The Decision to Patent, Cumulative Innovation, and Optimal Policy*

Nisvan Erkal[†]
University of Melbourne

First version: October 2001
This version: May 2003

Abstract

Optimal patent breadth is an issue that is still being vigorously debated at both the theoretical and empirical levels. This paper analyzes optimal patent policy in the context of cumulative innovation in a model that endogenizes the patenting decisions of early innovators. In the theoretical literature on cumulative innovation, it is generally assumed that all innovations are patented. However, studies such as Cohen et al. (2000) and Levin et al. (1987) report that firms frequently rely on secrecy to protect their discoveries. Cumulative innovation implies that innovators may have significant incentives to keep their innovations secret to get a head start in subsequent R&D races. This paper shows that if innovators cannot rely on secrecy to protect their innovations, it is optimal to have relatively narrow patent protection. This happens if the government has a weak trade secret policy or if innovators cannot monitor the flow of their technological information. This is because when innovators cannot rely on secrecy to protect their innovations, they have increased incentives to patent them and it is not necessary for the government to give them extra incentives to patent. In the case when innovators always prefer secrecy over patenting, it becomes optimal to have a flexible antitrust policy rather than a flexible patent policy. Since non-disclosure reduces the investment incentives in the second R&D race, allowing collusive licensing agreements between competing innovators becomes optimal in order to stimulate investment in the second R&D race.

JEL Classification: O31, O38, L40

Keywords: cumulative innovation, patenting decision, patent policy, antitrust policy

1

1 Introduction

Recent developments in the economic theories of innovation have challenged the notion that innovations are isolated events and emphasized the cumulative nature of innovation. The cumulative nature of innovation implies that the social value of innovations should include the value of subsequent innovations they inspire. This has significant implications for optimal patent design, especially optimal patent breadth. Hence, one of the main issues addressed in the cumulative innovation literature is how patent policy can be designed so as to divide the profits between sequential innovators in a way that provides them with optimal incentives to invest. Several papers have argued for broad patent protection in order to overcome the externality that exists between different generations of innovators.[1] One exception is Denicolo (2000). In a two-stage patent race framework, he shows that having broad protection may not always be optimal.

The theoretical debate over optimal patent policy has significant policy implications. Patent protection in the United States has become much stronger over the last two decades as a result of some recent reforms. The success of these reforms in stimulating innovation has been called into question in several empirical papers.[2] Therefore, the desirability of broad patent protection is an issue that is still being vigorously debated at both the theoretical and empirical levels.

In the theoretical literature on cumulative innovation, it is generally assumed that all innovations are patented. This paper challenges this assumption. I show that having broad

---

[1] See Scotchmer (1991 and 1996), Green and Scotchmer (1995), Chang (1995), and Matutes, Regibeau and Rockett (1996). Gallini and Scotchmer (2001) contains an excellent survey of the different approaches to the problem of optimal patent breadth.

[2] See Gallini (2002) for an excellent discussion of the effectiveness of patent protection and the recent patent reforms in the United States. Schankerman (1998) and Lanjouw (1998) find that patent protection is not the major source of private returns to R&D. Hall and Ziedonis (2001), and Lanjouw and Schankerman (2001) argue that stronger patent protection have resulted in socially wasteful patent portfolio races instead of stimulating innovation.

patent protection may not always be optimal if we take into account the fact that innovators choose between several mechanisms, including patents and trade secrecy, to protect their innovations. In fact, studies such as Cohen et al. (2000) and Levin et al. (1987) report that firms frequently rely on secrecy to protect their discoveries. Specifically, Cohen et al. (2000) find that patents tend to be the least preferred protection mechanism by firms while secrecy and lead time tend to be the most heavily used ones.[3] Moreover, by comparing their results with those of Levin et al. (1987), Cohen et al. (2000) conclude that there is an apparent growth in the importance of secrecy as an appropriability mechanism and a decline in the importance of patents. The importance of secrecy is further supported by Lerner (1994) who finds that 43% of all intellectual property litigation cases involve trade secrets.

The cumulative nature of innovation implies that early innovators may have significant incentives to keep their innovations secret to get a head start in subsequent R&D races. In an environment where firms compete to come up with improved versions of current products, the profits of early innovators can be significantly reduced if the following race is won by one of the rival firms. This is especially true under a policy of narrow patent protection. When early innovators prefer secrecy over patenting, the dynamics of subsequent R&D races change substantially since non-disclosure of early innovations can severely affect the investment incentives of rival firms in subsequent R&D races. Therefore, the analysis of optimal intellectual property policy should take into account the possibility that early innovators may hinder the pace of innovation by delaying the disclosure of their innovations.[4]

Does this mean that the government should have a policy of broad patent protection in order to encourage patenting? The two goals of the patent system are to encourage research

---

[3] According to their study, especially in case of product innovations, firms report that they use secrecy to protect just over 50% of their innovations.

[4] In this paper, I assume that patents have infinite duration and analyze the effect of patent breadth on innovators' disclosure incentives. Innovators' disclosure incentives could also be analyzed in a model where patent length is the critical policy instrument.

and development, and to promote the disclosure of innovations. The analysis of optimal policy so far has mainly focused on the first goal while this paper aims to draw attention to the second goal as well. Both a flexible patent policy and a flexible antitrust policy, which allows collusion between winners of two subsequent races, can potentially play a role in encouraging disclosure through patenting. The results of this paper reveal that in cases when secrecy is not a viable option, it is optimal to have relatively narrow patent protection. Moreover, there may be cases when patent policy is completely ineffective in achieving the two goals of the patent system and when it is optimal to have a tolerant antitrust policy instead.

The model consists of two consecutive R&D races as in Denicolo (2000).[5] The firms race to achieve two symmetric and competing innovations, and the winner of the first race can also participate in the second R&D race. At the end of the first race, the winner decides whether to patent the innovation or keep it a secret. Under the patent law, an applicant is required to disclose sufficient knowledge about the innovation in order to enable someone skilled in the art to make and use all the embodiments of the innovation claimed in the patent. Therefore, the winner of the first race may be reluctant to patent in order not to reveal crucial technological information to the rival firms. After observing the patenting decision of the innovator, firms invest to develop an improved version in the second R&D race. If the first innovation is not disclosed, the rival firms must spend resources to gain information about it or to re-invent it, which reduces their competitiveness in the second R&D race.[6]

---

[5] Although the literature on cumulative innovations is quite extensive, it has generally been assumed that there is no racing in the development of the innovations and that each innovation is developed by a different firm. This may be a realistic assumption to make especially if the second-generation product is an application of the first innovation. However, in many cases in the real world, where the second-generation product is an improvement of the first innovation, it may not be realistic to assume that initial innovators do not invest to improve their products. One paper that allows for competition in the R&D stage and repeated innovation by the same firm is Denicolo (2000).

