ISSN 1045-6333

PATENT SCOPE, ANTITRUST POLICY, AND CUMULATIVE INNOVATION

Howard F. Chang*

Discussion Paper No. 96

9/91

Program in Law and Economics
Harvard Law School
Cambridge, MA 02138

The Program in Law and Economics is supported by a grant from the John M. Olin Foundation.

*J.D., Harvard Law School and former John M. Olin Fellow.

PATENT SCOPE, ANTITRUST POLICY, AND CUMULATIVE INNOVATION

Howard F. Chang*

September 15, 1991

*Department of Economics, Massachusetts Institute of Technology

I wish to thank Franklin Fisher, Oliver Hart, Rebecca Henderson, Louis Kaplow, Andrea Shepard, Hilary Sigman, and Jean Tirole for helpful comments on earlier drafts, and the John M. Olin Foundation for financial support.

Howard F. Chang*

ABSTRACT

This paper presents a formal model of cumulative innovation to investigate what factors should influence a court's decision when a patentee alleges that an improved version of the patented product infringes the patent. The model reveals how the optimal patent policy, in order to promote both basic innovations and improvements that build upon such innovations, would limit the scope of patent protection for a basic innovation so that those improvements that are most valuable (among all possible improvements on that basic innovation) would be deemed not to infringe the patent. Most important, the model also indicates that the optimal patent policy would extend broad protection not only to those patents that are very valuable (standing alone) relative to the improvements that others may subsequently invent, but also to those that have very little stand-alone value relative to most possible improvements. Patents for such inventions should be permitted wide scope in the sense that only the most extraordinary improvements should be permitted to avoid infringement.

The paper also examines a related question: whether the courts should interpret the antitrust laws so that the holder of a patent and competing inventors with improved versions of the patented product can enter collusive agreements and thereby avoid competition that would erode the profits that encourage innovation. The model provides only limited support for such a permissive antitrust policy, because the policy would create incentives for inefficient entry by imitators who "invent around" the original patent. The optimal policy would permit collusion only between holders of basic patents with little stand-alone value and imitators who contribute relatively valuable improvements, and even in these cases the analysis suggests that the benefits from such collusion may be small.

*J.D., Harvard Law School and former John M. Olin Fellow.

The patent authorities, that is, the Patent Office and the courts, are constantly called upon to decide the proper scope of patents. Courts frequently must decide whether a second invention infringes upon the patent of a preceding invention. These decisions have important effects on the pace of technological progress through the incentives to invent not only the first invention but also later inventions that build upon the first. These effects are critical in the design of the optimal patent policy, because the dynamic losses that flow from a slower pace of innovation can easily swamp the static deadweight losses from patent monopolies. Nevertheless, until recently scholars have devoted little attention to the scope of patent protection and the issue of cumulative innovation.

This paper presents a formal model of cumulative innovation to determine if (and how) the scope of patents should be limited so as to encourage the invention of valuable improvements on the patented product or process. Despite some tension with the general principles of the patent system, courts have declined to find infringement when the patentee complains of a competing product that features a significant improvement over the patented product. Accordingly, this paper investigates how such an improvement should influence a court's decision in an infringement case.

The results suggest that, in light of the objective of promoting cumulative research, courts should curtail patent protection for an basic innovation so as to allow later innovators to develop improvements that are particularly valuable (among all possible improvements on that basic innovation) without violating the patent. Thus, the formal model provides a basis in economic theory for the general practice of declaring improved products not infringements of patented predecessors. The analysis, however, also implies that only the most valuable improvements should avoid infringement.

Most important, the formal model also suggests a rule for taking the value of these inventions into account that differs from that proposed by Merges and Nelson (1990, pp. 865-66), who suggest that courts decline to find infringement when the value of the original invention is small relative to the value of the improvement. Instead, courts should extend the broadest protection not only to the patent with very large stand-alone value relative to all possible subsequent improvements, but also to the patent with <u>very little</u> stand-alone value relative to the improvements that it may inspire. Thus, although the basic patent by itself may have almost no commercial value, the courts should allow the original patentee to claim rights over nearly all the improvements built upon the patented technology. Patents for such basic inventions should be permitted wide scope in the sense that only the most extraordinary improvements should be permitted to avoid infringement.

The paper also examines a related question: whether the courts should interpret the antitrust laws so that the holder of a patent and competing inventors with improved versions of the patented product can enter collusive agreements and thereby avoid competition that would erode the profits that encourage innovation. This paper evaluates this issue in a model designed to provide a synthesis of some of the opposing considerations raised in the preceding literature. Whereas Scotchmer (1991) and Green and Scotchmer (1990) suggest that courts should permit collusive licensing between competing patentees, Kaplow (1984) has argued that such an antitrust policy rewards innovation only at an excessive social cost. The analysis seeks to reconcile the views of Kaplow and of Green and Scotchmer regarding the antitrust treatment of collusion between competing innovators.

The results provide only limited support for a permissive antitrust policy regarding collusion between competing patentees, even in the context of cumulative innovation addressed by Green and Scotchmer. Such a policy would create incentives for inefficient entry by imitators who "invent around" the original patent. The optimal policy would permit collusion only between holders of basic patents with little commercial value and imitators who contribute relatively valuable improvements, and even in these cases the analysis suggests that the benefits from such collusion may be small.

Part I of this paper sets forth the issues that the formal model will address. First, Part I.A reviews the relevant economic theory developed in the preceding literature by economists and legal scholars. Next, Part I.B describes the law on patent scope, and Part I.C surveys the antitrust cases regarding patent licensing and patent acquisition. Part II sets forth a formal model of cumulative innovation and lays out the underlying assumptions. Part III presents the derivation of the optimal innovation policy. Part III.A a derives an optimal policy that corresponds to a compulsory licensing scheme with courts setting patent royalties. This policy proves to be a useful benchmark for comparison to the policies studied next. In Part III.B, the courts can only influence the inventors' incentives through decisions on patent scope and antitrust law. The optimal policy bases the infringement and antitrust decisions on the values of the basic invention and of the improvement that is alleged to infringe the basic patent. Part IV concludes with implications for policy from the results and suggests directions for further research and possible extensions of the model.

I. Background

A. Economic Theory

In an environment in which the regulatory authorities cannot observe research and development (R&D) costs, authorities seeking to encourage investment in R&D ex ante if and only if such efforts are socially desirable should set the inventor's private reward ex post equal to the social benefits of the invention. A system of patents cannot achieve this equality ex post unless firms can engage in perfect price discrimination and thereby appropriate all consumer surplus. Furthermore, in the absence of perfect price discrimination, patents impose the familiar ex post deadweight loss that results from monopoly pricing: such prices exclude some potential users of the invention that are willing to pay more than the marginal cost of production.

This paper focuses on one particular example of the ex post social costs of patents: the inhibition of R&D by other inventors with ideas for improvements in the patented product or process. Once another inventor has invested in developing an improvement, the original patentee may use its monopoly over its innovation to appropriate some of the value created by such a complementary innovation, even if the second innovator patents the improvement. If the second innovator can market its invention only with the consent of the original patentee, that patentee can increase its profits at the expense of the second innovator by bargaining to license the complementary technology at less than its full value. This "holdup" problem reduces R&D in complementary technologies by other inventors by reducing the expected return on their investment.[1]

As in the context of vertical integration and contracts discussed by Klein, Crawford, and Alchian (1978) and by Williamson (1979), opportunistic appropriation ex post (in this context, after R&D costs are sunk) discourages investments in assets (innovations) that have their highest value when used in conjunction with an asset (the original patent) held by another party. This underinvestment results whenever contracts are incomplete in that the parties cannot agree ex ante to specific levels of investment (because a court cannot observe levels ex post, for example) or to a particular division of the surplus generated by the investment. A complete contract ex ante could encourage the first-best level of investment by ensuring that the party investing receives an adequate return on its investment.

If the first innovation is a necessary condition for the second innovation, then the surplus from the second innovation is a joint product of both innovations. In this case, the holdup problem is bilateral: to the extent that the second innovator receives positive profits, then the first innovator will tend to underinvest in R&D relative to the first-best. As Scotchmer (1991, p. 34) and Green and Scotchmer (1990) observe, in markets with cumulative innovation, patent protection cannot offer both the first innovator and the second innovator the full surplus from the second innovation. As a result, some distortion of incentives is unavoidable under a patent system: at least one party will have too little incentive to invest in innovation.

Faced with such a bilateral holdup problem, the parties to an incomplete contract

[1]This particular social cost of patents merits special study because it is likely to have important implications for policy. The "holdup" problem is important, because the costs of a small dynamic effect can rapidly accumulate to dwarf the costs of a static effect such as the deadweight loss from monopoly pricing. See Scherer (1980, p. 407). The policy implications of the holdup problem will differ from those of the deadweight loss problem, because various policy instruments will have different effects on these two problems. For example, lax antitrust regulation of horizontal agreements may increase the static costs of patents, but as discussed below, it may decrease the dynamic costs of patents by encouraging subsequent innovation. On the other hand, whereas price discrimination ex post may reduce the static costs of patents, it cannot solve the "holdup" problem unless the original patentee can commit ex ante to reward a subsequent inventor ex post.

ex ante can allocate residual rights (i.e., ownership) to one party or the other before any investments are made and, as Grossman and Hart (1986) show, thereby influence any renegotiation to divide the surplus ex post. By setting the "threat point" in such ex post bargaining, the contract can trade off distortions in one party's ex ante incentives against those in the ex ante incentives of the other. Similarly, courts can affect ex ante incentives by allocating intellectual property rights: they decide how different a second invention must be from a prior patented invention to avoid infringement ex post. Green and Scotchmer observe that by adjusting patent scope in this manner, courts can shift surplus between the inventors ex post (by affecting their bargaining over a licensing agreement, for example) and thereby seek to minimize the distortion of R&D investment decisions ex ante.[2]

In particular, Merges and Nelson (1990) assert that the holdup problem posed by a finding of infringement becomes most acute when the value of the improvement embodied in the infringing invention is large relative to the value of the original invention. Therefore, Merges and Nelson (pp. 865-66) suggest that courts should find infringement when "the original invention contributes most of the value" of the improved invention, but find no infringement when "the original patent contributes very little value compared to the improvement." The model in this paper is designed to shed light on these recommendations regarding the proper criteria for such patent scope decisions.

The courts can affect not only the division of profits between innovators through decisions on patent scope, but also the division of surplus among consumers and innovators through antitrust policy. Kaplow (1984) addressed the inherent tension between antitrust and patent policies: antitrust regulation of the practices of patentees restrict their ability to exploit their market power. While such restrictions may reduce the static deadweight loss associated with patents, they also reduce the returns to the patentee's investment in

[2]Copyright protection presents similar problems, because it raises the costs of creating new works that borrow or build upon material from a prior work. Thus, Landes and Posner (1989) and Menell (1989, pp. 1079-88) analyze the scope of copyright protection as an instrument with which to balance these costs against the benefits flowing from enhanced incentives for the original work.

research and development. Kaplow observed that the optimal antitrust policies regarding various patentee practices should be determined simultaneously with each other and with other aspects of patent policy.

For example, horizontal agreements regarding patents pose an antitrust problem because they can be used by competing firms to collude. If patent licensing agreements may contain restrictions on the prices to be charged in the product market, then any patent may form the basis for a price-fixing cartel.[3] Kaplow (pp. 1855-60) concludes that courts should interpret the antitrust laws to prohibit price-restricted licenses. To protect its profits from erosion without any price restrictions, a patentee may instead raise the price charged by licensees through a royalty based on units of output by the licensee.[4]

If multiple firms that possess competing patents may cross-license their innovations to one another with price restrictions, they can also cartelize an industry. Even if antitrust law were to prohibit such collusive cross-licensing among competing patentees, however, these firms can achieve a similar result by selling their competing patents to one party.[5] Thus, Priest (1977, pp. 358-64) and Kaplow (pp. 1867-73) address the issue of horizontal combinations among competing patentees and conclude that such practices also should be

---

[3]One firm could license its innovation (even if the innovation is of little economic value) to all other firms in the industry and include a price floor in the licensing agreements. The patentee in such a case could realize profits far greater than the social value of the invention. Such large rewards for trivial inventions would misallocate R&D resources through perverse and excessive incentives.

[4]Competition between the patentee and its licensee would drive the licensee's prices toward the sum of the licensee's marginal costs and its per-unit royalty. In this case, the reward obtained through such a royalty cannot exceed the value of the patented innovation, because it cannot exceed what the licensee is willing to pay for it (the increase in the licensee's revenues attributable to the innovation).

