is therefore socially preferred to the weak one. In the area of Figure 3 with vertical lines, a strong novelty requirement is socially better than a weak novelty requirement because the competitor will drop out with a strong requirement, and this is efficient.

As with first-to-file, there are no parameter values for which patenting leads to negative profit when this would be nonegative without patenting. When costs are sufficiently high that *ex ante* profit might be negative if the first innovation were patented, it will not be patented for strategic reasons.

5. First-to-file versus first-to-invent: welfare comparison

■ As we have stressed, the first innovator retains property rights in a previous innovation under first-to-invent whether it is initially patented or not, and this should lead to a reluctance to patent, relative to first-to-file.

Figure 4 shows the parameter region for which the first innovator will patent in first-to-file, but not in first-to-invent. Since disclosure of the first innovation accelerates discovery of the final innovation, the first-to-file rule is superior to first-to-invent in the sense that it encourages disclosure.

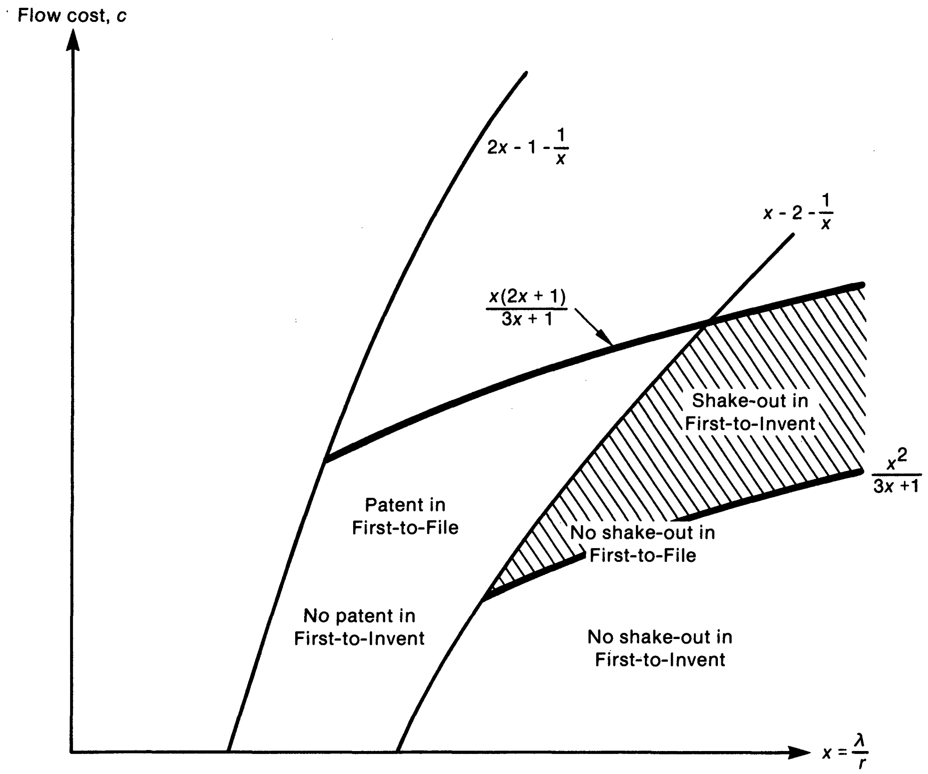
But first-to-file is not unambiguously better than first-to-invent. This is because a shakeout in the race may be socially beneficial, and a shakeout will occur for more parameter values with first-to-invent than with first-to-file. Figure 4 shows the parameter region in which B will drop out after the first unpatented innovation when the rule is first-to-invent, but will not drop out in first-to-file. As shown previously, it is efficient for the competitor to drop out only if $c \geq x(3x + 1)/(2x^2 + 4x + 1)$. Therefore, for parameter values in the shaded area of Figure 4, social welfare will be higher with first-to-invent than with first-to-file when the novelty requirement is weak. (If the novelty requirement is strong, there is no difference between first-to-file and first-to-invent.)

6. Conclusion

■ Our central observation is that disclosure of technologies is socially valuable because future research builds on previous technical knowledge. Disclosure of advances reduces the cost of seeking progress for other researchers. Firms may be reluctant to disclose (patent) interim technologies in a multistage race because they cannot profit from the cost reductions provided to competitors. Indeed, these externalities hurt them directly in that they make the competition more potent. As a consequence, the apparent social value of making the novelty requirement weak—to encourage disclosure of many small increments to technical

FIGURE 4

COMPARISON OF FIRST-TO-FILE WITH FIRST-TO-INVENT



knowledge—is undermined. Firms might not patent interim technologies, even if it were possible.