[6] If the innovator chooses to keep the innovation a secret, one issue that arises is what happens if the innovation is independently re-invented. In the United States, ownership is determined by the first-to-invent rule, which is what I assume in this paper. In most of the other countries, the first-to-file rule applies. Section 6 briefly discusses how a switch to the first-to-file rule may affect optimal policy.

The three domains of policy that influence the patenting decisions of innovators are patent policy, antitrust policy, and trade secret policy. In the context of sequentially developed competing innovations, the effective life of a patent is determined by how the courts interpret the patent and antitrust laws.[7] On the other hand, the government's trade secret policy determines the amount of protection innovators get against theft and unauthorized disclosure of their trade secrets. Following Chang (1995), I compare the behavioral implications and welfare properties of three different policy regimes. The courts first decide whether the innovation in question violates the patent of an earlier innovation. After deciding whether there is infringement, the courts can further decide whether to allow the firms to enter collusive agreements in case of no infringement. Such collusive agreements may increase the investment incentives of innovators by eliminating the competition between holders of competing but non-infringing patents. Focusing on the patenting decisions of innovators allows us to draw conclusions about the relationship between optimal patent and antitrust policy, and trade secret policy.

The results indicate that optimal patent protection depends on how viable an option secrecy is. Specifically, I find that if the government has a weak trade secret policy or if innovators cannot monitor the flow of their technological information, it may not be optimal to have strong patent protection. In other words, it is optimal to have relatively narrow

---

[7]The domain of patent policy cannot be isolated from the domain of antitrust policy since the ways in which patents are exploited by firms may give rise to antitrust concerns. For example, competing firms may use horizontal agreements regarding patents in order to restrict the prices they are going to charge in the product market. Or, they can give monopoly power to a single firm by selling all their competing patents to that firm. Historically, finding the right balance between the two areas of law has been an important policy question. The *Antitrust Guidelines for the Licensing of Intellectual Property,* issued jointly by the Department of Justice (DOJ) and the Federal Trade Commission (FTC), describes generally how the agencies will handle various forms of licensing deals that may contain anti-competitive provisions. See Tom (1998) and Pitofsky (2000) for recent comments on the purposes of intellectual property law and antitrust law. See also FTC Chairman Pitofsky's speech, "Antitrust and Intellectual Property: Unresolved Issues at the Heart of the New Economy," March 2, 2001, available at http://www.ftc.gov/speeches/pitofsky/ipf301.htm, and the former Deputy Assistant Attorney General Joseph Farrell's speech, "Thoughts on Antitrust and Innovation," January 25, 2001, available at http://www.usdoj.gov/atr/public/speeches/7402.htm.

patent protection, as in Denicolo (2000), if innovators cannot rely on secrecy to protect their innovations. This is because as the rival firm's probability of success in case of no patenting increases, innovating firms have increased incentives to patent their innovations. Therefore, the government may find it optimal to lower the degree of patent protection in order to encourage investment in the second R&D race. If, on the other hand, there is strong trade secret policy or innovators can monitor the flow of their technological information, then the rivals' probability of success in case of no patenting is relatively low and the government may find it optimal to have broad patent protection in order to encourage patenting.

Regarding antitrust policy, I find that having a flexible antitrust policy is justified only in cases when early innovators will always prefer to keep their innovations secret. If early innovators cannot be encouraged to disclose their innovations under any of the policy regimes, it is optimal have a flexible antitrust policy in order to encourage investment in subsequent R&D races. Non-disclosure of early innovations decreases the investment incentives of subsequent investors. Since second-generation innovators have the strongest incentives to invest when the courts find no infringement and allow collusion, it becomes optimal to allow collusive licensing agreements between competing innovators if there will be no patenting by early innovators under any of the policy regimes.

Thus, the paper makes a clear case for when it helps to have a lenient antitrust policy in order to protect intellectual property by taking innovators' decisions to patent into account. Chang (1995) argues that collusive licensing agreements should be allowed only under very limited circumstances, namely when the relative value of the second innovation is very high, while Priest (1977) and Kaplow (1984) oppose such collusive licensing on the grounds that such antitrust policy rewards innovation at an excessive social cost. This paper argues that there may be room for flexible antitrust policy if early innovators are likely to rely on secrecy rather than the patent system to protect their innovations.

6

Other papers that also focus on the issue of disclosure of early innovations are Matutes et al. (1996), Scotchmer and Green (1990), Gallini (1992), and Horstmann et al. (1985). Matutes et al. (1996), in contrast with this paper, consider disclosure issues in case of non-competing innovations. In their model, owners of fundamental innovations may be reluctant to commercialize their innovations in order to get a head start in developing applications of their new discoveries. They compare two types of patent protection regimes, "length" and "scope" protection, and conclude that policy-makers should target "scope" protection in order to eliminate the socially undesirable waiting period. Scotchmer and Green (1990) explore how stringent the novelty requirement in patent law, which affects the amount of technical information disclosed among firms, should be. This paper, in contrast, analyzes how broad patent protection should be and whether collusive agreements should be allowed.[8] Gallini (1992) analyzes patent policy in the context of costly imitation. In contrast with this paper, she allows imitation to take place both under patenting and no patenting. While I argue for narrower patent protection if the probability of imitation is high, she argues for broad patent protection in order to prevent imitation. Horstmann et al. (1985) investigate the patenting decisions of firms in case of a single R&D race while I consider the patenting decisions of firms in case of sequential R&D races.

The paper proceeds as follows. Section 2 presents the model and discusses the assumptions made regarding the legal background. In Section 3, I analyze the private incentives of the firms and explore what the optimal patent policy is. Section 4 concludes the paper. All proofs are in the Appendix.

---

[8] There are two legal instruments that determine the degree of protection that early innovators get. First, the novelty requirement determines how large subsequent innovations have to be in order to be patentable. Second, even if it is patentable, the courts may still decide that the second innovation infringes the patent of the first innovation. O'Donoghue et al. (1998) refer to the second instrument as a patent's leading breadth. This paper focuses on the second instrument while Scotchmer and Green (1990) focuses on the first one.

2 Model

2.1 Research environment and consumers

The research environment is adapted from Denicolo (2000). There are two sequential R&D races and free entry into each race. The size of each innovation is exogenously given and commonly known, but its timing is stochastic. For simplicity, I assume the two innovations are symmetric in terms of both their private and social values. The winner of the first race can participate in the second race, so the model allows for repeated innovation by the same firm. If the winner of the second race is different from the winner of the first race, the second innovator can obtain a patent for the technology and start using it provided that either the new technology does not infringe the patent of the old technology or the firms are able to make a licensing deal.[9]

The R&D game is modeled using a Poisson discovery process. Following Loury (1979), it is assumed that at the beginning of each race, participant i pays a lump sum amount equal to cx_i, where c is the per-unit R&D cost and x_i is the R&D effort level chosen by the firm. I assume the hazard function is linear. Both the Poisson and the linear hazard function assumptions are made in order to make the model tractable. The common social and private discount rate is represented by r.

