

This PDF is a selection from a published volume
from the National Bureau of Economic Research

Volume Title: Innovation Policy and the Economy,
Volume 2

Volume Author/Editor: Adam B. Jaffe, Josh Lerner
and Scott Stern, editors

Volume Publisher: MIT Press

Volume ISBN: 0-262-60045-5

Volume URL: <http://www.nber.org/books/jaff02-1>

Conference Date: April 17, 2001

Publication Date: January 2002

Title: Intellectual Property: When Is It the Best Incentive
System?

Author: Nancy Gallini, Suzanne Scotchmer

URL: <http://www.nber.org/chapters/c10785>

Intellectual Property: When Is It the Best Incentive System?

Nancy Gallini, *University of Toronto*

Suzanne Scotchmer, *University of California, Berkeley*

Executive Summary

Intellectual property is not the only mechanism used in the American economy for rewarding R&D. Prizes and contract research of various types are also common. Given the current controversies that swirl around intellectual property policies, we review the economic reasoning that supports patent and other intellectual property over the alternatives. For those economic environments where intellectual property is justified, we review some of the arguments for why it is designed as it is. We focus particularly on the issue of how broad awards should be and how much protection should go to the original inventor (as opposed to those who subsequently improve the invention). We emphasize that the ideal design of an intellectual property system depends on the ease with which rightsholders can enter into licensing and other contractual arrangements involving these rights.

I. Introduction

Intellectual property is the foundation of the modern information economy. It fuels the software, life sciences, and computer industries, and pervades most other products we consume. Although most inventors consider it essential, it is currently under attack by some academics and policymakers. One complaint is that intellectual property rewards inventors beyond what is necessary to spur innovation. Another is that intellectual property is a drag to innovation, rather than a spur, since it prevents inventions from being used efficiently, especially in creating further innovations. A third complaint is that some inventions should not be protected at all but, instead, be supported by public sponsors.

Controversies over what should constitute intellectual property swirl around business methods, computer software, research tools in the biomedical industry, and genetic sequences. However this is not new; controversies have swirled around every new technology in the

twentieth century. A sampler might include the question of whether player piano rolls should receive copyright protection, whether purification of chemical compounds constitutes "invention" for purposes of patent law, and whether mathematical algorithms such as public key encryption should be patentable. Technologies that fall outside the subject matter of patents and copyrights have sometimes received *sui generis* protections, such as computer chips under the Semiconductor Chip Protection Act.

For all these technologies, the same questions arise: Are there natural market forces that protect inventors so that formal protections or other incentives are not necessary? If not, is intellectual property the best incentive system, or would the technology more appropriately be developed by a public sponsor and offered freely in the public domain? How should intellectual property be designed so as to minimize deadweight loss due to monopoly pricing without undermining incentives to innovate?

Our objective in this paper is to review what economists have said about incentive schemes to promote R&D, including intellectual property. While we focus on environments in which other forms of protection are not available, we note that other protections can obviate the need for any formal reward system. For example, encryption offers the potential to protect digitally distributed products such as music, movies, and software, even in the absence of intellectual property (National Research Council 2000). In the realm of databases, for which formal protections have been mandated in Europe and proposed in the U.S. Congress, vendors are protecting their data with both clever business strategies and technology (Maurer 1999, Maurer and Scotchmer 1999). In markets with network effects, there may be natural barriers to entry, so that a vendor may capture the entire market even without formal protection (Farrell 1995). And, of course, trade secrecy can be an important protection, especially when firms devise clever nondisclosure agreements that enable them to license without leaking the secret to unauthorized users (e.g., see Anton and Yao 1994). In some of these examples, the alternative protection involves social costs that could be avoided by formal intellectual property. But if not, the case for intellectual property may be weak.

In section II, we compare intellectual property with alternative incentive schemes. Without losing the thread of the paper, the reader who is only interested in the design of intellectual property (as opposed to other incentive schemes) can skip the last three subsections of

section II. In section III we review optimal design issues for intellectual property, especially the question of patent breadth, and in section IV we turn to the special problems that arise when innovation is cumulative. In section V, we summarize the arguments for and against intellectual property. We comment on whether the design recommendations of economists can actually be implemented, and argue that intellectual property regimes should be designed so that the subject matter of each one has relatively homogeneous needs for protection.

II. Alternative Mechanisms for Rewarding Innovation

Competitive markets may not be conducive to innovation, for a reason that was well articulated by Arrow (1962). Inventions are *information*, and information is a public good. An invention such as a wireless palmtop is a combination of tangible embodiments and an intangible idea, as well as information about how to manufacture it. Typically, both the information and the tangible embodiments are costly to the inventor, but only the tangible components are costly to a rival. Without some sort of protection or reward, the inventor will therefore be at a market disadvantage relative to rivals, and will possibly be dissuaded from investing.

Arrow explained why *some* incentive scheme is needed, but not *which* scheme. Many schemes have been used in practice. In the seventeenth century, for example, a prize was offered in France for developing a workable water turbine (Reynolds 1983, p. 338). For about a century in the same era, a prize was outstanding for developing a method to calculate longitude at sea (Sobel 1995). In the modern era, R&D is sponsored to a large extent by government grants. According to the National Science Foundation (2000), in 1998 about 30% of U.S. research was funded by the Federal government. These examples raise the following question: In what environments are there better incentive schemes than intellectual property?

We shall use the term *intellectual property* (IP) to mean an exclusive right to market an invention for a fixed time period. It includes copyrights, patents, plant patents, protection under the Plant Variety Protection Act, and other *sui generis* types of protection. By a *prize* we mean a payment funded out of general revenue that is made to a researcher conditional on delivering a specified invention. Prizes can either be tailored individually to firms, depending on their efficiency characteristics, or can be offered symmetrically to any firm that wants to compete,

just as a patent is. By *procurement*, we mean a mechanism to solve the problem of getting an invention at minimum cost, in a timely manner, or otherwise efficiently (e.g., Laffont and Tirole 1986, 1987). A simple procurement mechanism would be an auction for the right to be paid when the invention is delivered.

A form of procurement commonly used in government-sponsored research appears, on its face, to be a fixed-price contract. For example, the National Institutes of Health give funding in advance for projects that are described in the proposals. Funds are not withheld if the output is not delivered, since the idea of the contract is to pay costs as they accrue. If such funding were a one-time event for each researcher, researchers might be inclined to "take the money and run." This moral hazard problem is overcome because future grants are contingent on previous success. The linkage between previous success and future funding seems even more specific in the case of the National Science Foundation. Fixed-price contracts thus operate much like prizes, with the wrinkle that a researcher must convince the sponsor in advance that his output might be worthy of a prize. For this purpose, his reputation might suffice, and in some cases, much of the research has already been completed.

We begin our analysis with a benchmark. When both the costs and values of innovations are publicly observable to both firms and a public sponsor, IP is not the best incentive scheme. A better scheme is for a public sponsor to choose the projects with the largest net social benefits, and pay for them on delivery, using funds from general revenue. With IP, projects are funded out of monopoly profits. Monopoly pricing is equivalent to taxing a single market, which is generally thought to impose greater deadweight loss than the broad-based taxation that generates general revenue. Thus, to justify intellectual property, there must be some type of asymmetric information about the costs and benefits of research programs.