[5]Suppose that a firm must possess a patent to sell in the relevant market and that each firm's patent is a perfect substitute for the others, so that no firm would be willing to license any patent from any other in the absence of price restrictions. Whereas product-market competition among the patentees would drive the profits of each toward zero, by selling their patents to one owner (who can then charge a monopoly price) the patentees can reap and divide the joint monopoly profits as a cartel would.

suspect under the antitrust laws.[6] Scotchmer (1991, pp. 33-34) and Green and Scotchmer (1990) suggest, however, that if a second innovation builds and improves upon the first, and the innovating firms cannot agree ex ante (before investment in the second innovation) to conduct joint R&D, then the courts should allow competing innovators to use collusive licensing agreements ex post to avoid competition that would dissipate their profits.[7]

Green and Scotchmer focus, however, on the possibility of agreements between firms ex ante and on what such research joint ventures imply for the optimal patent-antitrust policy. They do not do not explore the optimal tradeoff between the two innovators' incentives when such ex ante agreements are not possible. This paper will analyze the features of the optimal patent-antitrust policy under such circumstances, using a formal model that features some key elements excluded from their model.

For example, Green and Scotchmer do not allow the antitrust policy to be contingent on whether the improved invention competes with the first invention without infringing on the original patent or instead infringes and complements the original patent. Thus, their model sheds no light on the suggestion that complementary patents (which are complements in demand) should receive more permissive antitrust treatment than competing patents (which are substitutes in demand).[8] The model in this paper will

---

[6]In particular, Kaplow presents a model of process innovations that are perfect substitutes, notes that competition ex post would erode the incentive ex ante to invest in the first patent, and compares a competitive regime with one that allows combinations. Assuming free entry, Kaplow finds that a regime with combinations would entail more wasteful expenditures by imitators to invent around the initial patent and more deadweight loss from monopoly pricing per unit of profit for the initial patentee.

[7]Under such an antitrust regime, Green and Scotchmer find that patent policy can offer the second innovator correct incentives using a cutoff value for the improvement. If the value of the improvement is greater than the cutoff, the improved product does not infringe; otherwise, it does infringe. The cutoff value is an increasing function of the value of the original product. This policy, however, still offers insufficient incentives to the first innovator.

[8]Priest (pp. 357-58) distinguishes between complementary patents and competing patents. He asserts that price restrictions are essential when firms cross-license complementary patents, because the firms will otherwise compete away the patent rents. Kaplow (pp. 1860-62) disagrees, on the ground that per-unit royalties suffice to preserve

assume a richer set of policy instruments: courts may decide whether to allow collusive licensing agreements, both in cases of infringement and in cases of noninfringement. This model will also consider compulsory licensing for patents as a possible instrument.

Furthermore, in Green and Scotchmer's model, the expected value of the improvement (as well as the R&D cost for the first innovation) does not vary. Thus, their model does not address the considerations that arise when the authorities must commit ex ante to a policy that will apply to a large set of potential R&D projects of varying expected value and cost.[9] This paper models this tradeoff when the authorities know only the joint probability distribution over the possible values and expected costs of these innovations.

As this paper will show, both of these considerations have important implications for the optimal patent-antitrust policy. I shall evaluate (i) whether the court's decision on infringement should turn on the values of the two innovations, and if so, (ii) precisely how the optimal policy will depend on these values. I shall also determine when licensing agreements ex post between competing patentees are desirable, as Green and Scotchmer suggest, and when such collusive agreements are inefficient, as Kaplow argues.

B. Patent Law

Before describing the formal model, it will prove useful to review the law on patent scope which raises the issues that we will address. A patent application consists of a specification of the invention and a set of claims as to the subject matter that the applicant

---

adequate rewards for each patentee.

[9]For example, Green and Scotchmer's cutoff function, described above in footnote 7, optimizes incentives to the second innovator simply by randomizing between infringement (and partial appropriation by the first innovator) and noninfringement (which allows the second innovator to expropriate some value created by the first innovation through a collusive agreement with the first innovator). Any randomizing device, including a "lottery" based on a variable uncorrelated with the value of the improvement, could achieve this result. If a cutoff function dominates other devices, presumably, it is because when one must reduce incentives to the first innovator in order to offer greater incentives to the second innovator, it is best to sacrifice the first innovator's profits in those cases in which the expected value of the improvement is largest relative to the value of the original innovation.

regards as protected by the patent. See Patent Act, 35 U.S.C. § 112 (1988). A court may hold that a product infringes either because it falls within the patentee's claims, or if there is no such "literal infringement," because it is essentially the same as the patented invention. A court will hold that a device that falls outside a patent's claims nevertheless infringes the patent if the device has avoided those claims only through minor variations in the invention. This "doctrine of equivalents" applies if the inventions '"do the same work in substantially the same way, and accomplish substantially the same result, ... even though they differ in name, form, or shape."' Graver Tank & Manufacturing Co. v. Linde Air Products Co., 339 U.S. 605, 608 (1950) (quoting Machine Co. v. Murphy, 97 U.S. 120, 125 (1877)).

Patent authorities generally provide broad patent protection for particularly significant inventions. Patents for "pioneer" inventions, which cover "a function never before performed, a wholly novel device, or one of such novelty and importance as to mark a distinct step in the progress of the art" as defined in Westinghouse v. Boyden Power Brake Co., 170 U.S. 537, 561-62 (1898), are entitled to a broad range of equivalents. Inventions that represent less marked advances over prior technology are entitled to a correspondingly more limited range of equivalents. See Chisum (1990, Vol. 4, § 18.04[2]) and Lipscomb (1987, §§ 22:41-:45). Furthermore, the Patent Office often allow broad claims for pioneer inventions, leaving it to the courts to narrow the scope of the patent in subsequent infringement suits. See Merges and Nelson (1990, pp. 848-49).

A court may refuse to find an infringement under the doctrine of equivalents if the allegedly infringing device features major improvements rather than unimportant or insubstantial changes. See Chisum (Vol. 4, § 18.04[5]). A sample of the cases in this area provide examples of the phenomena that are the subject matter of this paper. In General Dynamics Corp. v. Whitcomb, 443 F.2d 630 (4th Cir. 1971), cert. denied, 404 U.S. 1016 (1972), the court found that an "antishock body" that General Dynamics used to improve the aerodynamic properties of its airplane wing did not infringe, because the device was attached to the wing in a different location than that specified in Whitcomb's patent. The court reasoned that because this "alteration does add to the effectiveness of the device, it cannot be considered a mere evasion of the patent." Id. at 633. Similarly, in Moleculon

Research Corp. v. CBS, Inc., 872 F.2d 407, 409 (Fed. Cir. 1989), for example, the court held that the Rubik's Cube and Rubik's Revenge puzzles did not infringe a patent owned by Moleculon because its design included changes that were not shown to be insubstantial in light of "its effects on the play value of the puzzle." Instead, the design was '"far superior"' and '"contributed significantly to the appeal and commercial success of those puzzles."' Id. (quoting 666 F. Supp. 661, 664 & n.3 (D. Del. 1987)). An improved invention can avoid infringement of even pioneer patents. In Texas Instruments, Inc. v. United States International Trade Commission, 805 F.2d 1558 (Fed. Cir. 1986), the court held that Texas Instuments' pioneering patent on the hand-held calculator did not encompass a competing version of the product subsequently developed with a variety of technological improvements.

On rare occasions, a court may hold that a patent does not include subsequent improved versions of the invention even in cases of literal infringement. See Chisum (Vol. 4, § 18.04[4]). The Supreme Court applied this "reverse" doctrine of equivalents in Westinghouse v. Boyden Power Brake Co., 170 U.S. 537, 568-73 (1898). In that case, George Westinghouse had invented a train brake that used reservoirs of compressed air for stopping power, and George Boyden had invented an ingenious improvement that provided added stopping power. Noting that the Westinghouse patent "did not prove to be a success until certain additional members had been incorporated into it," id. at 572, the Court refused to find that Boyden infringed, because his new design featured substantial improvements.

The doctrine of equivalents is somewhat at odds with some of the general principles of patent law. "The general design of the patent system avoids official decisions, by the Patent and Trademark Office or the courts, on the relative technological, social, or economic value of inventions. Such is left to the marketplace." Chisum (Vol. 4, pp. 18-101 to -102) (footnote omitted). Usually, improvements of (or additions to) a patented invention will not avoid infringement. See Lipscomb (§§ 22:29-:30).

One who invents an improvement on the patented product of another can obtain a patent on the improved feature alone. See Patent Act, 35 U.S.C. § 101 (1988). In this case the original patent is called the "dominant" patent, and the improvement patent is called the "subservient" patent. The two inventors then hold complementary "blocking patents": the original inventor cannot practice that particular improvement without a license from the second inventor, but the second inventor cannot practice the improvement either without a license from the original inventor. See Chisum (Vol. 3, § 9.03[2][ii]; Vol. 4, § 16.02[1][a]). For this reason, the second inventor would prefer a patent free of anyone else's claims and will not often voluntarily characterize its invention as subservient. The original inventor, on the other hand, would object to the second inventor's claim to a competing patent. In the course of litigation, a court may find that the second inventor's patent infringes the prior patent, even while upholding the infringer's patent as valid with respect to an improved feature. Merges and Nelson (pp. 861, 864-65) note that this practice is fairly common, and that such a holding in effect creates a complementary subservient patent.

As discussed above, however, a dominant patent poses a holdup problem for the holder of the subservient patent. For this reason, Merges and Nelson (pp. 857-68, 909-11) cite Texas Instuments approvingly as model for the application of the doctrine of equivalents and they defend the reverse doctrine of equivalents as applied in Westinghouse. These cases indicate that courts have broad discretion under these doctrines to decide the scope of patent protection, and Merges and Nelson propose that courts use these doctrines on a case-by-case basis to reduce the holdup problem posed by patents in the context of cumulative innovation.

Merges and Nelson (pp. 866 n.118) also suggest that a system of compulsory licensing, under which the patentee would be obliged to license its invention in exchange for a "reasonable royalty," would probably be the most efficient solution to the holdup problem. Although commentators, e.g., Adelman (1977), frequently have recommended greater use of compulsory licensing, political opposition has blocked efforts to enact reforms in this direction. See Scherer (1980, pp. 456-57). Courts in the United States, however, have sometimes ordered compulsory licensing as a remedy when they have found that the patentees have violated the antitrust laws. See Areeda and Kaplow (1988, ¶¶ 190, 284(c)-(d)). Although compulsory licensing is well established as an antitrust remedy, in general it remains disfavored in patent law. Merges and Nelson (p. 911) conclude that as long as compulsory licensing remains anathema, courts should rely on the doctrine of equivalents to solve the holdup problem.

Merges and Nelson (pp. 903-04, 914-15) also suggest applying the reverse doctrine of equivalents to another holdup problem raised by patents on natural products. These patents typically claim purified versions of substances that occur in nature. See Chisum (Vol. 1, § 1.02[9]). Although the patentees in these cases have in reality invented processes for purification of these substances, they seek and often obtain product patents. These patents pose a holdup problem even for those who invent entirely different (and better) processes for producing the same purified substance. Such R&D will be inhibited because the new processes will be held to infringe the product patent. For example, in Scripps Clinic and Research Foundation v. Genentech, Inc., 666 F. Supp. 1379 (N.D. Cal. 1987), Genentech had invented a recombinant DNA method for producing a human blood-clotting protein with major advantages over an earlier technique of purifying the protein from natural blood. The court held that Genentech had infringed the earlier product patent. The appropriate policy for such innovations, which are independent of the preceding substitute innovations, would appear to be the awarding process patents rather than product patents. Nevertheless, application of the reverse doctrine of equivalents after the fact would achieve the same result.

C. Antitrust Law

Antitrust law, like patent law, leaves courts with considerable discretion to make policy. This paper presents an economic model designed to test policy prescriptions proposed by various commentators to guide courts in the exercise of their discretion. Courts have recognized the tension between the patent system and general principle of antitrust law, and have modified antitrust doctrines accordingly. For example, although price-fixing and resale price maintenance are illegal per se under the antitrust laws, price restrictions in patent licenses are not illegal per se. In United States v. General Electric Co., 272 U.S. 476 (1926), the Supreme Court held that a patentee may set the price at which its licensee could sell the patented product.

Although the Court has never overruled General Electric, that doctrine "has not fared well," Sullivan (1977, p. 541), and many believe that it is "moribund," Areeda and Kaplow (1988, p. 447).[10] Areeda and Kaplow (p. 449) and Sullivan (pp. 545-46, 551-54) survey the case law and conclude that licenses may not fix prices if they cartelize an industry or cover substantial parts of the market. As Sullivan (p. 554) states, such cartel profits are not "'reasonably within the reward' to which the patentee is entitled."