Once an R&D race ends and the innovation is successfully developed, the good can be produced at zero marginal cost. There is a mass of consumers with homogeneous preferences. They buy at most one unit of the good.[10] I assume a model of successive market domination similar to Scotchmer and Green (1990). Each innovation has a market value of v. Firms

---

[9]Whether the second innovation should be patentable or not is a policy question that I do not consider in this paper. I assume the second innovation is always patentable. In cases when it is not, the bargaining powers of second-generation innovators may be significantly reduced. See Scotchmer and Green (1990), Scotchmer (1996), Denicolo (2000), and Denicolo and Zanchettin (2002) for optimal policy regarding the novelty requirement.

[10]Due to this assumption of inelastic demand, monopoly pricing in case of collusion does not result in any deadweight loss.

compete by setting prices. If the same firm is the owner of both of the innovations or if the firms sign a collusive licensing agreement, the profit is $2v$ per unit of time. If two different firms win the two races and the firms have to compete, the profits are given by 0 and v respectively. The consumers are willing to pay v for the basic version and $2v$ for the improved version of the product. Bertrand competition drives the price of the basic version to zero and that of the improved version to v minus some infinitesimal amount.

The sequence of events is as follows. At the beginning of the first R&D race, after the government announces its intellectual property protection policy, firms simultaneously decide how much to invest in order to develop the first innovation. The race ends when one of the participants successfully develops the innovation. The winner of the race decides whether to patent the innovation. In accordance with the goals of the patent system, I assume patenting means immediate full disclosure.[11] In other words, once an innovation is patented, the rival firms in the industry can easily build on it.[12,13] Therefore, if there is patenting, the hazard rates of the firms in the second R&D race are given by λx_w and λx_i, where x_w represents the investment of the winner of the first race in the second race. If there is no patenting, the innovator's hazard rate is λx_w while the rival firms' hazard rate is μx_i, where $\lambda \geq \mu$. If the innovation is kept a secret, rival firms can still learn about it through means such as employee mobility, informal communications between the employees of the firms, and espionage. Two factors determine whether innovators can protect themselves against such unauthorized disclosure of their innovations: the strength of the existing trade secret policy and the effectiveness of their own monitoring efforts. The magnitude of μ, which is the

---

[11] I assume that all patent applications are immediately disclosed to the public. In a more detailed analysis of the patenting process, it is necessary to differentiate between the time the innovator files a patent application, the time it becomes publicly available, and the time the patent is granted, if at all. Currently, in the United States, all patent applications are made available to the public 18 months after they are filed.

[12] That is, there is no learning by doing.

[13] I assume that even if an innovation is not patented, rival firms in the industry learn about its achievement with no time lag. According to Mansfield (1985), information about a new product or a process leaks out within about a year.

marginal benefit of the rival firms' R&D efforts in case of non-disclosure, depends on these two factors. Clearly, the first innovator's decision to patent is shaped by the magnitude of μ.

After the innovator decides whether to patent or not, the firms simultaneously choose how much to invest in the second R&D race. At the end of the second R&D race, the winner always patents since we assume the innovation is patentable and the winner does not lose anything by patenting. Finally, the firms compete in the product market by choosing prices.

2.2 Patent policy

Through its choice of policy, the government can affect the division of profits between different generations of innovators. I assume that the government can commit to a particular policy. The expected profits of firms depends on the specifics of that policy.

According to current patent law, if a second firm develops an improvement on a patented innovation and patents it, and if the courts find an infringement between the two sequential innovations, then the two firms end up holding "blocking patents." That is, if the courts find infringement (I), neither the first nor the second innovator can use the improved version without the other's permission. In such cases, the firms are allowed to maximize their joint profits.[14] This policy regime provides the first innovator with the broadest patent protection. If the courts do not find infringement (NI), they can either allow the firms to collude in the product market or they can ask them to compete. In other words, after deciding whether there is infringement, the courts can further decide whether to allow the firms to enter collusive agreements in case of no infringement. Such collusive agreements may increase the investment incentives of innovators by eliminating the competition between holders of competing but non-infringing patents. Thus, we may have either a policy of no infringement

---

[14]The purpose of patent protection is to give the patentholder a monopoly over the innovation covered by the patent. Legally, if the court decides that there is infringement, both the original innovation and the improvement are covered by the patent. Therefore, the courts maintain that the patentholder has the right to a monopoly over both of the products. I thank Howard Chang for making this point clear for me.

and collusion (NIC), or a policy of no infringement and no collusion ($NINC$). Clearly, the incentives to invest in the first R&D race are the highest under regime I while the incentives to invest in the second R&D race are the highest under regime NIC. If the courts rule that there is no infringement, the winner of the second race does not have to share its profits from the second innovation with the first innovator. If they furthermore allow collusive licensing agreements, the second innovator gets half of the surplus that the consumers would enjoy otherwise.

If the first innovator does not patent the innovation and continues to work on it, the first-to-invent rule in the United States allows the innovator to claim ownership in case of independent re-invention.[15] This implies that early innovators can strategically delay patenting of their innovations in an effort to get a head start in the second race without losing their right to patent. One important issue is whether the first innovator will use the innovation for commercial purposes if it chooses not to patent. If there is patenting, since commercialization could not result in any additional information leakage, the innovators would choose to offer the innovation for sale immediately. If there is no patenting, public disclosure of innovations significantly restricts the patenting rights of the original innovators. In the United States, innovators have a one-year grace period during which they can apply for a patent after they publicly disclose their innovations in any way. I assume that the length of the grace period is negligible compared to the expected duration of the R&D races. I also assume that reverse engineering is easy to achieve and that commercial use results in the same amount of information leakage as patenting.[16] This assumption implies that if an

---

[15] Under 35 U.S.C. § 102(g), a person is entitled to a patent unless "before such person's invention thereof, the invention was made in this country by another inventor who had not abandoned, suppressed, or concealed it." When an innovation should be regarded as abandoned, suppressed, or concealed is an issue that the courts have not been able to define with any level of precision. The prevailing view seems to be that if an innovator can prove that it is the first inventor and continues to work on the innovation even though it does not patent it, the innovation should not be considered abandoned, suppressed, or concealed.

[16] Section 4 contains a brief discussion of the implications of relaxing this assumption.

innovator does not patent an innovation but offers it for sale, the rival firms will have the same hazard rate as in the case of patenting. Therefore, the first innovator will always prefer patenting to not patenting if it is going to commercialize the product, and we will always have no commercialization under no patenting.

In the following analysis, we are interested in finding the subgame perfect Nash equilibria. Therefore, we start from the second R&D race and work backwards.