We first make some comparative remarks about intellectual property, prizes, and procurement contracts. These remarks are much in the spirit of Wright (1983), who gave the first formal treatment of how asymmetric information should inform our choice among incentive mechanisms. In the subsections that follow, we then show that these three mechanisms can generally be improved upon.¹

IP has an obvious defect as well as obvious virtues. The defect is the deadweight loss due to monopoly pricing. The virtues are several. Most importantly, if the costs and benefits of R&D investments are

known only to firms, and not to government sponsors, firms will use their superior knowledge to screen investments. A sponsor does not need to decide in advance which investments are meritorious. An investor knows that he will be punished by the market if he does not invest wisely. Another obvious virtue is that the prospect of valuable IP might elicit higher levels of effort than those generally associated with sponsored research. For example, much has been made of the human genome project, whose completion was accelerated by a private firm hoping to win IP rights on gene sequences. Finally, an IP system imposes the costs of an invention on its users. In other incentive mechanisms, the costs are borne more generally by taxpayers. Taxpayers might rightfully revolt if asked to bear the costs of developing, say, computer games.

Lest these advantages of IP be overstated, however, we note that prizes have many of the same virtues. If an investment's prospective value is known to the sponsor (or defined by the sponsor, as in the case of military wares), the sponsor can screen projects himself. A prize system then seems superior to IP. It avoids deadweight loss, and can be as good as IP at motivating effort.

Moreover, IP will not work as an incentive mechanism unless third parties can observe at least some aspects of value. A rightholder must be able to defend his right against potential infringers. He must be able to prove in court that his intellectual property meets the standard for protection, and that an alleged infringer is marketing a product that falls within the breadth of his claims. Aspects of the invention's value must therefore be observable *ex post*, although typically at the high cost of litigation and discovery.

The *ex post* observability requirement will typically impose less cost under an IP system than under a prize system. Under an IP system, the costs of discovery are incurred only if there is litigation. In contrast, for a prize, costs would have to be incurred for every invention in order for the sponsor to set a payment commensurate with the value.² Therefore, our distinction is not really between observability and nonobservability, but rather on whether the value is known to the sponsor without incurring cost. The most natural example is when the sponsor defines the value of the invention himself, as in military procurement.

Recently the World Health Organization and the World Bank have suggested prizes for developing vaccines that would not be developed or might not be widely enough distributed under a system of

proprietary rights. The problems are great: how to assess whether a vaccine merits a prize; how to ensure that the prizes are not given prematurely before higher-quality vaccines are brought forward; how to ensure that the prizes are actually given, when it is easy to manufacture reasons to withhold them. Prizes can be organized so that worthy projects need not be identified in advance, but administering the prize then becomes particularly burdensome. The problems are particularly acute where innovation is cumulative. See Kremer (2000) for a thoughtful and detailed analysis of how such a system might work.

Unlike IP, a procurement contract would typically not be offered to all comers. Instead there would be a negotiation phase in which the procurement officer tries to sort out which firm(s) are more efficient, and only offers the “prize” to those firms. A mechanism that allows such flexibility is more effective by definition than a prize offered to all comers. As for prizes, the sponsor must identify worthy projects. For traditional government procurement, such as for fighter jets, this is automatic. For medical research, the sponsor may solicit open-ended proposals, which entails administrative cost. In addition, the negotiation required for procurement might be politically infeasible as well as costly.

In the next subsections, we investigate optimal incentive mechanisms in specific research environments, with a view toward understanding how optimal mechanisms relate to IP, prizes, and simple reimbursements. We focus on environments in which no alternative mechanisms for protection (private or market) are available, and on single inventions that do not lead to future innovations. Following Scotchmer (1999b), we stylize the allocation problem as having three facets, which are intertwined. The first is the *decision* problem: should a project be undertaken? The second is the *delegation* problem: by which firms, or how many, and at what rates of investment? The third is the *funding* problem: Can the deadweight loss of monopoly pricing be avoided?

The Problem of Aggregating Information

To solve the decision, delegation, and funding problems jointly, all the information that is decentralized among firms may have to be aggregated. IP, prizes, and simple procurement mechanisms such as fixed-price contracts and auctions cannot aggregate information, and are therefore flawed at the outset.

To see this, consider a well-defined project, such as finding an AIDS vaccine or developing supersonic transport. Suppose that there are two potential researchers, $i = 1, 2$, and that each researcher i has an efficiency parameter c_i for this project, interpreted as the cost of success. The product will have a common value v regardless of which firm develops it, and each firm has a signal v_i of this value. The underlying value will typically be determined by the extent of demand or anything else that affects monopoly profit and social welfare. Because each v_i is a noisy signal of an underlying common value, it is natural to suppose that the signals $\{v_1, v_2\}$ are correlated. It is less obvious whether the cost parameters $\{c_1, c_2\}$ would be correlated. We shall assume that they are independent draws from a known distribution.

To make an efficient investment decision, each firm would like to know the other firm's signal. For example, a firm with a low signal of value, $v_1=L$, might invest if it knew the other firm had a high signal of value, $v_2=H$, but not otherwise. But neither the value nor its best estimate is known *ex ante* to other firm, since neither can observe the other's signal.

The importance of aggregating information is revealed in the following special case in which both the costs and the signals take on binary values: $c_i \in \{l, h\}$, $v_i \in \{L, H\}$. Suppose that the first-best, full-information rule for allocative efficiency is that the project should be undertaken unless (1) both firms have high costs, regardless of the signals of value or (2) both firms have low signals of value, regardless of costs. The project should be undertaken by a single firm if (3) at least one firm has low cost and at least one firm has a high signal of value, or (4) both firms have high cost and both have high signals of value.

Suppose $(c_1, v_1) = (l, L)$. Firm 1 should invest if $(c_2, v_2) = (h, H)$ but not if $(c_2, v_2) = (h, L)$. Without knowing firm 2's information, firm 1 could not make an efficient decision. Such could be the case under a patent system. Firm 1 may fail to invest because it is pessimistic about value ($v_1 = L$), and firm 2 may fail to invest because its costs are too high ($c_2 = h$). If the firms could share their information, firm 1 would invest based on firm 2's propitious information about the market. To some extent, the firms should be able to learn each other's private information by observing each other's investments. However, even if the firms know each other's costs, they might get stuck in an inefficient, but self-reinforcing, equilibrium where each invests because the other is investing, and each incorrectly thinks the other has a high signal of value (or vice versa) (Minehart and Scotchmer 1999). When the firms have

different, unobservable costs, the difficulties of making inferences from investment behavior are compounded. A firm that invests could be doing so either because it has low cost or because it has very propitious private information about the market. The observing firm cannot distinguish between these two cases.