Similarly, the acquisition of many patents is not illegal per se, but a combination of patents may violate the antitrust laws if it creates excessive market power. In Automatic Radio Manufacturing Co. v. Hazeltine Research, Inc., 339 U.S. 827, 834 (1950), the Supreme Court stated that the "mere accumulation of patents, no matter how many, is not in itself illegal." Courts have, however, looked upon acquisition of patents from others with greater suspicion than patents acquired by internal research. See, e.g., United States v. United Shoe Machinery Corp., 110 F. Supp. 295, 333 (D. Mass. 1953), aff'd per curiam, 347 U.S. 521 (1954). Sullivan (pp. 515-20) notes that the pooling of patents or the assignment of competing patents is analogous to a merger and argues that acquisitions of patents should be scrutinized under both Section 7 of the Clayton Act and Sections 1 and 2 of the Sherman Act for anticompetitive effects. See also Areeda and Turner (1978, Vol. 3, ¶¶ 704b, 705b, 819c) and Posner (1976, pp. 91-92).[11]

Courts will consider efficiency justifications for an otherwise suspect arrangement. If a patent pool (which share royalties among the participants) results from the settlement of patent disputes, for example, that may be a factor that militates in favor of the legality of the arrangement. In Standard Oil Co. (Indiana) v. United States, 283 U.S. 163 (1931),

---

[10]The Court has subsequently held such restrictions illegal in cross-licensing agreements, even if the patentees exchanged complementary blocking patents, United States v. Line Material Co., 333 U.S. 287 (1948), in the licenses of a patent pool including all the manufacturers in the industry, United States v. New Wrinkle, Inc., 342 U.S. 371 (1952), and in licenses granted to all competitors in an industry, United States v. United States Gypsum Co., 333 U.S. 364 (1948).

[11]In United States v. Hartford-Empire Co., 46 F. Supp. 541 (N.D. Ohio 1942), modified, 323 U.S. 386 (1945), for example, a group of firms purchased, accumulated, pooled, and cross-licensed patents so as to cartelize an entire industry. The court found that these arrangements violated the Sherman Act and ordered compulsory licensing as a remedy.

the Supreme Court upheld such a patent pool, citing the general policy in favor of settlements, which save the parties and society the costs of litigation. The Court also indicated that such arrangements are not anticompetitive if the defendants lack monopoly power over the relevant market. As Sullivan (pp. 567-70) notes, because a patent pool suppresses competition in royalty rates, courts should identify the market for the type of technology in question as the relevant market.

II. The Model

Consider a simple two-firm two-period model with no discounting. At the start of period 1, an idea for a product occurs exogenously to firm 1. For simplicity, I assume all consumers are identical and each demands one unit of the product per period. If firm 1 invests an amount c_1 in R&D, then firm 1 develops product 1, and the consumers are willing to pay an amount v_1 per period for the product. If and only if firm 1 develops product 1, then at the start of period 2 an idea for product 2 -- an improved version of product 1 -- occurs exogenously to a firm 2. If firm 2 invests an amount c_2 in R&D, then firm 2 develops product 2, which exceeds product 1 in value by an amount v_2 per period. There exist no other substitutes for these products. Once either product is developed, it can be produced at zero marginal cost.

At the time each of the innovators makes its investment decision, each knows the value and the R&D cost of its own innovation with certainty. Firm 1, however, faces uncertainty in period 1 over both the value and the cost of the second innovation in period 2. For simplicity, I assume that firm 1 knows only that v_2 is distributed continuously in the interval $[0, V_2]$ with positive density $f(v_2)$, where $F(v_2)$ is the corresponding cumulative distribution function and $V_2>0$, that c_2 is uniformly distributed in the interval $[0, C_2]$, where $C_2>0$, and that v_2 and c_2 are independently distributed.

Each firm is risk-neutral and chooses to develop its product if and only if its expected revenue exceeds its R&D cost. The expected revenue for each firm, however, depends upon public policy, and I shall use this model to explore the effects that alternative policies have upon the incentives to invent. For example, suppose that in the absence of any property rights for firm 1 in its invention, other firms would enter the market and use firm 1's technology for free. Suppose also that Bertrand competition then drives the price of product 1 to 0. Anticipating this outcome, firm 1 would not invest in product 1, and firm 2 would never have the opportunity to develop product 2.

Suppose instead that each firm receives an exclusive right to market its invention, i.e., a patent. If firm 1 sells its product under a patent monopoly, with unregulated prices, then firm 1 could appropriate all social surplus flowing from its sales, because I have assumed identical unit demands. Specifically, if firm 1 sells product 1 free from competition from product 2, it will choose the monopoly price v_1 and leave no consumer surplus. If firm 2 sells product 2 in competition with product 1, however, I assume that Bertrand competition drives the price of product 1 to 0 and that of product 2 to v_2 (minus some infinitesimal amount). Therefore, only product 2 is sold in equilibrium, but its price is constrained by the potential competition from product 1. Due to this competition, those who buy the product would enjoy v_1 in consumer surplus.

Again public policy ex post affects firm 1's incentives to invent, in this case not only through the likelihood that firm 2 decides to develop product 2, but also through the consequences of such a decision for firm 1's expected revenues. If firm 2 develops product 2, for example, suppose that firm 1 brings a lawsuit seeking to enjoin the marketing of product 2, claiming that product 2 infringes on the patent for product 1. If the court decides the second product infringes the patent granted to firm 1, then that patent includes the right to block entry by firm 2. Although this holdup right increases firm 1's expected revenues and encourages innovation by firm 1, it also reduces firm 2's revenues and discourages innovation by firm 2.

If firms can reach agreement ex ante (before investment in the second innovation) to conduct joint R&D in product 2 or to transfer the patent for product 1 to firm 2, then they can avoid this holdup problem and also avoid competing with one another in the product market in period 2. I assume that such transactions are either illegal or too difficult to arrange. There may be too many innovators with ideas for possible improvements for ex ante contracts to be a complete solution. See Scotchmer (1991, p. 37). As Merges and Nelson (1990, p. 877 n.160) note, in light of "problems of transaction costs ... it seems whimsical to assume that all improvers and potential improvers will be able to bargain with the holders of pioneering patents." Firm 2 may find, for example, that it cannot induce firm 1 to agree to an R&D joint venture without disclosing its idea. Such disclosure, however, would undermine the bargaining power of firm 2. Similarly, such asymmetric information may lead to strategic behavior that would inhibit the ex ante licensing or sale of firm 1's patent to firm 2, as the two firms bargain over the division of surplus.[12]

Now suppose that the courts act as a principal with limited information that seeks to design a general innovation policy that will induce desirable behavior on the part of its agents (the innovators). I assume that the courts can observe only the different flow values of the innovations ex post, v_1 and v_2 (not their costs, c_1 and c_2). The courts also know the information available to firm 1 in period 1, including the distributions of c_2 and v_2. Finally, the courts know that the two inventions are perfect substitutes aside from their vertical differentiation, that c_1 is uniformly distributed in the interval $[0, C_1]$, where $C_1>0$, and that v_2, c_1, and c_2 are all distributed independently of v_1 and of each other.

The courts establish a reputation for adherence to a particular policy (by following precedent), so as to give their policy the desired effect. This policy is a function mapping each point in (v_1, v_2) space to a pair (π_1, π_2), which represents the second-period revenues received by each of the two firms. More precisely, policy will allocate each pair of relative values, v_1/V_2 and v_2/V_2,[13] to a legal regime that determines both (1) the prices paid by consumers, and (2) how the resulting revenue is allocated between the two firms.

In Part III.A, I shall assume that the courts can pursue a hypothetical policy through which they can set the absolute value of these variables precisely for each (v_1, v_2) pair. To implement such a policy in practice, the courts would need to observe the absolute values of v_1 and v_2. In this model, this hypothetical policy is equivalent to a government

---

[12]See Caves, Crookell, and Killing (1983) for a discussion of market failures in the bargaining over technology licenses.

[13]Without loss of generality, we can normalize so that all variables and parameters express values relative to V_2. That is, the features of the optimal policy will depend only on values relative to V_2. Thus, the model is general enough to include cases with varying absolute values for V_2; one need only interpret the other variables accordingly.

agency that purchases each innovation from the corresponding patentee and sells the improved product to the public, but is required to balance its budget (i.e., set revenues equal to the total amount paid to the two innovators). This policy is also equivalent to the contract to which the parties would agree ex ante (before any opportunities for R&D occur), if the absolute values (but not the costs) of inventions are verifiable ex post in court. I refer to the optimal hypothetical policy as the "second best" insofar as it is the best that the court could implement given that it cannot observe R&D costs and given the "balanced budget" constraint (that is, the requirement that inventors be rewarded only with revenues from product sales).[14]

In Part III.B, I shall assume instead that the courts have only a cruder set of instruments: this "third best" policy will consist of a function mapping each point in (v_1,

---

[14]In reality, courts can observe the economic value of an invention only imperfectly. If the regulatory authorities could observe the absolute value of each invention perfectly, then one might think that an unconstrained system of rewards ex post for successful innovators (the "award" or "prize" system) would perform better than budget-constrained policies, including the patent system. Indeed a major rationale for the patent system is the fact that the regulatory authorities lack information that innovators possess, such as knowledge regarding the demand for inventions. For this reason, it may be better to rely on a patent system in which the monopoly profit extracted from the market is correlated (albeit imperfectly) with the social surplus created by the invention.

There may be other reasons, however, that one may prefer not to rely on an unconstrained government agency that transfers funds to innovators based on its appraisal of the values of their inventions. Laffont and Tirole (1990) model another problem that may arise if regulators are authorized to make transfers to firms: the agency may be "captured" by the industry. The award system also suffers from the problem that the authorities are tempted to "hold up" the inventor through conservative estimates of the invention's value.

In any event, the justification of the patent system itself is beyond the scope of this paper. Like much of the literature on the optimal patent policy, e.g., Green and Scotchmer (1990) and Klemperer (1990), this paper will instead take the constraints on policy as given and assume for the sake of argument that the regulatory authorities have full information regarding the demand for the innovations in question. The model makes this assumption simply to address the question: to the extent that the courts can acquire information on the value of inventions, how (if at all) should that affect their decisions? Even without any direct measure of value, courts can observe the degree of differentiation between the products, for example -- a characteristic that may be correlated with the economic value of the improvement.

v_2) space to one of four possible patent-antitrust regimes. To implement this policy, however, the courts need not know the absolute values of either v_1 or v_2; it will suffice if the courts observe the relative values v_1/V_2 and v_2/V_2. For each (v_1, v_2) pair, the court would decide both (1) whether the second product infringes the first patent, and (2) whether the two firms should be allowed to collude through a licensing agreement (or a transfer of one of the patents). These two binary decisions create a matrix of four possible regimes, each of which implies not only particular second-period revenues (π_1, π_2) for firm 1 and firm 2, respectively, but also a particular level of consumer surplus.

The second-best policy would dominate this set of four possible outcomes, because the hypothetical policy in Part III.A could replicate each of the four outcomes as a special case. Indeed, under that hypothetical policy, courts could divide the social surplus among the consumers and the two innovators in any way desired. As shown in Part III.A, courts in theory could implement the second-best policy through a compulsory licensing policy. Current law, however, restricts courts to the cruder third-best policy studied in Part III.B, perhaps because a general compulsory licensing policy would require courts to gather and process too much information.[15] In any event, analysis of the second-best policy will prove useful in understanding the features of the third-best policy.

Using either policy, the court can affect social welfare through the payoffs to each of the two firms. Let $\pi_2(v_1, v_2)$ denote the revenue (from the sales of product 2) received by firm 2 if it develops the product. This $\pi_2(v_1, v_2)$ will be a function of the courts' innovation policy. Thus, firm 2 would develop product 2 if and only if its expected net payoff is positive, that is, if and only if $\pi_2(v_1, v_2)$ exceeds c_2. Similarly, firm 1's R&D decision would depend on the innovation policy as specified in Part III.

Before proceeding to that discussion, it will prove useful to introduce some more notation. The social value flowing from sales of product 2 is v_2. Let $S(v_1)$ denote the expectation in period 1 of the net social value (conditional on firm 1 having developed

[15]See Areeda and Kaplow (1988, ¶ 284(d), pp. 444-45) (arguing that the difficulties of judicial administration and "the problems of calculation" of a "reasonable royalty" would be "formidable").

product 1 with value v_1) created by the possibility that firm 2 will develop product 2 in period 2. Thus, $S(v_1)$ equals the expected value of v_2-c_2, conditional on the development and sale of product 2, multiplied by the probability of the development and sale of product 2, $\Pr[\pi_2(v_1, v_2) > c_2]$. For any given v_1, I assume:

$$C_2 \geq \pi_2(v_1, v_2) \tag{1}$$

for all possible v_2 and for all possible innovation policies in period 2.[16] Then we may express $S(v_1)$ as follows:

$$S(v_1) = \int_0^{V_2} \int_0^{\pi_2(v_1,v_2)} \frac{v_2 - c_2}{C_2} dc_2 \, dF(v_2), \tag{2}$$

where we integrate with respect to v_2 to include all possible ideas for improvements and with respect to c_2 to include all possible R&D costs for any given v_2. For any given v_1, S will depend on how the innovation policy would treat a second invention (i.e., what π_2 each possible value of v_2 would imply).