3 Analysis

3.1 Second race

3.1.1 Patenting

If the policy regime is I, the first innovator can make a positive profit even if it loses the second race. The second innovator has exclusive rights to the second innovation, but it cannot use it without making a licensing agreement with the first innovator. If the firms can reach an agreement, the second innovation can be sold at a price of $2v$, which is the maximum amount the consumers are willing to pay for the product. If the firms cannot reach an agreement, the first innovator will continue to sell the first innovation at a price of v. I assume the parties have equal bargaining power and divide the surplus to be shared, v, equally. Therefore, if the first innovator loses the second race, it can still earn $\frac{v}{2}$ in addition to what it can make on its own without reaching an agreement with the second innovator. If it wins the second race, it earns $2v$.

Let $\pi_w^{P/I}$ stand for the expected profits of the winner of the first race when it patents and the courts find an infringement. Given the above discussion, it is equal to

$$
\begin{aligned}
\pi_w^{P/I} &= \int_0^\infty e^{-(\lambda X_2 + r)t}\left(\lambda x_w\left(\frac{2v}{r}\right) + \lambda X_o\left(\frac{v}{r} + \frac{v}{2r}\right) + v\right)dt - cx_w \\
&= \frac{\lambda x_w\left(\frac{2v}{r}\right) + \lambda X_o\left(\frac{v}{r} + \frac{v}{2r}\right) + v}{\lambda X_2 + r} - cx_w
\end{aligned}
\tag{1}
$$

12

where $X_o = \sum_i x_i$ represents the aggregate investment of all other firms, which are assumed to be symmetric, and $X_2 = X_o + x_w$ represents the aggregate investment in the second race. To get the aggregate investment level, we can simply add the individual investment levels since the research efforts of different firms are assumed to be independent of each other. The innovator maximizes $\pi_w^{P/I}$ with respect to x_w. Setting the first order condition equal to zero gives us the following condition.[17]

$$\frac{\lambda v \left(\lambda X_o + 2r\right)}{2cr} = \left(\lambda X_2 + r\right)^2 .\tag{2}$$

If one of the rival firms wins the second race, it earns $\frac{v}{2}$. Its payoff function is

$$\pi_i^{P/I} = \int_0^\infty e^{-(\lambda X_2 + r)t} \left(\lambda x_i \frac{v}{2r}\right) dt - cx_i = \frac{\lambda x_i \left(\frac{v}{2r}\right)}{\lambda X_2 + r} - cx_i.$$

Under symmetry, each rival firm maximizes $\pi_i^{P/I}$ with respect to x_i. Since there is free entry into the R&D race, the equilibrium individual investment levels and number of firms can be determined by solving the first order condition of the innovator, the first order condition of a generic rival firm, and the zero profit condition simultaneously. Since we are interested in figuring out the aggregate investment of all rival firms, we can simply use the zero profit condition to determine the rival firms' aggregate best response to the first innovator's investment choice. As a result, combining the zero profit condition together with (2) gives us $x_w^{P/I} = r/\lambda$ and $X_o^{P/I} = \left(\lambda v - 4cr^2\right)/2cr\lambda$.

Under policy regime NIC, the courts do not find an infringement but still allow the two winners, which hold competing patents, to collude in the product market. If the firms do not collude, the price competition between the firms drives the price of the first innovation to zero and the price of the second innovation to v minus some infinitesimal amount. If the firms decide to collude, they can charge $2v$ instead of v for the improved version. This

---

[17]It is straightforward to check that the second order condition is met.

implies that if a rival firm wins the second race, the first innovator receives $\frac{v}{2}$ and the second innovator receives $v + \frac{v}{2}$. Thus, the second innovator receives half of the market value of the first innovation in addition to the market value of the second innovation. The first innovator, on the other hand, receives much less than what it receives under policy regime I if it loses the second R&D race. If the first innovator wins the second race, it earns $2v$.

The first innovator solves

$$\max_{x_w} \pi_w^{P/NIC} = \frac{\lambda x_w \left(\frac{2v}{r} \right) + \lambda X_o \left(\frac{v}{2r} \right) + v}{\lambda X_2 + r} - c x_w$$

and a generic rival firm i solves

$$\max_{x_i} \pi_i^{P/NIC} = \frac{\lambda x_i \left(\frac{v}{r} + \frac{v}{2r} \right)}{\lambda X_2 + r} - c x_i$$

respectively. Due to free entry, we can set $\pi_i^{P/NIC} = 0$ and solve for X_o. Substituting for X_o in the first innovator's first order condition, we find that the first order condition is always negative. Therefore, the first innovator finds it optimal not to invest in the second race. We have $x_w^{P/NIC} = 0$ and $X_o^{P/NIC} = \left(3\lambda v - 2cr^2 \right) / 2cr\lambda = X_2^{P/NIC}$. Compared to regime I, the rival firms have substantially increased incentives to invest since their bargaining power is higher. When the courts rule that there is no infringement, the winner of the second race does not have to share the profits from the improved version of the product with the first innovator. The first innovator, on the other hand, faces a higher chance of losing in the second R&D race. Since it will still earn $\frac{v}{2r}$ even if it loses, it prefers not to invest at all.

Finally, under policy regime $NINC$, the courts do not find an infringement and they do not allow the firms to collude in the product market. The first innovator does not earn anything if it loses the second race. The winner of the second race, if different from the winner of the first race, receives v, the full market value of the innovation. This is higher than what it earns under I, but lower than what it earns under NIC since the firms are

14

not allowed to collude. Solving the zero profit condition and the first innovator's first order condition simultaneously, we get $x_w^{P/NINC} = X_o^{P/NINC} = \left(\lambda v - cr^2\right)/2cr\lambda$.

Comparing the individual investment levels, we have $x_w^{P/NINC} > x_w^{P/I} > x_w^{P/NIC}$ and $X_o^{P/NIC} > X_o^{P/NINC} > X_o^{P/I}$. For the rival firms, the returns to investment are the highest under NIC and the lowest under I. Thus, they invest the most under NIC and the least under I. For the first innovator, the case is a little bit more complicated. While choosing its investment level, it also takes into account its expected profits from losing in the second race and invests more aggressively if its expected profits from losing are low. Since it gets zero under $NINC$ if it loses the second race, the incentives to win the race are the highest in that case.

3.1.2 No patenting

If the first innovator chooses not to patent, we know it is not going to use the innovation for commercial purposes since public disclosure of innovations restricts their patentability and we assume reverse engineering is easy. If commercial use reveals the innovation fully, the innovator would always prefer patenting to not patenting as long as it is going to offer the innovation for sale. Therefore, while calculating the returns from not patenting, we only need to focus on the case when there is no commercialization.

Under policy regime I, the first innovator maximizes

$$\pi_w^{NP/I} = \frac{\lambda x_w \left(\frac{2v}{r}\right) + \mu X_o \left(\frac{v}{r} + \frac{v}{2r}\right)}{\lambda x_w + \mu X_o + r} - cx_w.$$

This payoff function differs from (1) in two ways. First, if the first innovation is not patented, the rival firms' hazard rate decreases. Second, the first innovator cannot make any interim profits from the first innovation until the second innovation is developed.