Neither IP nor prizes nor simple procurement mechanisms (e.g., auctions) can cope with the problem of aggregating information. Scotchmer (1999b) describes a procurement mechanism that bears little resemblance to auctions, prizes, or IP, but can achieve as good an outcome as when the signals of value are known, provided the firms' signals of value are correlated.³ While the mechanism described will delegate efficiently, it may not be realistic given the constraints of government procurement. The mechanism might entail payments from firms to the government, or payments to firms that are not asked to invest. Such payments would be difficult to enforce.

The problems with the efficient procurement mechanism may explain the use of prizes, IP, and simple procurement mechanisms, but, under the conditions presented in this example, no one has studied their relative merits as second-best mechanisms. In order to identify the relative merits of the simple schemes and other more realistic mechanisms, we now consider the decision and delegation problems separately.

The Delegation Problem

We isolate the problem of optimal delegation by assuming that the sponsor already knows the optimal decision, namely, to invest. That is, the sponsor knows the value of the project and that it exceeds the cost of delivery, but it does not know which firm(s) is (are) more cost-efficient. Optimal delegation has two components: choosing the most efficient firm or group of firms, and motivating the firm(s) to invest at efficient rates.

If the sponsor faced only a problem of selecting the more efficient firm(s), then the delegation problem would be easy to solve, e.g., by auctioning the right to invest. In contrast, IP and prizes could lead to inefficiency. If the market has room for only one firm, there is no reason to suppose that the lower-cost firm will be the entrant, especially when the relative efficiencies of the firms are not publicly observable.

But even an auction will not perform well when there is also a problem of inciting the right amount of effort, so that the invention is deliv-

ered in a timely manner. The appropriate rate of progress is key to the economics of R&D: How much additional cost should be tolerated in return for a higher rate of progress?

A firm's willingness to accelerate invention at higher total cost depends on the prize it will receive, conditional on delivering the product. Thus the size of the prize determines the rate of investment. However the optimal size of prize (and the optimal rate of investment) depend both on the researcher's "efficiency" and on his efficiency relative to other firms. For an inefficient firm, the optimal rate of investment may be zero if it is possible to delegate to a more efficient firm, but positive if the other firm is even less efficient. Thus, the problem is to tailor prizes both to the firms' *individual* efficiencies and to their *relative* efficiencies.

Gandal and Scotchmer (1993) study this problem, and show that the sponsor should offer a menu of options with both fixed fees and firm-specific prizes.⁴ The menu serves two purposes: it gets the firms to reveal their relative efficiencies, and, once the contracts are awarded, it gets the firms to invest at the efficient rates. The difficulty is in the coordination: each firm's efficient rate depends on both firms' efficiency parameters. A simple patent or prize system, where the IP or prizes are not tailored to the firms' relative efficiency, will not ensure that only the most efficient firm(s) invest, or at the efficient rates. And a simple fixed-price contract may not create incentives to invest fast enough, even if the contract is auctioned to the more efficient firm.

The message here is that, even when the value of the prospective invention is known prior to the investments, optimal procurement requires a mixture of prizes and fixed payments, rather than a pure prize system, a patent system, or an auction. Simple mechanisms can be resurrected as best in very simple contexts. An auction performs well when the only issue is to choose the most efficient firm, but there is no issue of eliciting the right amount of effort. A simple prize performs well when there is a single firm qualified to undertake the research. If the prize is set equal to the social value, the firm will have the same objective function as society and will invest efficiently. Since the best simple mechanisms are different for different simple contexts, it is no surprise that complicated research environments with several firms call for mechanisms that combine instruments.

In the next section we focus on the optimal decision problem, assuming that the value of the innovation is unknown. In order to avoid the

problem of optimal delegation, we also assume there is a single potential researcher.

The Decision Problem

We have just pointed out that if there is a single firm qualified for the research program, the optimal mechanism is a prize set equal to the social value. The firm's private incentives are then aligned with social incentives. However, to set such a prize, the sponsor must know the social value in advance or observe it *ex post*. Since IP automatically reflects the social value, at least to some extent, IP looks like an attractive alternative to a prize when the social value is unobservable. We now investigate whether this justification for IP holds up.

Kremer (1998) proposes a system to create a prize equal to the social value, even when the sponsor cannot observe it in advance. His proposal involves IP, but avoids deadweight loss by turning a patent into a prize. He proposes that the patent authority take possession of the patent, and auction it to the highest bidder, assuming that every firm can observe the value *ex post*. The rules of the auction are that with very small probability the patent will actually be sold to the highest bidder, and otherwise the invention will be put in the public domain. Firms will bid the true value, hence revealing it. The social value is estimated from the revealed private value, and the inventor receives a prize equal to the social value, paid out of general revenue. He will thus invest if the social value exceeds his cost, as is efficient.⁵

Another scheme to avoid deadweight loss is proposed by DeLaat (1996). To illustrate his idea in a very simple model, suppose that a potential R&D project is described by a pair (c, v) , where c is the cost, which is observable to the sponsor, and v is the value, which is not. But if the cost c is observable to the sponsor, he can ask the researcher to report the prospective value v , and then give a fixed-price contract to reimburse the cost c if and only if the prospective value exceeds the cost. Since the researcher earns zero profit whatever he reports (he is only reimbursed the cost), he will report the value truthfully to the sponsor, who will make the efficient decision whether to invest. Thus, IP is unnecessary.

But this scheme only seems credible if (contrary to the premise) the value of the invention is observable *ex post*, or if the sponsor can verify that the researcher is investing exactly as he promised (as deLaat assumes explicitly).⁶ If not, the researcher could use the contract money

for other purposes and deliver a shoddy product; there is a disabling problem of moral hazard, which IP could overcome.

Nevertheless, we can conclude from the arguments of Kremer and deLaat that if either cost or value is truly observable to a sponsor, there may be a better mechanism than IP. Consistent with this view, Scotchmer (1999a) justified patents by assuming that the cost and value are both unobservable. A similar interpretation can be made for Cornelli and Schankerman (1999). The latter present a model where the value of an invention is endogenous to the firm's investment effort, which, in turn, depends on an unobservable efficiency parameter. In effect, neither cost nor value is observable to the sponsor. Thus it is hard to see how any mechanism short of IP could be effective. Since the value of the patent increases with the value of the invention, a patent system gives the firm at least some incentive to spend more resources to create a product of greater value. Cornelli and Schankerman show how this incentive can be increased by using a patent renewal system.

The patent renewal system is a menu of options (F, T) , where F is a payment from the patentholder to the sponsor and T is a patent life.⁷ The fee F increases with the patent life, and might start out negative (a subsidy). The patentee can then "buy" a longer patent life by paying renewal fees. The value of the patent automatically increases with the value of the invention, but increases more for higher-value inventions, since those are the ones that will be renewed in return for fees. Thus the incentive to develop higher-value products is compounded.

Scotchmer (1999a) derives the renewal system as a multidimensional screening mechanism for ideas, (c, v) , where both are unobservable. Again, it is the higher-value ideas that will be renewed the longest, compounding their value. Thus the cost c that firms are willing to bear may go up faster than linearly with the value of the innovation, v .