Let $\Delta\pi_1(v_1, v_2)$ represent the change in firm 1's second-period revenues induced by the development of product 2. This $\Delta\pi_1(v_1, v_2)$ will be a function of the courts' innovation policy. Let $P(v_1)$ denote the expected net <u>private</u> benefit for firm 1 (conditional on having developed product 1) created by the possibility that firm 2 will develop product 2. Thus, $P(v_1)$ equals the expected value of $\Delta\pi_1(v_1, v_2)$, conditional on the development and sale of product 2, multiplied by the probability of the development and sale of product 2, $\Pr[\pi_2(v_1, v_2) > c_2]$. That is:

---

[16]This assumption ensures that greater rewards for firm 2 will always induce greater expected R&D expenditures. In this model, any policy that increases the expected payoff to firm 2 would encourage more expected innovation in period 2 only by inducing firm 2 to undertake projects that were otherwise too costly to be worthwhile. To create this effect, there must be some probability that the cost of developing product 2 exceeds the revenue to firm 2. Assumption (1) is necessary because c_2 can realize values only as high as C_2. In terms of equation (2), this simplifying assumption allows us to use $\pi_2(v_1, v_2)$ as an upper limit of integration rather than $\min[C_2, \pi_2(v_1, v_2)]$.

$$P(v_1) = \int_0^{v_2} \int_0^{\pi_2(v_1,v_2)} \frac{\Delta\pi_1(v_1,v_2)}{C_2} dc_2 \; dF(v_2), \tag{3}$$

where we integrate with respect to v_2 to include all possible ideas for improvements and with respect to c_2 to include all possible R&D costs for any given v_2. For any given v_1, P will depend on the $\Delta\pi_1$ and π_2 implied by each v_2 under the existing innovation policy. We can now analyze the optimal innovation policy.

III. The Optimal Innovation Policy

Suppose that the courts seek to maximize social welfare through their innovation policy. Given unit demands, as long as the price of a product does not exceed its value, that price merely determines how the social surplus from the product is divided between producer and consumer; there is no deadweight loss from prices above marginal cost in this model. To the extent that policies affect a product price within this range, then, they would affect social welfare only through the incentives for producers to invent and market products. The "first best" policy from this perspective would align the private incentives for producers with the social value derived from these products.

For example, to optimize the incentives to firm 1 in the absence of firm 2, courts would permit firm 1 to charge the monopoly price for product 1, v_1, which would allow firm 1 to appropriate as much social surplus as possible. To price any lower would lead firm 1 not to invest in R&D in cases in which the social benefit of the invention would exceed its cost. The lower the price below v_1, the greater the expected loss in social surplus, both because these cases become more numerous and because the average value (net of R&D costs) of the invention foregone increases. To price any higher than v_1, however, would eliminate any sales (and thus any benefits, private or social) of the product. Similarly, given the existence of product 1, to optimize incentives for firm 2 to invent product 2, a court would give firm 2 the full social value of product 2, v_2. That is, $S(v_1)$ is maximized by $\pi_2(v_1, v_2) = v_2$.

To maximize welfare, however, the courts must also consider the incentives for the first inventor, given the possibility of the second invention. Consider the optimal innovation policy in the relatively simple case in which the technology for the second invention is independent of the first invention, that is, in which the probability distribution for the cost and value of the second innovation is the same whether or not firm 1 develops product 1. This description would apply, for example, to the innovations in the Genentech infringement case. In such a case, a finding of noninfringement and a prohibition on collusion between firm 1 and firm 2 would provide optimal incentives to both firms. In competition with product 1, firm 2 would receive the social benefit of its invention, which would be the amount by which product 2 exceeds product 1 in value, v_2. Moreover, given the possibility of the invention of product 2, firm 1 receives the social benefit of its invention, which is v_1 for the first period, plus v_1 for the second period if firm 2 does not invent product 2, and 0 if firm 2 does invent product 2. This reasoning lends support both for Merges and Nelson's conclusion that the court should find no infringement in a case like Genentech and for the general policy against collusion among holders of competing patents.

In this model, however, as is often the case in reality, firm 2 would not have invented product 2 unless firm 1 had first invented product 1. The social benefit from the invention of product 1, then, would include the spillover to firm 2: product 1 provides firm 2 with the inspiration for product 2. To provide socially optimal incentives to firm 1, then, the court would increase firm 1's payoffs to include the expected value (in period 1) of this spillover (in period 2): if firm 2 receives socially optimal incentives, then this spillover is the expected value of $\max(0, v_2\text{-}c_2)$. The social value of firm 1's innovation, then, would be:

$$2v_1 + E[\max(0, v_2\text{-}c_2)]. \tag{4}$$

In the second period, however, a patent system cannot offer both producers together more than the v_1+v_2 that consumers are willing to pay for the improved product. The problem for the courts, then, is to allocate the social surplus so as to create the optimal trade-off between the incentives to the two firms.

In this model, for any possible patent policy (optimal or otherwise) that the court could implement, the following lemma will hold:

<u>Lemma 1</u>: Any patent policy will offer firm 1 less than the expected social value of its innovation.

<u>Proof</u>: The expected social value of firm 1's innovation equals the sum of firm 1's expected revenues, firm 2's expected surplus (its expected revenues π_2 minus its expected R&D costs c_2), and expected consumer surplus. A patent system can offer firm 1 and firm 2 no more revenue than what consumers are willing to pay for product 1, v_1, plus (if firm 2 invents product 2) what they are willing to pay for the improvement embodied in product 2, v_2, because the patent system has no other source of revenues. Furthermore, if firm 2 is to invent product 2, then a patent system must offer firm 2 at least its R&D costs, c_2. Indeed, because c_2 may vary, and the patent authorities cannot observe its realized value, any given revenue π_2 offered to firm 2 would leave firm 2 positive expected surplus, because the expected value of c_2 (conditional on firm 2's invention of product 2) must be less than π_2. Thus, a patent system must offer consumers a nonnegative share of the social surplus, and firm 2 a positive share. Therefore, firm 1 cannot appropriate all the social surplus flowing from its innovation. ■

Lemma 1, in turn, implies the following lemma:

<u>Lemma 2</u>: The welfare-maximizing court would allow firm 1 to charge the monopoly price v_1 in the first period, and if firm 2 does not invent product 2, in the second period as well.

<u>Proof</u>: Any lower price would transfer value to consumers without increasing social surplus ex post, but would reduce firm 1's incentives to invent ex ante. Reducing firm 1's incentives must reduce social welfare, because in this model, as Lemma 1 states, the patent system already offers firm 1 less than the social value of its innovation. ■

Given the price v_1 in Lemma 2, we can state:

$$\pi_1(v_1, v_2) = v_1 + \Delta\pi_1(v_1, v_2). \tag{5}$$

The expected payoff to firm 1 for inventing product 1 equals $2v_1$, which is its revenues in the absence of product 2, plus $P(v_1)$, which is the net change in its expected payoff due to the possibility of the second innovation. For any given v_1, I assume:

$$C_1 \geq 2v_1 + P(v_1) \tag{6}$$

for all the possible innovation policies in period 2.[17] Then we may express the social welfare function (conditional on a particular v_1), denoted by $W(v_1)$, as follows:

$$W(v_1) = \int_0^{2v_1+P(v_1)} \frac{2v_1+S(v_1)-c_1}{C_1} dc_1$$
$$= \{\frac{1}{C_1}[2v_1+P(v_1)]\}\{[2v_1+S(v_1)] - ½[2v_1+P(v_1)]\} \quad (7)$$

where we integrate with respect to c_1 to include all possible R&D costs for any given v_1. The second expression for $W(v_1)$ is the product of two terms. The first term is the probability that firm 1 will invent product 1. The second term is the net social benefit from that invention -- $2v_1$ plus $S(v_1)$ minus the expected R&D costs for the first invention.

Expected social welfare (unconditional) would be the expectation of $W(v_1)$ taken with respect to v_1. To maximize this expected welfare, however, the court need only maximize $W(v_1)$ for each v_1, because the court can observe v_1 and tailor its innovation policy accordingly. Furthermore, policy can also depend on v_2, and the optimal policy must be such that no small change in policy for any (v_1, v_2) pair can cause a welfare improvement.

To study the features of a welfare-maximizing policy, then, we shall consider the effect of changes in $\Delta\pi_1$ and π_2 for one (v_1, v_2) pair. In particular, consider the effect of such changes in innovation policy on welfare conditional on a given v_1. As we can see from (7), policy affects $W(v_1)$ only through $S(v_1)$ and $P(v_1)$. For small policy changes, such as changes pertaining only to one (v_1, v_2) pair, we can totally differentiate (7) to obtain:

$$dW(v_1) = \{[2v_1+P(v_1)]dS(v_1) + [S(v_1)-P(v_1)]dP(v_1)\}/C_1. \quad (8)$$

[17]Assumption (6), by reasoning similar to that in footnote 16, ensures that greater rewards for firm 1 would always induce greater expected R&D expenditures. In terms of equation (7), this simplifying assumption allows us to use $2v_1+P(v_1)$ as the upper limit of integration rather than $\min[C_1, 2v_1+P(v_1)]$.

Note that both the coefficient on $dS(v_1)$ and that on $dP(v_1)$ must be nonnegative.[18]

These two coefficients represent respectively the relative importance of the two objectives of the optimal innovation policy. Ceteris paribus, the courts should optimize the incentives to firm 2 so as to increase $S(v_1)$, the expected social surplus from the invention of product 2 (conditional on the invention of product 1 with value v_1). The importance of this objective is proportional to the probability of the first invention, $[2v_1+P(v_1)]/C_1$. Ceteris paribus, the courts should also optimize the incentives to firm 1, which entails increasing $P(v_1)$ insofar as $P(v_1) < S(v_1)$. The importance of this objective is proportional to the degree to which firm 1 cannot appropriate the social benefits of product 2, $S(v_1)-P(v_1)$, because the marginal gain in welfare from reducing this difference is proportional to this difference.

In general, then, any small policy change that increases both $S(v_1)$ and $P(v_1)$ must be desirable. To be precise, we can state the following lemma:

Lemma 3: Welfare given any particular v_1, $W(v_1)$, is nondecreasing in $S(v_1)$ and strictly increasing in $P(v_1)$. Furthermore, if either $v_1>0$ or $P(v_1)>0$, then $W(v_1)$ must be strictly increasing in $S(v_1)$ also.

A. The Second-Best Policy

Suppose that the courts can pursue the second-best policy described above in Part II. For any (v_1, v_2) pair, the optimal price will be the monopoly price, v_1+v_2, because there is no deadweight loss from monopoly pricing in this model. Any lower price could be raised to increase the royalties for firm 1, for example, which by Lemma 1 must improve the incentives to invent product 1. The more complex issue is the allocation of these monopoly profits between firm 1 and firm 2.

---

[18]First, as long as $v_1>0$, we know $2v_1+P(v_1) > 0$, because even the policy least favorable to firm 1 can only reduce its second-period payoff to 0. Therefore, $P(v_1)$ can be no less than $-v_1$. (If $v_1=0$, then $2v_1+P(v_1) = 0$ is possible.) Second, as long as there is some positive probability that firm 2 will invent product 2, we know $S(v_1)>P(v_1)$, because Lemma 1 and Lemma 2 together imply that firm 1's private benefit from the invention of product 2 must fall short of the social benefits from that invention.

Consider the cases in which $v_2>0$.[19] Then we can always express the payoff to firm 2 as a multiple of the value of product 2. That is, let $\pi_2(v_1, v_2) = \alpha(v_1, v_2)v_2$. The payoff to firm 1, then, will be the remaining revenue: $\pi_1(v_1, v_2) = v_1 + [1-\alpha(v_1, v_2)]v_2$, which together with (5), implies that $\Delta\pi_1(v_1, v_2) = [1-\alpha(v_1, v_2)]v_2$. The court's optimal policy would consist of some $\alpha>0$ for each (v_1, v_2) pair.[20] That is, the court would choose some $\alpha>0$ for each v_2 to maximize $W(v_1)$.

We can substitute for π_2 in (2) and express $S(v_1)$ as:

$$\begin{aligned} S(v_1) &= \int_0^{V_2} \int_0^{\alpha(v_1,v_2)v_2} \frac{v_2-c_2}{C_2} dc_2 \; dF(v_2) \\ &= \frac{1}{C_2} \int_0^{V_2} [\alpha(v_1,v_2) - \tfrac{1}{2}\alpha^2(v_1,v_2)]v_2^2 \; dF(v_2), \end{aligned} \tag{9}$$

and substitute for $\Delta\pi_1$ and π_2 in (3) and express $P(v_1)$ as:

$$\begin{aligned} P(v_1) &= \int_0^{V_2} \int_0^{\alpha(v_1,v_2)v_2} \frac{[1-\alpha(v_1,v_2)]v_2}{C_2} dc_2 \; dF(v_2)dv_2 \\ &= \frac{1}{C_2} \int_0^{V_2} [\alpha(v_1,v_2) - \alpha^2(v_1,v_2)]v_2^2 \; dF(v_2). \end{aligned} \tag{10}$$

Taking the derivatives of (9) and (10) with respect to α, and using (8), we find that the first-order condition for this maximization problem implies that:

$$\{[2v_1+P(v_1)](1-\alpha) + [S(v_1)-P(v_1)](1-2\alpha)\}v_2^2 f(v_2)/C_1C_2 = 0. \tag{11}$$

The left-hand side of (11) represents the net marginal benefit from increasing α.