15

Each rival firm has a payoff function of the form

$$\pi_i^{NP/I} = \frac{\mu x_i \left(\frac{v}{2r}\right)}{\lambda x_w + \mu X_o + r} - c x_i.$$

Solving for the investment levels, we get

$$x_w^{NP/I} = \frac{(\lambda - \mu)\,\mu v^2 + 2cr^2 \lambda\,(2v + 1)}{2cr\lambda^2 v}$$

and

$$X_o^{NP/I} = \frac{\mu^2 v - 8cr^2 \lambda}{2cr\lambda\mu}.$$

If the policy regime is NIC, the investment levels are

$$x_w^{NP/NIC} = \frac{9\,(\lambda - \mu)\,\mu v + 2cr^2 \lambda}{6cr\lambda^2}$$

and

$$X_o^{NP/NIC} = \frac{9\mu^2 v - 8cr^2 \lambda}{6cr\lambda\mu}.$$

Comparing these investment levels with the ones under patenting, we observe that under both I and NIC, not patenting results in an increase in the investment of the first innovator and a decrease in the investments of the rival firms. Since the marginal benefit of investment decreases for the rival firms, they choose to invest less. For the winner of the first race, not patenting decreases the level of competition it faces in the R&D market. Moreover, it decreases the payoff the first innovator will receive if it loses the second race. Thus, it has incentives to invest more in order to increase its chances of winning.

Finally, under regime $NINC$ we have

$$x_w^{NP/NINC} = \frac{(2\lambda - \mu)\,\mu v}{2cr\lambda^2}$$

and

$$X_o^{NP/NINC} = \frac{\mu^2 v - 2cr^2 \lambda}{2cr\lambda\mu}.$$

16

The rival firms still invest less than they would under patenting. For the winner of the first race we have $x_w^{P/NINC} > x_w^{NP/NINC}$ if and only if $(\lambda - \mu)^2 v - cr^2\lambda > 0$. The first innovator invests more under no patenting for μ close to λ and invests less under no patenting for low values of μ. This is because when the policy regime is $NINC$, the first innovator earns zero if it loses the second race whether or not it patents. Therefore, when μ is high, it invests more under no patenting than under patenting because under no patenting it is not making any interim profits and, thus, would like to win the second race as soon as possible. When μ is low, it does not have to invest as much since it does not face as intense competition.

Before moving onto the analysis of the patenting decision, we can note a few things about the aggregate investment levels in the second R&D race.

Proposition 1 $X_2^{P/NIC} > X_2^{P/NINC} > X_2^{P/I}$ and $X_2^{NP/NIC} > X_2^{NP/NINC} > X_2^{NP/I}$.

Under both patenting and no patenting, regime NIC results in the highest investment level in the second R&D race. This result is driven by the investment incentives of the rival firms since they have the highest incentives to invest under NIC. The incentives to invest in the second R&D race are the lowest under policy regime I since it favors first-generation investors at the expense of second-generation investors.

Remark 1 $\frac{\partial X_2^{NP}}{\partial \mu} > 0$.

This is an intuitive result that is again driven by the investment incentives of the rival firms. The rival firms' investment is an increasing function of μ: $\partial X_o^{NP}/\partial \mu > 0$. The first innovator's investment increases with μ for low levels of μ and decreases with μ for μ close to λ. At the aggregate level, the effect of μ on X_o^{NP} dominates the effect of μ on x_w^{NP}.

In the following analysis, I mainly focus on the case where the aggregate investment of the rival firms under both patenting and no patenting is always positive. Since their incentives

to invest are the lowest under regime I, this assumption implies that $\lambda v > 4cr^2$ in case of patenting and $\mu^2 v > 8cr^2\lambda$ in case of no patenting. Note that if μ is such that the first innovator can deter entry of the rival firms into the second R&D race by not patenting, it will have much lower incentives to patent.

3.2   Patenting decision

Under the first-to-invent rule, the innovator can delay its patenting decision without endangering its right to patent. In such a setting, the benefit of patenting is that the innovator can start to make profits from the innovation immediately. The benefit of delaying patenting is that it slows down the progress of rival firms. At the end of the first R&D race, the winner decides whether to patent by considering this trade-off. Clearly, μ, the rival firms' marginal benefit of investment in case of no patenting, plays a crucial role in the patenting decision.

Lemma 1 π_w^{NP} *is a decreasing function of μ.*

Under all of the policy regimes, the profits of the first innovator under no patenting decreases as μ increases. As μ increases, competition in the R&D market increases, which drives the profits of the first innovator down. The first innovator patents the first innovation if the rivals' probability of success in the second R&D race in case of no patenting is sufficiently high. In other words, if the first innovator can sufficiently impair the positions of the rival firms in the second R&D race by not patenting, it will choose not to patent. Let μ^I represent the critical μ value such that for $\mu > \mu^I$, the first innovator will choose to patent. We can define μ^{NIC} and μ^{NINC} in a similar fashion.

Proposition 2 *The first innovator's threshold for patenting is the lowest under regime I and the highest under regime $NINC$: $\mu^I < \mu^{NIC} < \mu^{NINC}$.*

Providing first-generation innovators with strong protection increases their incentives to patent. The regimes differ in the amount of licensing fees patentholders can collect if a rival firm develops an improved version of their innovation. If one of the rival firms wins the second R&D race, the licensing fees that the first innovator can collect are the highest under I and equal to zero under $NINC$. Not patenting decreases the investment incentives of the rival firms and relaxes the competition in the R&D industry. The first innovator is less concerned about the intensity of competition in the R&D market the higher the amount it will earn if it loses the second race. Its incentives to patent are the highest under policy regime I because it provides the innovator with the highest insurance against losing in the second R&D race.

Thus, we are likely to see patenting more frequently under regime I. There exists values of μ for which the first innovator chooses to patent under regime I while it chooses not to patent under the other two policy regimes.

3.3 First race

We can now analyze the investment incentives in the first R&D race. The prize to the winner in the first R&D race depends on the policy regime. If a firm loses the first race, it can still participate in the second R&D race. However, its expected profit in the second race is zero due to free entry.

Since all of the firms are symmetric in the first race, each firm has a payoff function of the form

$$\pi_i = \frac{\lambda x_i \pi_w}{\lambda X_1 + r} - c x_i$$

where X_1 represents the aggregate investment level in the first race and π_w represents the winner's expected profit at the beginning of the second race. Due to free entry, the aggregate investment levels can be determined by using the zero profit conditions. Comparing these investment levels, we get the following result.

Proposition 3 $X_1^{P/I} > X_1^{P/NINC} > X_1^{P/NIC}$ and $X_1^{NP/I} > X_1^{NP/NINC} > X_1^{NP/NIC}$.