As mentioned, the renewal system could start with subsidies, which are then reduced as firms pay fees in return for a longer patent life. Subsidies are advocated by Shavell and van Ypersle (1998) on grounds that they are a more efficient way to reward innovators than IP. Subsidizing low-value innovations allows the protection on high-value innovations to be shorter (thus reducing the deadweight loss), without jeopardizing incentives to innovate.

The problem with subsidies, of course, is that they may be exploited by opportunistic firms, which could collect the subsidy and either not invest or produce something worthless. To avoid this problem, subsidies, like their close kin, prizes, must be contingent on some aspect of

the resulting invention, such as its value. Thus it seems reasonable to suppose that subsidy schemes will not be used if the invention's value or success cannot be verified *ex post*. But then we have a contradiction. If subsidies are possible, it must be because some aspect of value is observable *ex post*. If so, IP should not be used at all, since prizes (rewards, fixed-price contracts) dominate. IP and prizes can serve the same screening function, and can motivate firms to the same levels of effort, but prizes avoid the deadweight loss. Consistent with this caveat, renewal schemes seen in practice do not provide for subsidies. (See Calandrillo (1998) for a broader set of criticisms of subsidies.)

In conclusion, IP can be justified in two ways. First, it can be justified as a screening mechanism to encourage investment in high-value projects, which may also have high cost. Second, it can be justified as a means to increase the rate at which firms invest, either to increase value or to accelerate progress. Without a means to link prizes to social value, there is no alternative to achieve these results. These virtues of IP should be weighed against the aggregation problems described earlier when more than one firm is capable of the research.

Assuming that, in a second-best analysis, IP would prevail, we now ask how the right should be designed. We have already discussed the benefits of a renewal system. But how broad and long should protection be?

III. Optimal Design: The Case of a Single Innovation

Perhaps the most influential work on patent design was that of Nordhaus (1969), who explained why patents (or other IP) should have finite length. If the sole concern is to encourage innovation, then IP should last forever. And if the sole concern is to avoid deadweight loss that occurs through proprietary prices, then IP should not exist at all. A finite length of protection balances these two concerns. Longer protection would encourage more innovation, but only by prolonging the deadweight loss on inventions that would be made anyway.

Nordhaus's simple framework spawned a large literature on the design of IP with consideration of patent races, imitation by rivals, technology licensing, and how the design question changes when technology is cumulative. In this section we focus on the design question of *breadth* (also called *scope*), which occupied considerable journal space in the 1990s. In the next section we turn to sequential or cumulative research, where breadth plays a different role.

We begin with Gilbert and Shapiro (1990; GS), who introduced the notion of patent scope into the Nordhaus analysis. They define patent scope as the price p that the innovator is able to charge for the product that embodies the innovation. Thus a patent policy is (T, p) , where T is the patent life. While such a definition is far removed from what a court might use, the analysis that arises from using it is still informative, as discussed below. Maximizing social surplus over all combinations (T, p) that yield enough revenue to cover the cost of research, GS find that optimal patent length is infinite, with the patent scope set at the level that just covers R&D investment.⁸ That is, the optimal design is for the patent to be narrow and long.

Gallini (1992) reversed this design conclusion in a model where patent breadth determines the ease of entry into the protected market. She defined scope technologically, as the cost K that rivals must incur to imitate the invention without infringement. Thus a patent policy is a pair (T, K) . The lower price that results from narrow scope arises from rivals' attempts to "invent around" the patent, rather than from some type of regulatory or antitrust action, as assumed by GS. In contrast to GS, the innovator's profit does not strictly increase with patent life, since a long patent life will encourage imitation (hence competition) before the patent expires. An increase in patent life provides incentives for wasteful imitation but not for productive innovation.

For a given imitation cost K , a sufficiently long patent will attract imitators, resulting in oligopoly pricing instead of monopoly pricing. Conversely, for a given patent life T , a sufficiently narrow scope will attract entrants. Patent life and scope are complementary in that *both* instruments must be increased or reduced to achieve most efficiently the required reimbursement to the innovator.

With imitation, the social cost of a patent may have two components: deadweight loss and the cost of imitation. The optimal patent policy minimizes these costs. Gallini shows that the optimal design is to avoid entry entirely by making the patent broad and short, in contrast to that proposed in GS. That is, the patent should be just long enough to generate the required revenue for the monopolist patentholder, and broad enough to prevent imitation.

However, this reversal depends on an assumption about licensing (or, rather, its absence). In the Gallini model, if the patent is too long or too narrow, the innovator is assumed to sit back passively and watch imitators erode her market share. Maurer and Scotchmer (1998; MS) point out that the duplicative waste could be avoided voluntarily

through licensing rather than by adjusting patent policy, which can again reverse the optimal design. Whatever the market outcome without licensing, the innovator and potential entrants can achieve the same market outcome (price and number of entrants) through a licensing agreement with appropriate royalties and other fees. Since both the innovator and potential entrants can jointly save the imitation costs, they prefer licensing to imitation. The innovator can do even better by fine-tuning the number of entrants.

An important point of agreement among GS, Gallini, and MS is that a narrow patent reduces market price. However, their arguments differ. GS have in mind some sort of regulatory mechanism; Gallini argues that the price reduction will occur through duplicative entry; and MS argue that the price reduction will occur through licensing to prevent duplication. In addition, the analyses of social cost differ, leading to different prescriptions about optimal length and breadth. Since GS do not recognize imitation costs, they simply ask whether the deadweight loss of monopoly pricing is smaller with a long patent and low price, or with a short patent and high price. Gallini argues that if the social cost includes the cost of imitation, the optimal policy should be aimed at avoiding it. MS argue that the imitation costs will not be borne in practice if licensing is available, so that the GS type of analysis is restored.

It is worthwhile expanding on why licensing will lower the market price, by considering what would happen if there were a single potential entrant. The latter situation was analyzed by Gallini (1984), who first pointed out that licensing can prevent entry. With a single potential entrant (or a fixed number), the optimal licensing strategy is to sustain the profit-maximizing (monopoly) price with high royalties, and to share the revenues by using other fees. The licensor has an incentive to keep the market price high regardless of the cost of imitation. In contrast, in the argument above, the licensor is worried about imitation by *nonlicensees* as well as by licensees; there is always an unlicensed potential entrant. The patentholder commits to a low market price precisely to reduce the attractiveness of entry by *nonlicensees*, who can be numerous and unidentifiable *ex ante*. This point stresses the significance of potential entry to the welfare analysis of licensing and, therefore, to the optimal design of IP.

The foregoing discussion shows that private contracting can dramatically alter the optimal design of patents, and that public and private instruments may be complementary in reducing social costs. Patent scope governs the market price in the proprietary market, and licens-

ing prevents wasteful imitation. In this environment where goods are homogeneous, licensing determines the design of patent policy: If licensing is available, a case can be made for narrow and long patents; if licensing is not available, the analysis points to patents that are broad and short.

Licensing may not occur for a variety of reasons, in which case we need a more thorough investigation of the relative merits of the GS and the Gallini arguments, in broader economic environments than they address. Such an analysis has been provided by Denicolò (1996). He explains that narrow (and long) or broad (and short) patents depend on the concavity or convexity, respectively, of the relationship between social welfare and postinnovation profit. Situations in which relatively short broad patents are optimal include costly imitation; Cournot duopoly with constant marginal costs; and horizontally differentiated firms and linear transportation costs, as in Klemperer (1990).