---

[19]The case of $v_2=0$ is trivial; then innovation by firm 2 is never desirable, and the optimal policy would always discourage such R&D by ensuring that firm 2 receives no revenues in such a case: $\pi_2(v_1, 0) = 0$.

[20]If α were nonpositive, then firm 2 would never develop product 2, even if the social value v_2 were to exceed the cost c_2. Therefore, increasing α slightly above 0 could impose only a welfare gain through the incentives for firm 1 and firm 2. One can see from (2) and (3) that any (v_1, v_2) pair with $\alpha=0$ (and thus $\pi_2=0$) would be making no contribution to either $S(v_1)$ or $P(v_1)$; with any α in the (0, 1) interval (and thus with $\pi_2>0$ and $\Delta\pi_1>0$), however, it would make a positive contribution to both $S(v_1)$ and $P(v_1)$. Given Lemma 3, these effects must improve welfare.

Lemma 4: For each (v_1, v_2), the optimal α is strictly less than 1 and must satisfy:

$$\alpha = [2v_1+S(v_1)]/[2v_1+2S(v_1)-P(v_1)]. \tag{12}$$

Proof: Note that the left-hand side of (11) must be negative if $\alpha \geq 1$ and $v_2>0$. Therefore, welfare must improve as α is reduced from such a value to below 1, and any $\alpha \geq 1$ cannot be optimal. Thus, α always lies strictly within the (0, 1) interval. One can see from (9) and (10), therefore, that once we exclude cases in which $v_2=0$, both $S(v_1)>0$ and $P(v_1)>0$.

Solving the first-order condition (11) for α yields (12). Note that the right-hand side of (12) must be strictly greater than ½ for any v_1, because $P(v_1)>0$, but also strictly less than 1, because $S(v_1)>P(v_1)$. Any such α would also meet the second-order condition. Taking the second derivative of $W(v_1)$ with respect to α, one can show that it must be negative for any α in the interval (½, 1). Thus, any α meeting the first-order condition is a maximum, and the optimal α must satisfy (12). ■

Note that the optimal α in (12) depends on v_1 but is independent of v_2. The marginal costs and benefits of increasing α for any given (v_1, v_2), as (11) indicates, are both simply proportional to v_2^2. Therefore, the optimal $\alpha(v_1, v_2)$ will be a function of v_1 only. We can now derive an explicit solution for the unique optimal policy $\alpha(v_1)$, which would allow firm 2 to capture most of the social value of its innovation, but no more than the total social value of the innovation:

Proposition 1: The second-best policy would give firm 2 a fraction $\alpha(v_1)$ of the social value of its invention, v_2, and give firm 1 the remaining share, where:

$$\alpha(v_1) = \frac{E(v_2^2) - 2v_1C_2 + \sqrt{[E(v_2^2)]^2 + 8E(v_2^2)v_1C_2 + 4v_1^2C_2^2}}{3E(v_2^2)} \tag{13}$$

for all (v_1, v_2). This $\alpha(v_1)$ must lie in the interval [⅔, 1) and is strictly increasing in v_1. If $v_1>0$, then $\alpha(v_1)$ is strictly greater than ⅔, and is strictly increasing in C_2, but strictly decreasing in $E(v_2^2)$.

Remark: Note that if $v_1=0$, then $\alpha(v_1)$ equals ⅔. The explanation for this result is as follows. If $v_1=0$, then the possibility of the second innovation is the only incentive for firm 1 to invent. Therefore, the probability of invention by firm 1 is proportional to $P(v_1)$. At $\alpha(v_1) = ⅔$, firm 1 and firm 2 evenly divide the expected social surplus (net of firm 2's expected R&D costs, which will consume half of firm 2's expected revenues) from the second innovation, so that $P(v_1) = ½S(v_1)$. As one can see from (8), this equality implies that increasing $P(v_1)$ and increasing $S(v_1)$ are equally important imperatives at this point. Furthermore, once $\alpha(v_1)$ rises to ⅔, $P(v_1)$ is declining at the same rate as $S(v_1)$ is rising. Thus, the marginal cost of further increases in α has risen to equal the (diminishing) marginal benefit.

If $v_1>0$, then as one can see from (8), increasing $S(v_1)$ is more important than increasing $P(v_1)$ at $\alpha(v_1) = ⅔$. The marginal benefit of further increasing α above ⅔ is now greater than the marginal cost, rather than just equal to marginal cost at ⅔ (as it would if $v_1=0$). Thus, as v_1 grows larger, the marginal benefit schedule rises, and so does the optimal α. Indeed, as v_1 rises toward infinity, $P(v_1)$ makes a proportionally less significant contribution to firm 1's incentives. Then increasing $S(v_1)$ becomes far more important than increasing $P(v_1)$. As the implications for the incentives for firm 2 become relatively more important, the optimal α grows arbitrarily close to 1.

Increasing C_2 or decreasing $E(v_2^2)$ similarly increases the relative importance of v_1 in firm 1's incentives. A larger C_2 makes the second innovation less likely; a smaller $E(v_2^2)$ makes the second innovation less important. Thus, if $v_1>0$, then as $C_2/E(v_2^2)$ ranges toward infinity, the optimal α again grows arbitrarily close to 1.

<u>Proof</u>: See Appendix A.

Finally, note that courts can implement this second-best policy through an idealized compulsory licensing policy. Suppose the court finds that the second product infringes the first patent, but allows the second innovator to retain a subservient (and therefore complementary) patent. The court can then require each innovator to license its innovation to anyone who desires to do so at what the court determines to be a "reasonable royalty." Thus, firm 2 (or its licensee) could license the basic technology from firm 1 and market product 2. Following Tandon (1982), suppose that the court has complete discretion in setting the per-unit royalty rate for use of each innovation and thereby controls the flow of revenue received by each firm. Therefore, the court can set firm 1's royalty at $v_1 + [1-\alpha(v_1)]v_2$ and firm 2's royalty at $\alpha(v_1)v_2$ minus some infinitesimal amount. If firm 2 invents product 2, then competing firms will seek to license the complementary innovations from firm 1 and firm 2, driving down the price paid by consumers until that price just covers the cost of the royalties. Consumers would choose to buy the improved product at a price equal to v_1+v_2 minus some infinitesimal amount, and given that firm 1 would make v_1 on each such sale, firm 1 would not have any incentive to undercut this price by selling product 1 for less than v_1.

B. <u>The Third-Best Policy: Antitrust Regulation and Patent Scope</u>

I now restrict the court's control over the innovators' payoffs. Under current law, courts in general cannot resort to compulsory licensing. Suppose the court can only implement the third-best policy described above in Part II; that is, the court can affect the firms' behavior only by it decisions on infringement and on collusive licensing. The court observes v_1 and v_2, then assigns that case to one of four patent-antitrust regimes. The firms may then bargain over a licensing agreement. Let the firms evenly divide the joint surplus from any licensing agreement, collusive or otherwise; that is, the parties reach the Nash bargaining solution. The court's decisions would set the threat point for such bargaining and thereby determine both $\Delta\pi_1$ and π_2 for that (v_1, v_2) pair. Thus, the patent-antitrust policy implies a pair of functions, $\Delta\pi_1(v_1, v_2)$ and $\pi_2(v_1, v_2)$, which in turn determine the welfare function, $W(v_1)$ in (7), through $S(v_1)$ in (2) and $P(v_1)$ in (3). I shall next describe each of these four regime and the payoffs that each would imply in the second period:

1. <u>Infringement</u>

If the court finds that product 2 does infringe (let I denote this finding), then it declares firm 2's patent to be subservient to firm 1's dominant patent: firm 2 cannot sell product 2 unless firm 1 gives its consent, that is, unless firm 1 licenses its technology to firm 2. In this case, the second innovator may compensate the first for its consent through the licensing agreement. Firm 2 may also license its innovation to firm 1, which cannot use the patented improvement without the consent of firm 2. The courts also may allow the two firms to avoid competition in the product market through collusive licensing or a transfer of one of the patents (let I∩C denote this regime), or they may prohibit such collusive agreements (let I∩NC denote this regime). Under either I regime, if the firms do not agree on a licensing agreement (collusive or otherwise), firm 1 would deny its consent to firm 2 (and obtain v_1) in the second period, rather than compete with firm 2 (and obtain 0). These payoffs constitute the threat point for the bargaining over any surplus created by any licensing agreeement.

a. Infringement with No Collusion: Suppose the courts prohibit both price restrictions in patent licensing agreements and transfers that enable one party to own both patents. Nevertheless, the two firms can enjoy the maximum possible joint producer surplus: v_1+v_2. For example, as Kaplow (pp. 1860-62) notes, owners of complementary patents can cross-license and charge per-unit royalties to one another that sum to the monopoly price, v_1+v_2. This arrangement leads to the monopoly price as the equilibrium price, because only at that price will each party have no incentive to change price and thereby shift sales between the two firms: each would receive the same revenue from each unit sold regardless of which party sells the unit. Furthermore, if the two firms do not cross-license, they can reap monopoly profits even if the courts prohibit per-unit royalties. In particular, firm 2 can license its technology to firm 1, and as long as firm 1 does not also license its technology to firm 2, firm 1 can sell the improved product as a monopolist. Firm 2 cannot compete with firm 1, even if firm 1 cannot tie sales of product 1 and of product 2, because firm 2's complementary patent is subservient. In any event, the firms would choose some licensing arrangement that produces monopoly profits, then divide the surplus between them. Specifically, they divide the surplus v_2 that licensing creates beyond the v_1 that firm 1 could obtain on its own. Thus, under the I∩NC regime, firm 1 receives $v_1+\frac{1}{2}v_2$, and firm 2 receives $\frac{1}{2}v_2$.

b. Infringement with Collusion: Under the I∩C regime, the firms can agree to sell the improved product only at the monopoly price, v_1+v_2. Again, firm 1 receives $v_1+\frac{1}{2}v_2$, and firm 2 receives $\frac{1}{2}v_2$. Thus, each of the two I regimes produce the same payoffs: the first innovator can appropriate half of the value created by the improvement. The court's antitrust policy does not matter in cases of infringement, so we can refer to the I regime without specifying an antitrust policy. Under this regime, neither firm would receive the full social value generated by its innovation: firm 1 receives less under any regime, as we know from Lemma 1, and firm 2 receives less than v_2. In terms of the policy described above in Part III.A, this regime is equivalent to choosing $\alpha = 1/2$. Compared with the second-best policy, then, firm 2 receives too small a share of v_2, and firm 1 receives one too large, and as Proposition 1 indicates, the magnitude of these deficiencies increase with v_1.

2. No Infringement

If the second product does not infringe (let NI denote this finding), it may be sold in (Bertrand) competition with the first. In this case, the court may either allow the holders of competing patents to collude through a licensing agreement or by a transfer of one of the patents (let $NI \cap C$ denote this regime) or it may prohibit such collusion (let $NI \cap NC$ denote this regime).

a. No Infringement with No Collusion: Under the $NI \cap NC$ regime without licensing, competition yields 0 for firm 1 and v_2 for firm 2. With competition ex post, producer surplus can be no more than v_2; therefore, there is no incentive for firm 2 to agree to a licensing agreement.[21] Under this regime, then, firm 2 appropriates the full social value of its innovation, v_2, while firm 1 receives nothing. For firm 2, this regime dominates the I regime. Firm 1, however, receives the lowest possible payoff under this regime; thus, for firm 1, all other regimes dominate this regime. Furthermore, firm 1 receives less profit, and firm 2 receives more, than it would under the second-best policy.

b. No Infringement with Collusion: Under the $NI \cap C$ regime, the firms would agree

---

[21]Even cross-licensing with per-unit royalties cannot by itself sustain prices any higher. To suppress competition, a more elaborate agreement would be necessary. For example, an agreement can maintain monopoly prices if the royalties are based on all of a firm's sales regardless of which patent the firm uses for the particular units sold. The royalties might then be pooled and the surplus divided between the firms. Alternatively, each firm could agree to sell only its rival's product, a per-unit royalty of v_1 for product 1 could maintain monopoly prices, and the royalty for product 2 could provide firm 2 with the desired share of the monopoly profits. I assume that courts would recognize such collusive arrangements and prohibit them whenever they would prohibit explicit price-fixing.

to sell only the improved product, fix its price at the monopoly level, and divide the surplus from such a collusive agreement. Therefore, firm 1 receives $\frac{1}{2}v_1$, and firm 2 receives $\frac{1}{2}v_1+v_2$. Thus, under the NI∩C regime, the second innovator can expropriate half of the value of the first product. Firm 1 fares better under this regime than under the NI∩NC regime, but not as well as it does under the I regime. For firm 2, this regime dominates all others: firm 2 receives the highest possible payoff. Indeed, firm 2 receives <u>more</u> than the full social value of its innovation; under this regime, firm 2 has an <u>excessive</u> incentive to innovate. In terms of the policy described above in Part III.A, this regime is equivalent to choosing $\alpha(v_1, v_2) = 1 + v_1/2v_2$. Compared with the second-best policy, then, firm 1 receives too little profit, and firm 2 receives too much profit. For any given v_1, however, the magnitude of these deficiencies decrease with v_2.