Under both patenting and no patenting, we have the same ranking of the aggregate investment levels in the first race. Since the winner of the first race has the strongest patent protection under regime I, it is going to have the highest investment incentives then. Comparing the expected profit of the winner under regimes NIC and $NINC$, we observe that $\pi_w^{P/NINC} > \pi_w^{P/NIC}$ and $\pi_w^{NP/NINC} > \pi_w^{NP/NIC}$. Since the winner of the first race faces stronger competition under NIC, its expected profits are lower in that case. This causes the investment level in the first race under regime NIC to be relatively lower.

Comparing Proposition 1 and 3 reveals that allowing collusion favors second-generation investors at the expense of first-generation investors while having broad patent protection, policy regime I, favors first-generation investors at the expense of second-generation investors. We can also show that while investment in the second race increases with μ, investment in the first race decreases with μ.

Remark 2 $\frac{\partial X_1^{NP}}{\partial \mu} < 0$.

An increase in μ increases the competition in the second R&D race. Consequently, it affects the investment level in the first R&D race adversely.

3.4 Welfare Analysis

We assume the courts seek to maximize expected social welfare at the beginning of the first R&D race. Social welfare is defined as the sum of producer surplus, consumer surplus, and the non-appropriable values of the innovations. As pointed out in Arrow (1962), for a variety of reasons investors may not always be able to appropriate for themselves the entire social benefit of their innovations. Let $s \geq 0$ stand for the non-appropriable value of the innovations. It represents the increase in social welfare that other firms may enjoy due to either knowledge

20

spillovers or demand spillovers.[18] Note that since we assume the innovators in our model are able to capture the entire consumer surplus, s does not include any consumer surplus.

While writing the expected social welfare function, we should take into account the spillovers between the first and the second innovations. As pointed out in Chang (1995) and Green and Scotchmer (1995), the social benefit from an innovation should include the option value of the subsequent innovations it inspires. The second innovation would not have been possible without the first one. Therefore, if the first innovation is patented, the expected social welfare can be expressed as

$$W = P\left(X_1\right)\left[\left(\frac{v+s}{r}\right) + \left(P\left(X_2\right)\left(\frac{v+s}{r}\right) - cX_2\right)\right] - cX_1$$

where $P\left(X_1\right) = \lambda X_1/\left(\lambda X_1 + r\right)$ and $P\left(X_2\right) = \left(\lambda X_2\right)/\left(\lambda X_2 + r\right)$. As in Denicolo (2000), these expressions can be interpreted as the probability of innovating adjusted by the discount rate.

Due to free entry, expected profits are bid down to zero in both races. Social welfare becomes a function of the non-appropriable values of the innovations and consumer surplus. We have

$$\begin{aligned}
W^{P/I} &= P\left(X_1^{P/I}\right)\left[\frac{s}{r} + P\left(X_2^{P/I}\right)\frac{s}{r}\right] \\
W^{P/NIC} &= P\left(X_1^{P/NIC}\right)\left[\frac{s}{r} + P\left(X_2^{P/NIC}\right)\frac{s}{r}\right] \\
W^{P/NINC} &= P\left(X_1^{P/NINC}\right)\left[\frac{s}{r} + P\left(X_2^{P/NINC}\right)\frac{s}{r} + \left(\frac{\lambda X_o^{P/NINC}}{\lambda X_2^{P/NINC} + r}\right)\frac{v}{r}\right].
\end{aligned}$$
(3)

The extra term in $W^{P/NINC}$ represents the consumer surplus. If the firms are not allowed to collude, there is competition between the first and second innovators. Consumers receive a surplus of v per period after the second innovation is developed. Consumer surplus exists

---

[18]In fact, Jones and Williams (1998) find that the optimal R&D investment is at least two to four times of actual investment.

if and only if the second innovation is developed by one of the rival firms. This is the reason we have $\lambda X_o^{P/NINC}$ instead of $\lambda X_2^{P/NINC}$ in the numerator. Under policy regimes I and NIC, the investors can capture all of the consumer surplus, which increases their willingness to invest. Note that in this model, charging a price above marginal cost does not result in any deadweight loss.

If the first innovation is not patented, we have

$$
\begin{aligned}
W^{NP/I} &= P\left(X_1^{NP/I}\right) P\left(X_2^{NP/I}\right) \frac{2s}{r} \\
W^{NP/NIC} &= P\left(X_1^{NP/NIC}\right) P\left(X_2^{NP/NIC}\right) \frac{2s}{r} \\
W^{NP/NINC} &= P\left(X_1^{NP/NINC}\right) \\
&\quad \left[P\left(X_2^{NP/NINC}\right) \frac{2s}{r} + \left(\frac{\mu X_o^{NP/NINC}}{\lambda x_w^{NP/NINC} + \mu X_o^{NP/NINC} + r} \right) \frac{v}{r} \right].
\end{aligned}
\tag{4}
$$

Under no patenting, since the first innovator does not start to use the innovation for commercial purposes until the second innovation is developed, the social benefit of the first innovation as well as the second innovation can only be realized after the second innovation is developed. This is the main difference between the welfare functions under patenting and no patenting.

We first analyze whether the first innovator's patenting decision is socially optimal. Since in R&D race models the private incentives to invest may exceed the social incentives to invest, disclosure of innovations, which increases the investment incentives of subsequent innovators, may not always be desirable. However, Proposition 4 states that there is never over-patenting.

Proposition 4 *Whenever the first innovator decides to patent, it makes the socially optimal choice.*

In order to analyze the optimal policy, consider first the case when μ is low. If the rival firms' probability of success in case of no patenting is sufficiently low, there will not be any patenting under any of the policy regimes. In that case, we need to compare $W^{NP/I}$,

$W^{NP/NIC}$, and $W^{NP/NINC}$. Let s^{low} stand for the value of s such that $W^{NP/NIC} = W^{NP/NINC}$.

Proposition 5 *For $\mu^I \geq \mu$, $W^{NP/NIC} > W^{NP/I}$. Moreover, $W^{NP/NIC} \lesseqqgtr W^{NP/NINC}$ as $s \lesseqqgtr s^{low}$.*

Since, for $\mu^I \geq \mu$, social welfare is always higher under NIC than it is under I, it is never optimal to have regime I. This is because regime I gives the first innovator the strongest patent protection and it is not necessary to have such strong protection when the first innovator already protects itself by not patenting. Second-generation investors, on the other hand, have substantially low incentives to invest under non-disclosure. Among the three policy regimes, the incentives to invest in the second R&D race are the strongest under NIC. Since the incentives to invest in the second R&D race are low when μ is low, having non-infringement and allowing collusion may be optimal in order to encourage investment in the second R&D race.[19] We will see that this is the only case when it is optimal to allow collusion.

Proposition 5 also states that allowing collusive licensing becomes optimal when the non-appropriable value of the innovations is sufficiently high. Since the aggregate investment levels in both of the R&D races are higher under NIC than they are under $NINC$, as the non-appropriable value of the innovations increases, a switch to NIC becomes desirable in order to encourage investment in both races. This is the case despite the fact that when the innovators are allowed to collude, consumer surplus is reduced.