We now turn to cumulative innovation in which subsequent research activity is directed toward the development of improvements or applications of a previous innovation.

IV. Optimal Design: The Case of Cumulative Innovation

In the above discussion, IP is designed for isolated innovations that may be imitated. In reality, research is cumulative. Innovations build upon each other, and subsequent research activity is directed toward improvements or applications of previous discoveries. This fact changes the problem of patent design in interesting and complex ways.

The first and most fundamental complexity, articulated by Scotchmer (1991), is that early innovators lay a foundation for later innovations. The later innovations could not be made without the earlier ones. So that the first innovator has enough incentive to invest, she should be given some claim on profit of the later innovations; otherwise, early innovators could be underrewarded for the social value they create. This is particularly evident in the case of a research tool for which all the social value resides in the innovations it facilitates. If the innovator could not profit from the later products, she would have no incentive to create the tool. The incentive problems are particularly vexed in the case of *creative destruction*, discussed by Schumpeter (1943): an innovator's descendants can actually become the instruments of his destruction.

The Schumpeterian perspective highlights an important problem that arises in the cumulative context: that of dividing the profit between innovators in a way that respects their costs. If, for example, only one pot of money is available for distribution between two innovators and most is allocated to the first firm, the second inventor's incentive for research is reduced, and vice versa. Green and Scotchmer (1995) argue that because of the difficulties in dividing profit, patent lives will have to be longer than if the whole sequence of innovations occurred in a single firm. Ex ante licensing—licensing before investments are made—is a way of mimicking the latter outcome. As in the case of a single invention, the availability of private contracting influences the optimal patent scope when innovation is cumulative (see below).

Cumulativeness changes the design instruments that are relevant to the length of protection. The statutory life can be irrelevant when a noninfringing substitute, such as an improvement, can displace a protected product. What matters is the *effective* life, that is, the time until the noninfringing substitute appears (Scotchmer 1991; O'Donoghue, Scotchmer, and Thisse 1998 (OST)). The effective life is determined by patent *scope* or *leading breadth*, which is interpreted as the minimum quality improvement that avoids infringement. As in the case of costly imitation discussed above, the effectiveness of patent life as an instrument for R&D may be limited when subsequent innovation can undermine profitability.

Finally, cumulativeness makes a third instrument—the minimum standard for protection, or minimum inventive step—relevant to the optimal design of IP. For copyrighted works the standard for protection is low (as is the breadth of protection), while for patents, the patentability standard (or novelty requirement) can be quite stringent. In our discussion of isolated inventions above, we assumed that the invention was protectable, since there would be no incentives to innovate if there were no IP or other incentive instruments. But in the cumulative context, patentability on second-generation inventions is less essential, since an innovation can be protected by an exclusive license on a previous patent it infringes, rather than by its own patent. Leading breadth and the standard for patentability together determine the level of "forward protection" each innovation has.

Several arguments favoring both weak and strict standards for IP protection have been advanced. Scotchmer and Green (1990) argue with caution for a weak standard (a weak "novelty requirement"), so that firms are encouraged to disclose every small bit of progress. While

these disclosures could speed up invention by giving a technological boost to competitors, Scotchmer and Green warn that the weak novelty requirement could also encourage firms to choose trade secrecy over patents. In contrast, a tightening of the standards for patentability can encourage firms to be more ambitious in the improvements they attempt to develop (O'Donoghue 1998) or can direct their investments toward more socially useful inventions (Eswaran and Gallini 1996). Even when the standard for protection does not reorient research efforts, it can affect the division of profit among sequential researchers. Scotchmer (1996) argues that the strictest novelty requirement (no protection) on second-generation products would tilt the joint profit of a sequence of innovations in favor of earlier innovators without jeopardizing second-generation advances. A second-generation product can be protected by an exclusive license on the infringed patent of the earlier generation. Denicolò (2000a, b) makes a case for a patent policy with a weak patentability standard and narrow leading breadth. In a model in which firms race for the first- and second-generation patents, he shows that tilting profits in favor of earlier innovators might only encourage a socially wasteful patent race at the stage of basic research and underinvestment in the second stage.

Although the complexities of cumulateness seem to defy clear, unqualified design implications, one lesson is clear: The optimal design of IP depends importantly on the ease with which rights holders can contract around conflicts in rights. Contracting is especially relevant to the question of breadth, which determines the likelihood that a follow-on innovation will infringe a prior patent.

A danger of IP that has been debated from its inception to the present (see Machlup and Penrose 1950) is that IP can stifle innovation and slow progress. Merges and Nelson (1990) link this danger to breadth, using examples from the aircraft, radio, and pharmaceutical industries to argue for narrow patents. An earlier example concerned steam engines. James Watt refused to license his patents for improvement, with the result that there was a flood of pent-up invention when his patents expired in 1800 (Derry and Williams 1993, p. 324).

In contrast, Kitch (1977) argues that broad patents are socially beneficial precisely *because* they stimulate further developments. Scotchmer (1991) and Green and Scotchmer (1995) take the same point of view, but focus on how *ex ante* contracting affects division of profit. With *ex ante* contracting, the role of breadth is not to determine whether subsequent products are made (they will be made if they add to joint profit), but rather to determine how the profit is divided. This

theme is also carried forward in later papers, e.g., Merges (1998, 1999), Scotchmer (1996), Lemley (1997) (who compares how copyright and patent doctrines respectively treat the possibility of blocking), O'Donoghue, Scotchmer, and Thisse (1998), and Schankerman and Scotchmer (2001). Matutes, Regibeau, and Rockett (1996) and Chang (1995) argue for broad patents even without assuming that *ex ante* contracts can be made.

To some extent, broad patents are also supported by the arguments of O'Donoghue, Scotchmer, and Thisse (1998), who study breadth in a model with with an infinite sequence of improved products (quality ladder). If patents are relatively narrow, the effective life of each patent ends when a noninfringing improvement arrives, and is thus endogenous. But if the patent is broad, then the statutory life is also the effective life: Every subsequent innovation on the quality ladder infringes during the statutory life and must be marketed under license. To achieve the same rate of progress under both regimes, the effective patent life with a narrow patent must be longer than the (effective) statutory life with a broad patent. Broad, short patents are more efficient at rewarding innovators along the quality ladder, because less of the total profit in the system accrues to high-value innovations that would be made in any case, and more goes to the innovators who need additional incentives.

Thus, with some caution, we can extract from the literature a case for broad (and short) patents. Broad patents can serve the public interest by preventing duplication of R&D costs, facilitating the development of second-generation products, and protecting early innovators who lay a foundation for later innovators. However, these benefits disappear if licensing fails. Heller and Eisenberg (1998) argue that licensing will likely fail when researchers must negotiate multiple licenses, as now occurs in the biomedical industry. Mazzoleni and Nelson (1998) caution that these transaction costs may limit the use of contracts for coordinating innovations that follow from a broad patent.