An optimal policy would assign all (v_1, v_2) pairs to one of these three regimes: I, NI∩C, or NI∩NC. Each pair would be assigned such that it could not raise $W(v_1)$ to switch it to another regime. The resulting configuration implies a particular $S(v_1)$ and a particular $P(v_1)$, with $P(v_1)$ strictly less than $S(v_1)$.

Consistent with assumption (1), let $C_2 \geq \frac{1}{2}v_1+v_2$ for all possible v_2 (i.e., let $C_2 \geq \frac{1}{2}v_1+V_2$). Then we may express $S(v_1)$ as follows:

$$\begin{aligned} S(v_1) &= \int_{I}\int_{0}^{\frac{1}{2}v_2} \frac{v_2-c_2}{C_2}dc_2dF(v_2) + \int_{NI\cap C}\int_{0}^{\frac{1}{2}v_1+v_2} \frac{v_2-c_2}{C_2}dc_2dF(v_2) + \int_{NI\cap NC}\int_{0}^{v_2} \frac{v_2-c_2}{C_2}dc_2dF(v_2) \\ &= \frac{1}{C_2}\left[\int_{I}(\tfrac{3}{8}v_2^2)dF(v_2) + \int_{NI\cap C}(\tfrac{1}{2}v_2^2-\tfrac{1}{8}v_1^2)dF(v_2) + \int_{NI\cap NC}(\tfrac{1}{2}v_2^2)dF(v_2)\right] \end{aligned} \tag{14}$$

where the probability of the second invention would depend on the regime to which it would be assigned. We must integrate with respect to c_2 over a different interval for each regime, and integrate with respect to v_2 within each regime separately. Note that the NI∩NC regime induces the optimal behavior for firm 2. Thus, the patent-antitrust policy that assigns all v_2 (from 0 to V_2) to that regime would maximize $S(v_1)$.

We may also express $P(v_1)$ as follows:

$$
\begin{aligned}
P(v_1) &= \int_I \int_0^{\frac{1}{2}v_2} \frac{\frac{1}{2}v_2}{C_2} dc_2 dF(v_2) + \int_{NI \cap C} \int_0^{\frac{1}{2}v_1 + v_2} \frac{-\frac{1}{2}v_1}{C_2} dc_2 dF(v_2) + \int_{NI \cap NC} \int_0^{v_2} \frac{-v_1}{C_2} dc_2 dF(v_2) \\
&= \frac{1}{C_2} \left[\int_I (\tfrac{1}{4}v_2^2) dF(v_2) + \int_{NI \cap C} (-\tfrac{1}{4}v_1^2 - \tfrac{1}{2}v_1 v_2) dF(v_2) + \int_{NI \cap NC} (-v_1 v_2) dF(v_2) \right]
\end{aligned}
\tag{15}
$$

where the probability of the second invention would depend on the regime to which it would be assigned. Again, we must integrate within each regime separately. Note that only under the I regime can firm 1 appropriate some value from the improvement; in any NI regime, the entry of firm 2 in period 2 will reduce firm 1's payoff. Thus, the patent-antitrust policy that assigns all v_2 (from 0 to V_2) to the I regime would maximize $P(v_1)$; nevertheless, in that case, $P(v_1) < V_2^2/4C_2$. Consistent with assumption (6), then, let $C_1 \geq 2V_1 + V_2^2/4C_2$. We can then derive the following propositions:

Proposition 2: For any $v_1 > 0$, the optimal patent-antitrust policy is infringement for v_2 below a cutoff value, which lies strictly between 0 and V_2, but noninfringement (either with collusion or without collusion) for v_2 above this cutoff value.

Remark: In general, courts should find that improvements infringe on the original patent, so as to preserve the incentives to the first inventor. Of the three regimes, this one most often provides the payoffs closest to the second-best described above in Part III.A. In terms of the second-best policy, this regime chooses the α that maximizes the contribution of the (v_1, v_2) pair to $P(v_1)$, but at the expense of making its contribution to $S(v_1)$ too low.

For a sufficiently large gain in terms of encouraging the second innovator, it will be worthwhile to curtail the first innovator's patent protection at the margin. In particular, consider the noninfringement regime without collusion as an alternative for some (v_1, v_2) pairs. This regime maximizes the contribution a pair makes to $S(v_1)$, but minimizes its contribution to $P(v_1)$. Thus, switching some (but not all) pairs to this alternative, then, would serve to offset the infringement regime's deficiencies relative to the second-best policy.

Placing especially valuable improvements outside the scope of the original patent provides a relatively large gain compared to the sacrifice in the first inventor's expected profits. The benefits (in terms of $S(v_1)$) of switching to this alternative would derive from increasing π_2 from $\frac{1}{2}v_2$ to v_2. These benefits, then, would increase with v_2 and would be independent of v_1. The costs (in terms of $P(v_1)$) would derive from reducing $\Delta\pi_1$ from $\frac{1}{2}v_2$ to $-v_1$. For any $v_1>0$, then, these costs would also increase with v_2, but at a proportionally slower rate than the benefits would increase, because the sacrifice in terms of firm 1's profits would depend on v_1 as well as v_2. For any given $v_1>0$, then, a switch to this alternative yields the greatest net benefit (i.e., at a relatively small cost) for the largest v_2.[22] Thus, it will prove optimal to allow a subsequent inventor to market a particularly valuable improvement -- one with a high v_2/V_2 ratio -- without sharing its value with the first inventor.

Thus, for any given $v_1>0$, only the most valuable improvements should avoid infringement. Note that this result emerges not because subservient improvement patents, which permit holders of dominant patents to "holdup" subsequent inventors, are a less attractive option for the courts as v_2 rises to such high levels, as Merges and Nelson (p. 866 & n.117) suggest. Proposition 1 reveals that ideally the original inventor would be entitled to the same share of v_2, whether that improvement is very valuable or not. Rather, the cutoff value for infringement emerges because the alternative (noninfringement) becomes relatively more attractive as v_2 grows larger: its social benefits grow faster than its social costs.

Proof: See Appendix B.

Proposition 3: As v_1 approaches 0, the cutoff value for noninfringement (either with or without collusion) rises toward V_2. For $v_1=0$, the optimal patent-antitrust policy is infringement for any v_2.

Remark: Recall that under the second-best policy, the optimal share α for firm 2

[22]Conversely, for any v_1, the infringement regime will be better than this alternative for sufficiently small v_2. The same will be true of the other alternative, noninfringement with collusive licensing. See the Remark accompanying Proposition 4.

increases in v_1, because the need to increase $S(v_1)$ becomes more important relative to the need to increase $P(v_1)$. Therefore, the infringement regime comes closer to the second-best policy as v_1 approaches 0. As v_1 grows smaller, there is less scope for a welfare improvement by switching any (v_1, v_2) pairs to another regime.

When the first invention has no economic value except as a necessary condition to subsequent innovation, then the only policy that can permit the first inventor to appropriate any reward at all is a finding of infringement. Indeed, if $v_1=0$, then the probability of the first invention is so small that the need to improve firm 1's incentives (i.e., increase $P(v_1)$) outweighs the countervailing objective of rewarding subsequent inventors (i.e., to increase $S(v_1)$). Thus, although the economic value of the second invention is much greater than the original invention, a finding of noninfringement would never improve welfare.

The most important implication of this proposition is a criterion for the court's finding of noninfringement that is more complex than that described by Merges and Nelson (1990, pp. 865-66). The decision is not simply a function of the v_2/v_1 ratio. Indeed, the cutoff value v_2 for the second product not to infringe is not even monotonically increasing in v_1. Instead, this proposition reveals that this cutoff value is <u>decreasing</u> in v_1, at least when v_1 is small relative to the value of possible improvements. Thus, the v_2/v_1 ratio is not sufficient to distinguish cases in which the courts should find infringement from those in which the courts should not. Instead, the optimal policy depends on both v_1/V_2 and v_2/V_2. Specifically, in cases of basic inventions with low v_1/V_2 ratios, courts should establish a very high v_2/V_2 hurdle for subsequent inventors to avoid infringement.

This result contradicts the reasoning of the Supreme Court in the <u>Westinghouse</u> case, in which the Court noted that although the basic patent in that case was a "pioneer" insofar as it broke new ground in the technology for train brakes, the design was not a success until later improvements perfected it. The Court cited the patent's shortcomings as a reason for a narrower range of equivalents. This proposition, however, indicates that the argument for noninfringement is weakest in just such a case: because the basic patent by itself lacked great commercial value, the original inventor should be entitled to receive a large share of the value produced by the improvements inspired by his own pioneering invention. Only the most extraordinary improvement among the universe of possible improvements -- the improvement that contributes such value that an invention of its importance would be deemed very unlikely ex ante -- should be permitted to escape infringement of such a pioneer invention.

Proof: See Appendix B.

Proposition 4: For sufficiently small v_1, collusive licensing agreements will be optimal for all noninfringement cases. For sufficiently large v_1, including any $v_1 > v_2$, however, it will be optimal to prohibit collusive licensing agreements in all noninfringement cases.

Remark: Switching a case from infringement to noninfringement with collusion can increase the contribution a (v_1, v_2) pair makes to $S(v_1)$, but must reduce its contribution to $P(v_1)$. For any v_1, this alternative is most attractive when v_2 is large. Thus, switching some (but not all) of the largest v_2 values to this alternative, just as switching some to noninfringement without collusion, can serve to offset the infringement regime's deficiencies relative to the second-best policy.

Noninfringement with collusion, however, provides firm 2 with excessive rewards that may encourage inefficient invention, simply to threaten entry and share in firm 1's revenues.[23] This excess incentive grows as v_1 grows larger. Indeed, if $v_2 < v_1$, then this alternative would be unambiguously inferior to the infringement regime: both $S(v_1)$ and

[23]In the Green and Scotchmer (1990) model, the courts can always avoid this problem. In their model, the second innovator is uncertain of the value of its invention, and the distribution of this variable is the same for all inventors of improvements who could be subject to the patent policy. Therefore, even with collusive licensing, the courts can adjust the patent scope so that the expected payoff to the second innovator is always exactly equal to the expected value of the second innovation. In contrast, the model in this paper incorporates the possibility that second innovators face different prospects and thus explores the implications of inefficient entry. The Kaplow (1984) model addresses the problem of wasteful R&D by imitators who "invent around" the original patent, but does not introduce the possibility of "imitators" who contribute valuable improvements. Nor does he include patent scope as a policy instrument, which allows courts to decide whether imitators have competing patents or merely complementary patents that infringe on the original patent.

$P(v_1)$ would be lower. That is, this alternative would not only reduce firm 1's profits and incentives to invent, but also distort firm 2's incentives even more than the infringement regime would. If the improvement is less valuable than the original invention, then the overinvestment by firm 2 under that alternative regime would be larger than the underinvestment by firm 2 under the infringement regime.

Nevertheless, for small v_1, noninfringement with collusion can pose less of a threat to the profits for firm 1 than noninfringement with competition does. Although competition gives the second inventor the full surplus from product 2, and so maximizes the contribution of a (v_1, v_2) pair to $S(v_1)$, it also erodes the incentives for the first invention more (than noninfringement with collusion does) in the event of the second invention. If v_1 is large relative to v_2, on the other hand, then the excess incentive for firm 2 under the collusion regime would be significant: it could increase the probability of entry by firm 2 by such a large factor that the regime with collusion erodes firm 1's expected profits more than the regime with competition.

The relative merits of the two noninfringement regimes, then, will vary significantly with v_1. If the first invention is not so valuable, then the risk of inefficient R&D by firm 2 is relatively insignificant, and the need to maintain the firm 1's incentives (i.e., increase $P(v_1)$) is great relative to the need to optimize firm 2's incentives (i.e., increase $S(v_1)$). Therefore, collusion is optimal in such noninfringement cases (as Green and Scotchmer suggest). If the first invention is valuable relative to improvements, however, then the risk of inefficient entry is significant, and the first inventor's incentives are less important. Therefore, competition is optimal in such noninfringement cases (as Kaplow argues).

Proof: See Appendix B.

Proposition 5: As v_1 grows very large relative to V_2, the cutoff value for noninfringement (without collusion) rises toward V_2.