For $\mu^I < \mu$, there will always be patenting under I.

---

[19]If μ is so low that the innovator can deter entry into the second R&D race by not patenting, there may be more scope for allowing collusive licensing agreements because the rivals firms would have more incentives to enter the second race under regime NIC than they would under the other two regimes.

Proposition 6 *If there is going to be patenting under policy regime I, it is never optimal to allow collusion.*

Although having strong patent protection discourages investment in the second R&D race, social welfare is always higher under I than it is under NIC as long as there will be patenting under regime I.[20] Under both of the regimes, the innovating firms capture all of the consumer surplus. Policy regime I favors first-generation innovators while policy regime NIC favors second-generation innovators. Proposition 6 states that in case of symmetric innovations, having broad patent protection is socially preferable to having narrow patent protection (no infringement) and collusion. Although having infringement provides first-generation innovators with strong incentives to invest at the expense of second-generation innovators, having no infringement with collusion may provide second-generation innovators with unnecessarily high incentives to invest because under NIC second-generation innovators obtain a share of the first innovator's profits.

The next proposition states that as the non-appropriable value of the innovations increases, it becomes socially more beneficial to have policy regime I, which transfers rents from consumers to producers, than to have policy regime $NINC$. Let s^{int} and s^{high} stand for the critical s values for intermediate and high values of μ respectively. We define them by setting $W^{P/I} = W^{NP/NINC}$ in each case.

Proposition 7 *For $\mu^{NINC} > \mu > \mu^I$, $W^{P/I} \lesseqgtr W^{NP/NINC}$ as $s \lesseqgtr s^{int}$. For $\mu \geq \mu^{NINC}$, $W^{P/I} \lesseqgtr W^{P/NINC}$ as $s \lesseqgtr s^{high}$.*

When the first innovator participates in the second R&D race and there is going to be patenting under regime I, it becomes optimal to have regime I as the non-appropriable value

---

[20]It is important to keep in mind that we have been assuming that the innovations are symmetric. Having a policy of NIC may become more attractive if it is the case that the second innovation is a much more significant one than the first one.

of the innovations increases. Policy regime I provides the participants in the first R&D race with the highest incentives to invest. As the non-appropriable value of the innovations increases, it becomes optimal to stimulate investment in the first race. If the non-appropriable value of the innovations is low, it is optimal to have $NINC$, which stimulates investment in the second R&D race and allows consumers to benefit from the competition between sequential innovators.

We finally analyze how optimal policy changes as μ increases. The size of μ depends on the government's trade secret policy and the firms' ability to monitor the flow of their technological information. Clearly, different industries may be characterized by different values of μ. That is, for the innovators, the effectiveness of secrecy as a protection mechanism may vary across industries. Proposition 8 implies that having strong patent protection becomes less desirable as μ increases.

Proposition 8 *For all values of μ, $s^{high} > s^{int}$.*

If the probability of imitation under secrecy is high, the innovator has increased incentives to patent the innovation in order to protect itself. In fact, when μ is sufficiently high, there will be patenting under all of the policy regimes. The first innovator will always choose to patent the innovation since it cannot significantly decrease the success rate of the rival firms by not patenting. Therefore, it is not necessary to have strong patent protection in order to encourage the disclosure of innovations and regime I becomes less attractive from the social point of view. On the other hand, if the probability of imitation under secrecy is low, it becomes more attractive to have strong patent protection in order to encourage disclosure.

These results indicate that it may be optimal to have strong patent protection even in cases when several firms race to develop certain innovations if we endogenize the patenting decisions of innovators. Narrower protection may be optimal, as it is Denicolo (2000), under

a weak trade secret policy. However, in cases when secrecy is an attractive option for the innovators, it is optimal to have broader patent protection. As far as collusive licensing agreements are concerned, the results extend the ones in Chang (1995). Chang argues that collusion should be allowed if the relative private value of the second innovation is significantly higher. In our case, collusion should be allowed if early innovators will prefer non-disclosure.

4 Conclusion

Innovators can try to protect their innovations by relying on either the patent system or secrecy. Although the patent system provides innovators with the right to exclude others from using the innovation for a fixed period of time, it also requires them to disclose the innovation. This paper emphasizes that while designing patent and antitrust policy, it should be taken into consideration that early innovators may not always choose to patent. Attention should be paid to how profits are divided between different generations of innovators as well as to how to discourage early innovators from restricting the supply of new technologies to other, possibly rival, firms.[21]

Three policy regimes have been compared in order to explore when courts should find an infringement and when they should allow collusive licensing agreements in case of competing innovations. I have assumed that if an innovation is not patented, the first-to-invent rule applies. Under the first-to-invent rule, innovators do not lose their ownership rights if they do not patent an innovation immediately. The patenting decisions of innovators depend on the ease with which rival firms can learn about new technologies. This indicates that policy

---

[21]The trade-off that policy makers face between decreasing the incentives for initial innovations and increasing the incentives for subsequent innovations while determining the optimal patent protection implies that we may be facing a time-inconsistency problem as far as policy making is concerned. In terms of emerging industries, while the policy makers may find it optimal to pursue a strong patent protection policy in the initial stages of the industries, their optimal policy may change as the industry ages. Unless they can credibly commit to a single policy, the expectation that patents may not be as effectively enforced in the future may cause the R&D incentives of initial innovators to diminish.

geared towards industries where innovators can rely on secrecy to protect their innovations may be different from policy geared towards industries where innovators cannot rely on secrecy.

The main findings of the paper can be summarized as follows. Suppose we allow for repeated innovation by the same firm. As far as disclosure is concerned, although second–generation investors would have decreased incentives to invest under strong patent protection, providing first-generation innovators with strong protection increases their incentives to both invest and patent. They have the highest incentive to patent under a policy of strong patent protection, and the lowest incentive to patent under a policy of no infringement and no collusion. Patenting allows them to make profits from the first innovation while they race to develop the second innovation. However, it has the disadvantage of transferring crucial technical information to rival firms. Since strong patent protection provides them with the highest amount of insurance against losing in the second race, they have the highest incentives to patent then.

Optimal patent policy depends on how viable an option secrecy is. If innovators cannot rely on secrecy to protect their innovations, it is optimal to have relatively narrow patent protection. This happens if the government has a weak trade secret policy or if innovators cannot monitor the flow of their technological information. If secrecy is a feasible option, it becomes optimal to have relatively broad patent protection. This is because when innovators cannot rely on secrecy to protect their innovations, they have increased incentives to patent them and it is not necessary for the government to give them extra incentives to patent. In the case when innovators always prefer secrecy over patenting, it becomes optimal to have a flexible antitrust policy rather than a flexible patent policy. Since non-disclosure reduces the investment incentives in the second R&D race, allowing collusive licensing agreements between competing innovators becomes optimal in order to stimulate investment in the second

R&D race. This implies that in such cases, policy should target encouraging investment in future races rather than encouraging disclosure of past innovations.