Another problem with licensing is that it can lessen competition both in innovation markets⁹ and in product markets. It thus raises antitrust issues, even in the simpler context where there is no cumulative aspect. One of the difficult issues is determining whether an *ex ante* merger of research activities through licensing is efficient or inefficient from a social perspective. On the efficiency side, *ex ante* licensing can enable firms to avoid duplicated costs and to delegate efficiently, much as discussed in section II. But on the inefficiency side, *ex ante* licensing can retard progress, e.g., by nullifying the acceleration that would other-

wise come from a patent race. See Gilbert and Sunshine (1995) for a discussion of these issues.¹⁰

The cumulative context raises another issue. Above we focused on the salutary effects of licensing, namely that *ex ante* licensing can ensure investment in infringing follow-on products that would add to joint profit. Turning this argument on its head, licensing can stifle noninfringing follow-on products that would *detract* from joint profit. Gallini and Winter (1985) analyze a situation where a potential competitor is licensed *ex ante* in order to dissuade him from investing in a noninfringing cost reduction that would have lowered prices in the market. Such licensing clearly reduces product market competition relative to what the Congress apparently intended in designing patent law. If such licensing occurred *ex post* to prevent production of the cost-reducing innovation after it had been developed, it would presumably be an antitrust violation. Chang (1995) analyzes precisely that type of *ex post* collusion and advocates a strict antitrust rule against collusion. For a discussion of how principles of competition policy might be formulated to distinguish *ex ante* licensing that is procompetitive from that which is anticompetitive, see Scotchmer (1998).

Besen and Maskin (2000) argue that if firms do not license in a way that takes full advantage of their IP, e.g., because of antitrust restrictions, then licensing may reduce industry profits below those available without licensing, and the broad patents that support such licensing are counterproductive. In this sense, Besen and Maskin's paper is consistent with the above observation that impediments to contracting may strengthen the case for narrow patents.

In light of these qualifications, what conclusions can we make for patent design in the cumulative context? One interpretation is that, when research is cumulative, relatively broad patents may be efficient if *ex ante* contracting is available. However we prefer to be cautious; the jury is still out.

What is conclusive is the importance of private contracting. Whether property rights are helpful or counterproductive in encouraging innovation depends on the ease with which innovators can enter into agreements for rearranging and exercising those rights, as constrained by the rules of antitrust law.

V. Conclusions

In the past two decades, academic interest in the economics and law of intellectual property has exploded. The renewed interest has been

fueled by controversies surrounding new technologies, by international agreements, and by changes in the nature of protection, e.g., see Mazzoleni and Nelson (1998). It is generally thought that IP rights have been strengthened, but there is also evidence that some forms of IP, in particular, patents, have previously been ineffective (Cohen et al. 2000). Contrary to the apparent intent, strengthening of IP is thought by some commentators to impede research rather than to promote it (Heller and Eisenberg 1998). In this environment, economists have had much to say about both the optimal design of IP and the advisability of substituting other incentive mechanisms.

Although it comes as no surprise that a property system has defects, we hope we have illuminated some offsetting virtues, and some circumstances where other mechanisms, such as prizes, fixed-price contracts, and auctions can dominate. Our main conclusions on the effectiveness of intellectual property are that:

1. IP is probably the best mechanism for screening projects when value and cost are not observable by the sponsor, since the private value of IP reflects the social value, and firms automatically compare some measure of value with the cost of innovation. In addition, IP encourages firms to accelerate progress, since the reward is conditional on success. Prizes could serve the same purposes if the size of the prize could be linked to the social value but without the deadweight loss of monopoly pricing.
2. Neither IP nor prizes can aggregate the information that is decentralized among firms, and neither will be completely effective at delegating research effort efficiently. A procurement system that restricts prizes to certain firms, or differentiates prizes according to firms' relative efficiencies, can improve on a simple prize system or patent system, but then there must be an *ex ante* negotiation to select the favored firms.

For circumstances where IP is justified, we asked how the property right should be designed. Every IP regime has provisions on length, breadth, and the standard for protection. The economics literature on design of IP concerns the appropriate choice of these provisions. The optimal length, breadth, and standard for protection depend on the economic environment, e.g., the shape of the demand curve, the rate at which improvements to existing technologies are developed, or the relative costs of sequential innovators.

How much flexibility is there in designing IP rights differently for different economic environments? In fact, there is a lot of flexibility. Different IP regimes are targeted at different subject matter, and the subject matter is an important defining aspect of the IP regime. Copyright has traditionally been targeted at literature, other printed matter, and art. Patents have traditionally been targeted at manufactured items. The subject matters of *sui generis* laws typically have been very specific, e.g., the Plant Patent Act, the Plant Variety Protection Act, the Semiconductor Chip Protection Act, and the proposed database legislation.

The IP regimes that cover different subject matter are noticeably varied in the three important features: length, breadth, and standard for protection. On the matter of length, copyrights last essentially forever, patents last 20 years, and chip protection lasts 10 years. On the matter of breadth, copyright protection is restrained by fair use exemptions and by the fact that the underlying "ideas" are not protected; patents have the doctrine of equivalents; and copying of chips is allowed for some uses but not others. We thus believe that it is incorrect to criticize the economic design arguments on grounds that, in IP, "one size fits all." While we do not think it would be appropriate to define new IP regimes for every small category of technology, we wish to emphasize that the Congress can exercise as much flexibility as it wishes, and that courts also have some flexibility.

Each IP regime should cover subject matter with similar needs for protection, especially if heterogeneous needs cannot be remedied by courts. Many controversies arise because of heterogeneity within IP regimes. For example, business methods probably do not need the strong protection provided by the Patent Act, even though such protection is appropriate for other patentable subject matter. A new regime could have been created for business methods, but protection under the Patent Act could alternatively be weakened through the courts' interpretation of novelty and nonobviousness.

Finally, there are the design recommendations themselves. We have not been specific in this review about the exact ways in which length, breadth, and standards for protection should reflect the economic environments, and refer the reader to the underlying papers for more detail. Instead, we have emphasized a message of a different sort: the optimal design of the property right should depend on whether firms contract with others for the use of their protected innovations. With fluid contracting, policies that otherwise would be inefficient may be

optimal. For example, licensing can avoid wasteful imitation, making an otherwise inefficient narrow patent optimal. In the cumulative context, there is a danger that broad patents will inhibit future innovators from making product improvements. But with contracting, the patentholder can profit from, instead of being threatened by, new improved products, and will ensure that they arise even if infringing. The most striking message of the literature is that IP and private instruments may be complementary in reducing social costs from an overreaching or insufficient protection regime.

However, contracting also has the potential to undermine competition in ways that were not anticipated or approved by the Congress when designing IP. Contracting that we have not covered includes cross-licensing and patent pools. We have also not discussed joint ventures and other alliances for avoiding litigation, duplicated efforts, and holdups. A recurring theme, especially evident in these contexts, is that despite the efficiencies that contracting can ensure, contracting may also facilitate anticompetitive behavior. See Hall and Ziedonis (2001), Shapiro (2000), Denicolò (2000a, b). To understand whether the property system is too strong, too weak, or necessary at all requires us to understand the incentives for contracting, and its potential anticompetitive consequences.