Remark: For any given policy, as v_1 grows larger, the need to increase $S(v_1)$ tends to grow relative to the need to increase $P(v_1)$. This effect tends to make the noninfringement regime (without collusion) attractive relative to the infringement regime, because it gives firm 2 better incentives. A second effect, however, opposes the first: an increase in v_1 implies that such a noninfringement finding entails a greater sacrifice in firm 1's incentives, and thus $P(v_1)$ would decline (whereas $S(v_1)$ would be unaffected) insofar as patent policy includes the noninfringement regime. This increase in the distortion of firm 1's incentives tends to make the need to increase $P(v_1)$ more important relative to the need to increase $S(v_1)$, which in turn makes the infringement regime (which better protects firm 1's profits) more attractive. Furthermore, if v_1 is large, then so is the cost of a noninfringement finding in terms of $P(v_1)$, relative to the corresponding gain in terms of $S(v_1)$. Therefore, if v_1 is large, then the second effect dominates the first effect: the noninfringement regime becomes less attractive as v_1 grows still larger.

In the limit, the value of all possible improvements becomes trivial in comparison with the original invention. In this case, the social costs of diminishing firm 1's incentives outweigh the social benefits of encouraging improvements, even those with high v_2/V_2 ratios. If v_1/V_2 is large, then the optimal policy would confront firm 1 with only a small ex ante probability of noninfringement by setting a high v_2/V_2 hurdle for firm 2.

Proof: See Appendix B.

IV. Conclusions and Issues for Further Research

Proposition 2 supports the suggestion that courts respond to the holdup problem in cumulative innovation by placing particularly valuable improvements outside the scope of the original patent, at least as long as the original patent has some positive economic value. Thus, the paper offers a basis in economic theory for the courts' reliance on the value of the inventions in infringement cases. Courts, however, should not find that any improvement avoids infringement. Given the option of subservient improvement patents, courts should be circumspect in allowing imitators to "invent around" a patent. Thus, the optimal patent policy may entail broader or narrower patent protection than courts currently provide.

Although the model supports the general practice of declaring improved products not "equivalents" of patented predecessors, the model also suggests a refinement in how courts take the value of these inventions into account. In particular, pioneer inventions should include two distinct classes of basic inventions: not only those that are very valuable relative to possible improvements (Proposition 5), but also those that have very little value relative to the improvements that one would expect to follow (Proposition 3). Courts should grant a broad patent range of equivalents for inventions falling toward either extreme; if instead a basic invention would be expected to generate improvements with values on the same order of magnitude as the value of the basic invention, then it should receive relatively narrow patent protection.

That is, contrary to the Supreme Court's suggestion in the Westinghouse case, courts should recognize an invention with little stand-alone value as a pioneer if it provides "stepping stones" for others to render the invention much more valuable. Thus, the definition of a pioneer invention should not turn simply on the commercial value it has standing alone. Although a basic invention may have trivial commercial value by itself, it may also be a technological breakthrough in that it generates great spillovers in the form of improvements likely to be far more valuable than the basic invention itself. Like basic research, such pioneer inventions present a strong case for some form of subsidy because a private inventor is unlikely to undertake such R&D at levels commensurate with their social value. The doctrine of equivalents, then, can help such pioneer inventors appropriate the external benefits of their research.

The foregoing discussion presumes that the original invention is a necessary condition for the allegedly infringing innovation that follows. As the discussion in Part III indicated, if the patented innovation does nothing to facilitate the subsequent innovation, then a finding of noninfringement (with collusion prohibited) would be appropriate. This analysis supports the critique of cases like Genentech, which uphold product patents for those who develop means of producing natural substances. As Merges and Nelson point out, such a policy inhibits R&D in new biotechnology like recombinant DNA research. Furthermore, because such new techniques do not build on the older patented techniques, there is no reason that holders of patents on those older technologies should profit from such unrelated subsequent research. The holdup problem in such cases is unilateral, not bilateral: it could be eliminated entirely by assigning control over the new technology to the second innovator alone. Patents should protect against imitators who would otherwise free-ride on the work of patentees, not independent innovators who could not possibly gain any benefit from the patentee's work.

Finally, the model provides very little support for lax antitrust scrutiny of collusion between patentees, even though the model excludes considerations of deadweight loss. Such collusion would be not only undesirable outside of the context of cumulative innovation, as in the Genentech situation, but also unnecessary between holders of dominant and subservient patents. Collusion between holders of competing patents, even in the context of cumulative innovations which Green and Scotchmer discuss, would be desirable only in limited circumstances. Once we recognize the potential for inefficient entry, a policy favoring collusion proves to be an unappealing alternative. Proposition 4 indicates that such a policy would only be superior to the alternatives only if v_2 is large, both relative to v_1 (so that the risk of inefficient entry is relatively small) and relative to most other possible improvements (so as to justify a finding of noninfringement, which should be rare when v_1 is small).

It is difficult to imagine an antitrust defense tailored to these circumstances. If a court would find these cases at all difficult to identify, the best policy would be a uniform rule against collusion. After all, in the circumstances described by Proposition 4, noninfringement without collusion is likely to be an adequate substitute for noninfringement with collusion: if v_1 is small relative to v_2, then the improvements in the two inventors' payoffs when they collude will be correspondingly small. Collusive licensing, then, seems to be a poor response to the bilateral holdup problem.

These conclusions, however, should be qualified: the model presented here include many simplifying assumptions to render the analysis tractable. Further research is necessary to determine how robust the results in this paper are when these assumptions are relaxed. Given the many complexities excluded from the model, one must be cautious in drawing implications for public policy. Even given this uncertainty, however, the preliminary analysis undertaken here does contribute some basic insights that one would expect to survive further study.

For example, the model assumes that the parties litigate and do not settle out of court, as they did in the Standard Oil case. Insofar as litigation costs are symmetric, however, settlements should reflect the expected outcome of trials, and the ex ante incentives should be the same. Similarly, the model assumes that courts have full information and make no errors. Even with judicial error, however, courts could create the same incentives ex ante as long as they offer the firms the correct payoffs on average. Indeed, to the extent there is a stochastic element in infringement decisions, courts can attempt to do better than the third-best policy described in Part III.B: they can pursue a randomized "mixed strategy" that better approximates the expected payoffs of the ideal compulsory licensing policy described in Part III.A.[24]

Settlements may not reflect the expected outcome of a trial, however, to the extent that settlements allow collusion that would not be permitted otherwise. Antitrust scrutiny of such settlements, then, should ensure that patentees with substitute (rather than complementary) innovations do not fix prices or otherwise collude. To the extent that rival patents actually derive from complementary innovations, particularly those that arise in the context of cumulative innovation, the patentees should be allowed to cross-license with per-unit royalties. This arrangement may lead to monopoly pricing, but as Part III.A revealed, the optimal policy would allow such innovators to maximize their joint producer surplus. To the extent that courts find complementary and competing patents difficult to distinguish, however, antitrust policy should not be so permissive as to allow explicit price-fixing. See Kaplow (1984, p. 1862) and Priest (1977, p. 358).

A complete analysis of the effects of monopoly pricing would include the deadweight loss created by pricing above marginal cost. One could relax the assumption of unit demands,[25] address the case of downward-sloping demand curves, and thereby introduce

---

[24]This task is easier if the first innovation is necessary for the second innovation, as the model in this paper assumes. In reality, courts may be uncertain whether the second innovation required the first. One could extend the model to consider other possible relationships between the innovations: the second innovation may be independent of the first, it may have occurred earlier than it would otherwise because of the first, or it may have cost less because of the first. To the extent a finding of infringement leads to excessive rewards for the first innovator in these circumstances, a finding of noninfringement may be preferable.

[25]This simplifying assumption allowed us to focus on the trade-off between the incentives to the two innovators. This assumption may be reasonable for a process

the deadweight loss from patent monopolies. It is ambiguous, however, whether this change in the model would make collusive pricing more or less attractive. Of course, the deadweight loss militates in favor of competition and lower prices. Downward-sloping demand curves, however, would also imply consumer surplus even at monopoly prices, which in turn would imply that innovators always reap too little profit to encourage all socially desirable R&D. This consideration militates in favor of monopoly pricing.[26]

This paper includes assumptions in the model that ensure that patents tend to offer inventors too little incentive to invest in R&D, in the belief that this problem is most important. For example, whereas the model assumed that there is no rivalry in R&D, the literature using "patent race" models indicate that rivalry in R&D can lead to excessive R&D expenditures. See Reinganum (1989). To include the case of excessive incentives for R&D, one might also extend the model to include this effect of patent races.[27]

---

innovation that reduces the cost of producing a given product. Although this paper describes the model as one of product patents, the analysis applies to process patents as well. The assumption of unit demands will also approximate reality whenever consumer surplus and deadweight loss are small, which may be the case when producers can price discriminate among consumers or when the market demand curve is concave.

[26]The optimal balance between these considerations would depend on the particular shape of the demand curve, the nature of the improvement, and other factors that would be difficult for a court to observe. Further research may prove useful. The simple model in this paper, however, may be a better guide to policy than a more complex model that implies a different result only in special cases that would be difficult to distinguish from the usual case.

[27]Further research could also explore whether the results in this model generalize to a sequence of innovations in an n-stage model. Other extensions that could qualify the results in this paper include horizontal product differentiation, see Klemperer (1990) and Waterson (1990), ex ante agreements to form research joint ventures, see Green and Scotchmer (1990), and patent length as another policy instrument, see Gilbert and Shapiro (1990), Klemperer (1990), and Tandon (1982).

APPENDIX A

Proof of Proposition 1: The expressions involving α in (9) and (10) may be brought outside the integrals, because α is independent of v_2. In each equation, what is left of the integral is simply $E(v_2^2)$. Substituting these revised expressions for $S(v_1)$ and $P(v_1)$ in (12), and rearranging terms, yields the following quadratic equation:

$$3E(v_2^2)\alpha^2 + [4v_1C_2 - 2E(v_2^2)]\alpha - 4v_1C_2 = 0. \tag{A1}$$

Solving for α using the quadratic formula, and excluding any solutions for α that are negative, we obtain (13) as the unique solution for $\alpha(v_1)$.

Total differentiation of (A1) yields:

$$[(6\alpha-2)E(v_2^2) + 4v_1C_2]\, d\alpha + [4C_2(\alpha-1)]\, dv_1 + [4v_1(\alpha-1)]\, dC_2 + [\alpha(3\alpha-2)]\, dE(v_2^2) = 0, \tag{A2}$$

which implies our comparative statics results. The partial derivative of α with respect to v_1 is strictly positive, because we know from the proof of Lemma 4 that $\frac{1}{2}<\alpha<1$. Similarly, the partial derivative of α with respect to C_2 is nonnegative, and strictly positive if $v_1>0$. For $v_1>0$, α must be greater than $\frac{2}{3}$, because $\alpha'(v_1) > 0$ and (13) implies that if $v_1=0$, then $\alpha=\frac{2}{3}$. Therefore, the partial derivative of α with respect to $E(v_2^2)$ is nonpositive, because $\alpha\geq\frac{2}{3}$, and is strictly negative if $v_1>0$, because then $\alpha>\frac{2}{3}$. ■

APPENDIX B

In this appendix, I prove the propositions presented in Part III.B. To characterize the optimal policy, it is useful to describe first the regions in (v_1, v_2) space where each regime dominates the others. I compare the three regimes in a series of three pairs.

First, given the optimal policy, consider the hypothetical change in welfare resulting from switching a point in (v_1, v_2) space from I to NI∩NC. Switching one such point in (14) would entail a $dS(v_1)$ proportional to $\frac{1}{8}v_2^2f(v_2)/C_2$, and in (15), a $dP(v_1)$ proportional to $-(\frac{1}{4}v_2^2+v_1v_2)f(v_2)/C_2$. Therefore, we know from (8) that $dW(v_1)$ would be positive or negative as:

$$\tfrac{1}{8}[2v_1+P(v_1)]v_2^2 - [S(v_1)-P(v_1)](\tfrac{1}{4}v_2^2+v_1v_2)$$

is positive or negative. Excluding points where $v_2=0$ (and thus where $dW(v_1)=0$), then, $dW(v_1)$ would be positive or negative as:

$$2v_1v_2 - [2S(v_1)-3P(v_1)]v_2 - 8[S(v_1)-P(v_1)]v_1 \tag{B1}$$

is positive or negative. We can see from (14) and (15) that $0 \leq 2S(v_1)-3P(v_1)$. Thus, if $2v_1 \leq 2S(v_1)-3P(v_1)$, then (B1) implies that $dW(v_1)<0$ for any $v_1>0$, and I will strictly dominate NI∩NC for any v_2. If $2v_1 > 2S(v_1)-3P(v_1)$ instead, then (B1) implies that NI∩NC will dominate I for any:

$$v_2 > 8[S(v_1)-P(v_1)]v_1/\{2v_1-[2S(v_1)-3P(v_1)]\}, \tag{B2}$$

and I will dominate NI∩NC if the inequality is reversed.