On the other hand, there is often nothing to be gained by having a strong rather than a weak novelty requirement. A weak novelty requirement permits interim technologies to be patented and this is socially valuable, provided it does not undermine *ex ante* profit so that firms are dissuaded from research. In our model, firms will not be dissuaded from research by the erosion of profit, because whenever profit would be eroded in this way, the firm will choose not to patent. Hence, the weak novelty requirement is attractive. Its attractiveness might be reduced if disclosure at patenting is incomplete or if reverse engineering is impossible so that, under a weak novelty requirement, firms would be willing to market their advances without patenting them.

Firms that achieve an initial competitive advantage can use that advantage to force a shakeout in the industry by announcing the discovery without patenting it. For the announcement to be credible, it must serve another purpose, such as attracting venture capital, in which fraud would be very costly to the firm. Shakeouts may or may not be socially valuable, depending on the costs of research. Neither the strong nor the weak novelty requirement (with either dispute-resolution rule) induces shakeouts precisely when shakeouts are efficient. The weak novelty requirement gives too strong an incentive for the competitor to stay in when the dispute-resolution rule is first-to-file, and too strong an incentive for the competitor to drop out when the dispute-resolution rule is first-to-invent. Since it is more difficult for the first innovator to shakeout the competitor in first-to-file, the first innovation will be patented for more parameter values than in first-to-invent. The strong novelty requirement also gives too strong an incentive for the competitor to stay in.

In this article we have focused on a tension between protecting profit and encouraging disclosure. When the novelty requirement is weak, first-best efficiency will not always be achieved because firms will not disclose all of their technical progress. If firms could be forced to disclose, first-best efficiency might not be achieved because disclosure could undermine profit, so that firms might not enter the race at all. In companion papers (Green and Scotchmer, 1989; Scotchmer, 1990) we investigate an obvious resolution of this tension: if firms whose products are close substitutes can form cooperative agreements, they can avoid competition and keep all the social surplus as profit. Licensing allows firms to patent and disclose all their innovations without undermining profit, and can therefore provide correct incentives to innovate if the profit is divided appropriately between the innovators. We discuss how the novelty requirement should be chosen in order to ensure that the licensing agreement divides the profit efficiently.

Klemperer (1990) and Gilbert and Shapiro (1990) have approached the novelty issue in a way that does not focus on cost externalities or sequential innovation. Assuming that the social planner wants to guarantee the innovator a specified level of profit, they ask whether patent protection should be "long" or "broad." Length and breadth should jointly be chosen to minimize the social cost of monopoly pricing. We have omitted monopoly pricing distortions so as to isolate the profit incentives for R&D in markets with information externalities among innovators.

Schmitz (1989) ignores the strategic incentive not to patent (disclose) interim technologies and shows that, when the value of the sequence of innovations is large relative to its cost, the novelty requirement should be weak.

References

GILBERT, R. AND SHAPIRO, C. "Optimal Patent Length and Breadth." *The RAND Journal of Economics*, Vol. 21 (1990), pp. 106–112.

GREEN, J. AND SCOTCHMER, S. "Antitrust Policy, the Breadth of Patent Protection, and the Incentive to Develop New Products." Working Paper 1067. Harvard Institute of Economic Research, 1989.

——— AND ——— "The Novelty Requirement and Diversity in Research Lines." Mimeo, 1988.

GROSSMAN, G.M. AND SHAPIRO, C. "Optimal Dynamic R&E Programs." *Economic Journal,* Vol. 97 (1987), pp. 372–387.

HORSTMANN, I., MACDONALD, G.M., AND SLIVINSKI, A. "Patents as Information Transfer Mechanisms: To Patent or (Maybe) Not To Patent." *Journal of Political Economy,* Vol. 93 (1985) pp. 837–856.

KLEMPERER, P. "How Broad should the Scope of Patent Protection Be?" *The RAND Journal of Economics,* Vol. 21 (1990), pp. 113–130.

MANSFIELD, E. "R&D and Innovation: Some Empirical Findings." In *R&D, Patents and Productivity,* Zvi Griliches, ed., Chicago: University of Chicago Press for the National Bureau of Economic Research, 1984.

SCHMITZ, J.A., JR. "On the Breadth of Patent Protection." Mimeo, SUNY-Stonybrook, 1989.

SCOTCHMER, S. "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law." *Journal of Economic Perspectives,* Symposium on Intellectual Property Law, forthcoming.

——— AND GREEN, J. "Novelty and Disclosure in Patent Law." Working Paper 1388 Harvard Institute of Economic Research, 1988.

SHAPIRO, C. "Patent Licensing and R&D Rivalry." *American Economic Review,* Vol. 75 (1985), pp. 25–30.

YOSHIKAWA, A. "Japanese Biotechnology: New Drugs, Industrial Organization, Innovation and Strategic Alliance." Berkeley Roundtable on the International Economy, 1988.