Notes

This article will also appear in *Legal Orderings and Economic Institutions*, F. Cafaggi, A. Nicita and U. Pagano, eds., Routledge Studies in Political Economy (2002).

1. For example, in the environment discussed by Wright (1983), none of the three mechanisms is optimal. The first best can be achieved with a mechanism similar to the one mentioned in note 3 below.
2. Prizes might also require enforcement. John Harrison's longitude prize was delayed for decades while the prize committee attempted to prove that astronomical solutions were superior to his clock. Harrison eventually sought redress in Parliament, and was partially rewarded.
3. She suggests a two-part procedure. The sponsor first asks the firms to reveal their information on value and then, if warranted, employs the best procurement mechanism to delegate to the least-cost firm(s). Following Cremer and McLean (1988), it is costless to get the firms to reveal their correlated information on value. They are asked to report their signals of value, and then rewarded if they agree and punished if they disagree. Due to the correlation, an equilibrium is to report truthfully, and the payments can be chosen so that each firm makes zero expected profit.
4. A related problem is studied by Bhattacharya et al. (1998). Instead of assuming that firms have different efficiency parameters, they assume that firms have different *knowl-*

edge about the cost of achieving an innovation. If the knowledge is revealed, then all firms have the same cost. Their mechanisms also use payments conditional on delivery (prizes).

5. We caution, however, that the Kremer scheme is efficient only if there is a single researcher. A prize equal to the social value could easily attract other firms to a race in which the firms overinvest (Loury 1979). Not only is there a problem of overinvesting, but inefficient firms as well as efficient firms may invest. This is the problem avoided by the more complex procurement mechanism discussed above, where prizes are tailored to the firms' relative efficiency in order to make sure that the investment effort is undertaken by the more efficient ones.

6. In deLaat's model, the sponsor chooses the "size of the invention," which is observable, given the firm's report of the market size (value), which is unobservable to the sponsor. DeLaat assumes that the sponsor can verify which invention is made, but not the market conditions (e.g., demand) for the invention.

7. For recent empirical investigations of how firms exercise their option to renew, and implications for the values of innovations, see Lanjouw (1997) and Schankerman (1998).

8. The intuition for this result can be found in the familiar economic principle that underlies Ramsey pricing. Ramsey pricing solves the problem of maximizing consumer surplus in multiple markets subject to the constraint that revenues cover cost. The solution is to set prices below monopoly prices so that the markup of price in each market is inversely proportional to the elasticity of demand in each market. In the patent problem, the different time periods are parallel to different markets, and since the demands are assumed to be constant over time, the markup of price over cost in each period is identical.

9. See the U.S. Department of Justice–Federal Trade Commission *Antitrust Guidelines for the Licensing of Intellectual Property* (1995) for a discussion of innovation markets.

10. One of the thorny questions that arise is whether competition policy should view licensing practices more leniently than otherwise if incentives to innovate are at stake. See Gallini and Trebilcock (1998) for a discussion of this issue.

References

Anton, J., and D. Yao. 1994. "Expropriation and Inventions—Appropriable Rents in the Absence of Property Rights." *American Economic Review* 84(1): 191–209.

Arrow, K. J. 1962. "Economic Welfare and the Allocation of Resources for Invention," in Universities-National Bureau of Economic Research Conference. Series, *The Rate and Direction of Economic Activities: Economic and Social Factors*. Princeton: Princeton University Press.

Besen, James, and Eric Maskin. 2000, January. "Sequential Innovation, Patents and Imitation." Working Paper no. 00-01. MIT.

S. Battacharya, D'Aspremont, C., and L.-A. Gerard-Varet. 1998. "Knowledge as a Public Good: Efficient Sharing and Incentives for Development Effort." *Journal of Mathematical Economics* 30: 389–404.

Calandrillo, Steve. 1998. "An Economic Analysis of Intellectual Property Rights: Justifications and Problems of Exclusive Rights, Incentives to Generate Information, and

the Alternative of a Government-Run Reward System." *Fordham Intellectual Property Media and Entertainment Law Journal* 9: 301–348.

Cohen, Wesley M., Richard R. Nelson and John P. Walsh (2000), "Protecting their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)", National Bureau of Economic Research Working Paper 7552.

Chang, H. 1995. "Patent Scope, Antitrust Policy, and Cumulative Innovation." *RAND Journal of Economics* 26: 34–57.

Cornelli, F., and Mark S. Schankerman. 1999. "Patent Renewals and R&D Incentives." *RAND Journal of Economics* 30(2).

Cremer, J., and R. P. McLean. 1988. "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions." *Econometrica* 56(6): 1247–57.

De Laat, Eric A. A. 1996. "Patents or Prizes: Monopolistic R&D and Asymmetric Information." *International Journal of Industrial Organization* 15(3): 369–90.

Denicolò, V. 1996. "Patent Races and Optimal Patent Breadth and Length." *Journal of Industrial Economics* 44: 249–265.

Denicolò, V. 2000a. "Two Stage Patent Races and Patent Policy." *RAND Journal of Economics* 31: 488–501.

Denicolò, V. 2000b. "Combination of Competing Patents and Antitrust Policy." Working paper. Department of Economics, University of Bologna.

Derry, T. K., and Williams, Trevor I. 1993. *A Short History of Technology: From Earliest Times to A.D. 1900*. Mineola, NY: Dover.

Eswaran, M. and N. Gallini. 1996. "Patent Policy and the Direction of Technological Change," *RAND Journal of Economics* 27(4): 722–746.

Farrell, J. 1995. "Arguments for Weaker Intellectual Property Protection in Network Industries." In Brian Kahin ed., *Standards Policy for Information Infrastructure*. 368–377. Cambridge, MA: MIT Press.

Gallini N. T. 1984. "Deterrence by Market Sharing: A Strategic Incentive for Licensing." *American Economic Review* 74: 931–41.

Gallini, N. T. 1992. "Patent Length and Breadth with Costly Imitation." *RAND Journal of Economics* 44: 52–63.

Gallini, N., and M. Trebilcock. 1998. "Intellectual Property Rights and Competition Policy—A Framework for the Analysis of Economic and Legal Issues." In R. Anderson and N. Gallini, eds., *Competition Policy and Intellectual Property Rights in the Knowledge-Based Economy*. Industry Canada Research Series, University of Calgary Press: 17–64.

Gallini, N. T., and R. Winter. 1985. "Licensing in the Theory of Innovation." *RAND Journal of Economics* 16: 237–52.

Gandal, N., and S. Scotchmer. 1993. "Coordinating Research through Research Joint Ventures." *Journal of Public Economics*, 51: 173–93.

Gilbert, R., and C. Shapiro. 1990. "Optimal Patent Length and Breadth." *RAND Journal of Economics* 21: 106–12.

Gilbert, R., and G. C. Sunshine. 1995. "Incorporating Dynamic Efficiency Concerns in Merger Analysis: The Use of Innovation Markets." *Antitrust Law Journal* 63: 569–602.