Second, consider the hypothetical change in welfare resulting from switching a point in (v_1, v_2) space from I to NI∩C. Switching one such point in (14) would entail a $dS(v_1)$ proportional to $\frac{1}{8}(v_2^2-v_1^2)f(v_2)/C_2$, and in (15), a $dP(v_1)$ proportional to $-(\frac{1}{4}v_1^2+\frac{1}{4}v_2^2 + \frac{1}{2}v_1v_2)f(v_2)/C_2$. Therefore, we know from (8) that $dW(v_1)$ would be positive or negative as:

$$\tfrac{1}{8}[2v_1+P(v_1)](v_2^2-v_1^2) - [S(v_1)-P(v_1)](\tfrac{1}{4}v_1^2+\tfrac{1}{4}v_2^2+\tfrac{1}{2}v_1v_2)$$

is positive or negative. Excluding the origin, where $v_1+v_2 = 0$ (and thus where $dW(v_1)=0$), then, $dW(v_1)$ would be positive or negative as:

$$2v_1v_2 - 2v_1^2 - [2S(v_1)-P(v_1)]v_1 - [2S(v_1)-3P(v_1)]v_2 \tag{B3}$$

is positive or negative. We can see from (14) and (15) that $0 < 2S(v_1)-P(v_1)$. Thus, if $2v_1 \leq 2S(v_1)-3P(v_1)$, then (B3) implies that $dW(v_1)<0$ for any $v_1>0$, and I will strictly dominate NI∩C for any v_2. If $2v_1 > 2S(v_1)\text{-}3P(v_1)$ instead, then (B3) implies that NI∩C will dominate I for any:

$$v_2 > \{2v_1^2+[2S(v_1)-P(v_1)]v_1\}/\{2v_1-[2S(v_1)-3P(v_1)]\}, \tag{B4}$$

and I will dominate NI∩C if the inequality is reversed.

Finally, consider the hypothetical change in welfare resulting from switching a point in (v_1, v_2) space from NI∩NC to NI∩C. Switching one such point in (14) would entail a $dS(v_1)$ proportional to $-\frac{1}{8}v_1^2 f(v_2)/C_2$, and in (15), a $dP(v_1)$ proportional to $(\frac{1}{2}v_1v_2-\frac{1}{4}v_1^2)f(v_2)/C_2$. Therefore, we know from (8) that $dW(v_1)$ would be positive or negative as:

$$-\tfrac{1}{8}[2v_1+P(v_1)]v_1^2 + [S(v_1)-P(v_1)](\tfrac{1}{2}v_1v_2-\tfrac{1}{4}v_1^2)$$

is positive or negative. Excluding $v_1=0$ (and thus where $dW(v_1)=0$), then, $dW(v_1)$ would be positive or negative as:

$$4[S(v_1)-P(v_1)]v_2 - 2v_1^2 - [2S(v_1)-P(v_1)]v_1 \tag{B5}$$

is positive or negative. Thus, NI∩C will dominate NI∩NC for any:

$$v_2 > \{2v_1^2+[2S(v_1)-P(v_1)]v_1\}/4[S(v_1)-P(v_1)], \tag{B6}$$

and NI∩NC will dominate NI∩C if the inequality is reversed.

Given this information, we can now derive the following propositions:

<u>Proof of Proposition 2</u>: For any $v_1>0$, we know the optimal policy must be the I regime for some positive v_2 sufficiently close to 0. Consider the limit of (B1) or (B3) for any v_1 as v_2 grows arbitrarily small. These limits are negative: therefore, a switch from either NI∩NC or NI∩C to I would improve welfare. That is, the I regime dominates either alternative for sufficiently small v_2.

Nevertheless, for any $v_1>0$, the optimal policy cannot rely on the I regime exclusively. Otherwise, we know from (14) and (15) that $S(v_1)-P(v_1) < V_2^2/8C_2$, and $2S(v_1)-3P(v_1) = 0$. These relationships, together with (B1), would imply that for any $v_1>0$ a shift from I to NI∩NC for a v_2 sufficiently close to V_2 will yield a welfare improvement, which would contradict the supposed optimum. (Recall that $C_2>V_2$.) Therefore, for any $v_1>0$, the optimal policy must instead feature at least one cutoff function, either the critical v_2 in (B2) or that in (B4), strictly between $v_2=0$ and $v_2=V_2$ to separate the infringement from the noninfringement regimes. ■

Proof of Proposition 3: We know that $2v_1 > 2S(v_1)$-$3P(v_1)$ for all $v_1>0$. If we were to suppose otherwise, then both (B1) and (B3) would be negative for all $v_1>0$, as a comparison of (14) and (15) confirms. If both (B1) and (B3) were negative, however, then a shift for any $v_2>0$ from either NI$\cap$NC or NI$\cap$C to I would yield a welfare improvement, which would contradict Proposition 2.

We also know from Proposition 2, (14), and (15) that $2S(v_1)$-$3P(v_1) > 0$ for any $v_1>0$. As v_1 becomes arbitrarily close to 0, then, $2S(v_1)$-$3P(v_1)$ must stay between $2v_1$ and 0, which can occur only if fewer and fewer v_2 values are assigned to either NI$\cap$NC or NI$\cap$C. Therefore, the lowest cutoff value for noninfringement, either the critical v_2 in (B2) or that in (B4), must rise toward V_2 as v_1 approaches 0.

If $v_1=0$, then no $v_2>0$ can be assigned to either NI$\cap$NC or NI$\cap$C. If we were to suppose otherwise, then we know from (14) and (15) that $0 < 2S(0)$-$3P(0)$. This inequality, together with (B1) and (B3) evaluated at $v_1=0$, imply that a shift for any $v_2>0$ from either NI$\cap$NC or NI$\cap$C to I would yield a welfare improvement, which would contradict the supposed optimum. (Note that at the origin in (v_1, v_2) space, policy is irrelevant, because any one of the three regimes would yield the same welfare ex post: zero.) ■

Proof of Proposition 4: Consider the critical value for v_2 in (B6), which denotes indifference between NI$\cap$C and NI$\cap$NC. This critical value for v_2 falls to 0 as v_1 grows arbitrarily close to 0. Furthermore, Proposition 3 implies that the lowest cutoff value for noninfringement -- either that in (B2) or that in (B4) -- grows close to V_2 (and is therefore bounded away from $v_2=0$) as v_1 approaches 0. Therefore, for sufficiently small v_1, all noninfringement cases fall in the region where NI$\cap$C will dominate NI$\cap$NC.

Nevertheless, NI$\cap$NC must dominate NI$\cap$C for larger values of v_1. In particular, consider (B3) where $v_2=v_1$. Note that (B3) must be negative along the $v_2=v_1$ line. Therefore, the cutoff function for NI$\cap$C in (B4) must always lie strictly above this line. For any $v_1 \geq v_2$, then, collusive licensing cannot be optimal in noninfringement cases. For any $v_1 > V_2$, for example, courts should prohibit collusive licensing agreements in all cases of noninfringement. ■

Proof of Proposition 5: Proposition 4 implies that for large v_1, we can exclude the NI∩C regime from (14) and (15). In these cases, the critical v_2 in (B2) is the relevant cutoff value for noninfringement. Let $v_2^*(v_1)$ denote this cutoff function. Thus, in our expressions for $S(v_1)$ in (14) and for $P(v_1)$ in (15), we would integrate with respect to v_2 from 0 to $v_2^*(v_1)$ for the I regime and from $v_2^*(v_1)$ to V_2 for the NI∩NC regime. Although we cannot solve explicitly for $v_2^*(v_1)$, we can substitute our new expressions for $S(v_1)$ and $P(v_1)$ into (B2). Multiplying both numerator and denominator by C_2/v_1, we derive the following condition:

$$v_2^*(v_1) = \frac{\int_0^{v_2^*(v_1)} v_2^2 dF(v_2) + \int_{v_2^*(v_1)}^{V_2} (4v_2^2 + 8v_1v_2) dF(v_2)}{2C_2 - \int_{v_2^*(v_1)}^{V_2} [(v_2^2/v_1) + 3v_2] dF(v_2)}. \tag{B7}$$

Consider the right-hand side of (B7) and suppose that $v_2^*(v_1)$ is simply a constant. As v_1 grows larger, both the numerator and the denominator tend to increase. For a sufficiently large v_1, however, the numerator will increase by a greater proportion than the denominator as v_1 increases; that is, the right-hand side of (B7) is increasing in v_1. Then the equality (B7) cannot hold if $v_2^*(v_1)$ is constant as v_1 increases. Nor can (B7) hold if $v_2^*(v_1)$ is decreasing in v_1, for this would only cause the right-hand side of (B7) to rise faster: $v_2^{*\prime}(v_1) < 0$ would add another positive effect upon the numerator and a negative effect upon the denominator. Instead, for sufficiently large v_1, $v_2^*(v_1)$ must increase in v_1, which will dampen the increase in the right-hand side of (B7). Proposition 2 implies that $v_2^*(v_1)$ must remain below V_2, and to keep the right-hand side of (B7) bounded as v_1 increases without bound, $v_2^*(v_1)$ must approach V_2 asymptotically. ■

REFERENCES

Adelman, Martin J. (1977), "Property Rights Theory and Patent-Antitrust: The Role of Compulsory Licensing," New York University Law Review, Vo. 52, 977-1013.

Areeda, Phillip and Louis Kaplow (1988), Antitrust Analysis: Problems, Text, Cases, 4th ed. (Boston: Little, Brown & Co.).

Areeda, Phillip and Donald F. Turner (1978), Antitrust Law, Vol. 3 (Boston: Little, Brown & Co.).

Caves, Richard E., Harold Crookell, and J. Peter Killing (1983), "The Imperfect Market for Technology Licenses," Oxford Bulletin of Economics and Statistics, Vol. 45, 249-267.

Chisum, Donald S. (1990), Patents (New York: Matthew Bender & Co., Inc.).

Gilbert, Richard and Carl Shapiro (1990), "Optimal Patent Length and Breadth," RAND Journal of Economics, Vol. 21, 106-112.

Green, Jerry R. and Suzanne Scotchmer (1990), "Antitrust Policy, the Breadth of Patent Protection and the Incentive to Develop New Products," School of Law, University of California, Berkeley, John M. Olin Working Paper in Law and Economics No. 90-12.

Grossman, Sanford J. and Oliver D. Hart (1986), "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," Journal of Political Economy, Vol. 94, 691-719.

Kaplow, Louis (1984), "The Patent-Antitrust Intersection: A Reappraisal," Harvard Law Review, Vol. 97, 1813-1892.

Klein, Benjamin, Robert G. Crawford, and Armen A. Alchian (1978), "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process," Journal of Law and Economics, Vol. 21, 297-326.

Klemperer, Paul (1990), "How Broad Should the Scope of Patent Protection Be?", RAND Journal of Economics, Vol. 21, 113-130.

Laffont, Jean-Jacques and Jean Tirole (1990), "The Politics of Government Decision Making: Regulatory Institutions," Journal of Law, Economics, and Organization, Vol. 6, 1-31.

Landes, William M. and Richard A. Posner (1989), "An Economic Analysis of Copyright Law," Journal of Legal Studies, Vol. 18, 325-363.

Lipscomb, Ernest B. (1987), Walker on Patents, 3d ed., Vol. 6 (Rochester, N.Y.: Lawyers Cooperative Publishing Co.).

Menell, Peter S. (1989), "An Analysis of the Scope of Copyright Protection for Application Programs," Stanford Law Review, Vol. 41, 1045-1104.

Merges, Robert P. and Richard R. Nelson (1990), "On the Complex Economics of Patent Scope," Columbia Law Review, Vol. 90, 839-916.

Posner, Richard A. (1976), Antitrust Law: An Economic Perspective (Chicago: University of Chicago Press).

Priest, George L. (1977), "Cartels and Patent License Arrangements," Journal of Law and Economics, Vol. 20, 309-377.

Reinganum, Jennifer F. (1989), "The Timing of Innovation: Research, Development, and Diffusion," Ch. 14 of Handbook of Industrial Organization, Vol. 1, edited by Schmalensee, Richard and Robert D. Willig (New York: North-Holland).

Scherer, F.M. (1980), Industrial Market Structure and Economic Performance, 2d ed. (Boston: Houghton Mifflin Co.).

Scotchmer, Suzanne (1991), "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law," Journal of Economic Perspectives, Vol. 5, 29-41.

Sullivan, Lawrence A. (1977), Antitrust (St. Paul, Minn.: West Publishing Co.).

Tandon, Pankaj (1982), "Optimal Patents with Compulsory Licensing," Journal of Political Economy, Vol. 90, 470-486.

Waterson, Michael (1990), "The Economics of Product Patents," American Economic Review, Vol. 80, 860-869.

Williamson, Oliver E. (1979), "Transaction-Cost Economics: The Governance of Contractual Relations," Journal of Law and Economics, Vol. 22, 233-261.