- Green, Jerry, and Suzanne Scotchmer. 1995. "On the Division of Profit in Sequential Innovation." *RAND Journal of Economics*, 26: 20–33.
- Hall, B. H. and R. Ham Ziedonis. 2001. "The Patent Paradox Revisited: An Empirical Study of Patenting in the U.S. Semiconductor Industry, 1979–1995," *RAND Journal of Economics* 32(1): 101–128.
- Heller, Michael A., and R. S. Eisenberg. 1998. "Can Patents Deter Innovation? The Anticommons in Biomedical Research." *Science* 280: 698–701.
- Kitch, E. W. 1977. "The Nature and Function of the Patent System." *Journal of Law and Economics* 20: 265–90.
- Klemperer, P. 1990. "How Broad Should the Scope of Patent Protection Be?" *RAND Journal of Economics* 21: 113–30.
- Kremer, Michael. 1989. "Patent Buyouts: A Mechanism for Encouraging Innovation." *Quarterly Journal of Economics* 113: 1137–67.
- Kremer, Michael. 2000. "Creating Markets for New Vaccines." Conference Paper, Innovation Policy and the Economy. Cambridge, MA: National Bureau of Economic Research.
- Laffont, J.-J., and J. Tirole. 1986. "Using Cost Observation to Regulate Firms." *Journal of Political Economy* 94: 614–41.
- Laffont, J.-J., and J. Tirole. 1987. "Auctioning Incentive Contracts." *Journal of Political Economy* 95: 921–937.
- Lanjouw, J. 1998. "Patent Protection in the Shadow of Litigation: Simulation Estimations of Patent Value." *Review of Economic Studies*. Vol 65, 671–710
- Lemley, M. A. 1997. "The Economics of Improvement in Intellectual Property." *Texas Law Review* 75: 989.
- Loury, G. C. 1979. "Market Structure and Innovation." *Quarterly Journal of Economics* 93: 395–410.
- Machlup, Fritz, and Edith Penrose. 1950. "The Patent Controversy in the Nineteenth Century." *Journal of Economic History* 10(1): 2.
- Matutes, C., P. Regibeau, and K. E. Rockett. 1996. "Optimal Patent Protection and the Diffusion of Innovation." *Rand Journal of Economics* 27: 60–83.
- Maurer, S. 1999. "Raw Knowledge: Protecting Technical Databases for Science & Industry." Report commissioned by the National Academy of Sciences. www.nap.edu/html/proceedings_sci_tech/appC.html
- Maurer, S. M., and S. Scotchmer. 1998. "The Independent Invention Defense in Intellectual Property." John M. Olin Working Paper no. 98-11. Boalt School of Law, University of California, Berkeley.
- Maurer, S. M., and S. Scotchmer. 1999. "Database Protection: Is it Broken and Should We Fix it?" *Science* 284: 1129–1130.
- Mazzoleni, Roberto, and Richard R. Nelson. 1998. "The Benefits and Costs of Strong Patent Protection: A Contribution to the Current Debate." *Research Policy* 27: 273–84.
- Merges, R. P. 1998. "Antitrust Review of Patent Acquisitions: Property Rights, Firm Boundaries, and Organization." In Robert D. Anderson and Nancy T. Gallini, eds., *Com-*

petition Policy and Intellectual Property Rights in the Knowledge-Based Economy. Calgary: University of Calgary Press.

Merges, R. P. 1999. "Intellectual Property Rights, Input Markets and the Value of Intangible Assets." Working Paper. Berkeley: University of California.

Merges, R. P., and R. R. Nelson. 1990. "On the Complex Economics of Patent Scope." *Columbia Law Review* 90(4): 839–916.

Minehart, D., and S. Scotchmer. 1999. "Ex Post Regret and the Decentralized Sharing of Information." *Games and Economic Behavior* 27(1): 114–131.

National Research Council. 2000. *The Digital Dilemma: Intellectual Property in the Information Age.* Committee on Intellectual Property Rights and the Emerging Information Infrastructure, Computer Science and Telecommunications Board. Washington DC: National Academy Press.

National Science Foundation. 2000. *Science and Engineering Indicators 2000.* www.nsf.gov/sbe/srs/seind00/frames.htm

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

O'Donoghue, T. 1998. "A Patentability Requirement for Sequential Innovation." *Rand Journal of Economics* 29: 654–67.

O'Donoghue, T., S. Scotchmer, and J. F. Thisse. 1998. "Patent Breadth, Patent Length and the Pace of Technological Progress." *Journal of Economics and Management Strategy* 7: 1–32.

Reynolds, T. 1983. *Stronger than a Hundred Men: A History of the Vertical Water Wheel.* Baltimore: Johns Hopkins Press.

Schankerman, M. 1998. "How Valuable is Patent Protection? Estimates by Technology Field." *RAND Journal of Economics* 29: 77–107.

Schankerman, M., and S. Scotchmer. 2001. "Damages and Injunctions in Protecting Intellectual Property." *RAND Journal of Economics* 32: 199–220.

Schumpeter, J. 1943. *Capitalism, Socialism and Democracy.* London: Unwin University Books.

Scotchmer, S. 1991. "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law." Symposium on Intellectual Property Law. *Journal of Economic Perspectives*, 5(1): 29–41.

Scotchmer, S. 1996. "Protecting Early Innovators: Should Second-Generation Products Be Patentable?" *RAND Journal of Economics* 27: 322–31.

Scotchmer, S. 1998. "R&D Joint Ventures and other Cooperative Arrangements." In *Competition Policy and Intellectual Property Rights in the Knowledge-Based Economy*, R. Anderson and N. Gallini, eds. Industry Canada Research Series. University of Calgary Press: 203–22.

Scotchmer, S. 1999a. "On the Optimality of the Patent Renewal System." *RAND Journal of Economics* 30, 181–196.

Scotchmer, S. 1999b. "Delegating Effort in a Common Value Project." IBER Working Paper no. E99-266. Berkeley: University of California.

Scotchmer, S., and J. Green 1990. "Novelty and Disclosure in Patent Law." *RAND Journal of Economics* 21: 131–46.

Shapiro, C. 2000. "Navigating the Patent Thicket: Cross Licenses, Patent Pools, and Standard-Setting." Working Paper. Berkeley: University of California Competition Policy Center. Working Paper No. CPC00-11. Now published in *Innovation Policy and the Economy*, Vol. 1, MIT Press, 2001. Adam B. Jaffe, Josh Lerner, and Scott Stern, eds.

Shavell, S., and T. van Ypersele 1998. "Rewards versus Intellectual Property Rights." Olin Discussion Paper no. 246. Harvard University.

Sobel, D. 1995. *Longitude: The True Story of a Lone Genius Who Solved the Greatest Scientific Problem of His Time*. New York: Walker.

U.S. Department of Justice and Federal Trade Commission. 1995. *U.S. Department of Justice—Federal Trade Commission Antitrust Guidelines for The Licensing of Intellectual Property*.

Wright, B. 1983. "The Economics of Invention Incentives: Patents, Prizes and Research Contracts." *American Economic Review*, 73, 691–707.