The analysis can be extended in several ways by relaxing some of the assumptions we have made. First, we have assumed that the first-to-invent rule determines ownership. Currently, there is a lot of debate on whether the United States should continue with its first-to-invent policy or adapt the first-to-file policy to be in harmony with the rest of the world. Therefore, we can analyze what a switch to the first-to-file rule would imply in terms of optimal patent policy. The first-to-invent rule gives innovators more flexibility about when to disclose their innovations. Under the first-to-file rule, innovators may feel more threatened and have higher incentives to patent. Therefore, having strong patent protection may become less attractive. Second, we have assumed that reverse engineering is very easy. Therefore, the first innovator never uses the innovation for commercial purposes in the case when it does not patent. If reverse engineering is difficult, this would allow the first innovator to have a third option, which is to keep the innovation secret and commercialize it at the same time. In such cases, the innovator would have lower incentives to patent and the courts may find it optimal to have strong patent protection more frequently. Third, we have assumed that the second innovation is always patentable. Whether the second innovation should be patentable is a policy question that can be explored. Especially if the first innovation is not patented, this may make the second innovation more patentable. The government may decide whether to make the second innovation patentable or not depending on the patenting decision of the first innovator.

References

[1] Chang, H. F. 1995. "Patent Scope, Antitrust Policy, and Cumulative Innovation," *RAND Journal of Economics*, 26, 34-57.

[2] Cohen, W. M., R. R. Nelson, and J. P. Walsh. 2000. "Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)," Working Paper No. W7552, National Bureau of Economic Research, 2000.

[3] Denicolo, V. 2000. "Two-Stage Patent Races and Patent Policy," *RAND Journal of Economics*, 31(3), 488-501.

[4] Denicolo, V. and P. Zanchettin. 2002. "How Should Forward Patent Protection be Provided?" *International Journal of Industrial Organization*, 20, 801-27.

[5] Gallini, N. T. 1992. "Patent Policy and Costly Imitation," *Rand Journal of Economics*, 23(1), 52-63.

[6] Gallini, N. T. and S. Scotchmer. 2001. "Intellectual Property: When is it the Best Incentive System?" in *Innovation Policy and the Economy*, Vol. 2, Adam Jaffe, Joshua Lerner, and Scott Stern eds., MIT Press.

[7] Gallini, N. T. 2002. "The Economics of Patents: Lessons from Recent U.S. Patent Reform," *Journal of Economic Perspectives*, 16(2), 131-154.

[8] Green, J., and S. Scotchmer. 1995. "On the Division of Profit in Sequential Innovation," *RAND Journal of Economics*, 26, 20-23.

[9] Hall, B. H. and R. H. Ziedonis. 2001. "The Patent Paradox Revisited: An Empirical Study of Patenting in the U.S. Semiconductor Industry," *Rand Journal of Economics*, 32(1), 101-128.

[10] Horstmann, I., G. M. MacDonald, and A. Slivinski. 1985. "Patents as Information Transfer Mechanisms: To Patent or (Maybe) Not to Patent," *Journal of Political Economy*, 93(5), 837-858.

[11] Jones, C. I. and J. C. Williams. 1998. "Measuring the Social Return to R&D," *Quarterly Journal of Economics*, 113, 1119-1135.

[12] Kaplow, L. 1984. "The Patent-Antitrust Intersection: A Reappraisal," *Harvard Law Review*, 97, 1813-1892.

[13] Lanjouw, J. O. 1998. "Patent Protection in the Shadow of Infringement: Simulation Estimations of Patent Value," *Review of Economic Studies*, 65(4), 671-710.

[14] Lanjouw, J. O. and M. Schankerman. 2001. "Characteristics of Patent Litigation: A Window on Competition," *Rand Journal of Economics*, 32(1), 129-151.

[15] Lerner, J. 1994. "The Importance of Trade Secrecy: Evidence from Civil Litigation," Harvard Business School Working Paper #95-043.

[16] Levin, R. C., A. K. Klevorick, R. R. Nelson, and S. G. Winter. 1987. "Appropriating the Returns from Industrial Research and Development," *Brookings Papers on Economic Activity*, 3, 783-820.

[17] Mansfield, E. 1985. "How Rapidly Does New Technology Leak Out," *The Journal of Industrial Economics*, 34, 217-223.

[18] Matutes, C., P. Regibeau, and K. E. Rockett. 1996. "Optimal Patent Protection and the Diffusion of Innovation," *Rand Journal of Economics*, 27, 60-83.

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

[20] Pitofsky, R. 2000. "Challenges of the New Economy: Issues at the Intersection of Antitrust and Intellectual Property," *Antitrust Law Journal*, 68, 913-924.

[21] Priest, G. L. 1977. "Cartels and Patent License Arrangements," *Journal of Law and Economics*, 20, 309-377.

[22] Scotchmer, S. and J. Green. 1990. "Novelty and Disclosure in Patent Law," *Rand Journal of Economics*, 21, 131-46.

[23] Scotchmer, S. 1991. "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law," *Journal of Economic Perspectives*, 5, 29-41.

[24] Scotchmer, S. 1996. "Protecting Early Innovators: Should Second-Generation Products be Patentable?" *Rand Journal of Economics*, 27, 322-31.

[25] Schankerman, M. 1998. "How Valuable is Patent Protection? Estimates by Technology Field," *Rand Journal of Economics*, 29(1), 77-107.

[26] Tom, W. K. 1998. "Licensing and Antitrust: Common Goals and Uncommon Problems," Address before the American Conference Institute, *9th National Conference on Licensing Intellectual Property*, http://www.ftc.gov/speeches/other/aciippub.htm.

[27] U.S. Department of Justice and Federal Trade Commission. 1995. "Antitrust Guidelines for the Licensing of Intellectual Property," http://www.usdoj.gov/atr/public/guidelines/ipguide/htm.

Appendix

A Proof of Proposition 1

In case of patenting, we have

$$X_2^{P/NIC} - X_2^{P/NINC} = \frac{v}{2cr}$$

and

$$X_2^{P/NINC} - X_2^{P/I} = \frac{v}{2cr}.$$

Clearly, $X_2^{P/NIC} > X_2^{P/NINC} > X_2^{P/I}$.

In case of no patenting, we have

$$X_2^{NP/NIC} - X_2^{NP/NINC} = \frac{3\left(3\lambda - 2\mu\right)\lambda\mu^2 v - 2\left(\lambda - \mu\right)cr^2\lambda^2}{6cr\lambda^3\mu}$$

and

$$X_2^{NP/NINC} - X_2^{NP/I} = \frac{\mu^2 v - 6\left(\lambda - \mu\right)cr^2}{2cr\lambda\mu}.$$

Given the assumptions that aggregate investment by the rival firms in the second R&D race is positive and that $\lambda \geq \mu$, we get $X_2^{NP/NIC} > X_2^{NP/NINC} > X_2^{NP/I}$.

B Proof of Lemma 1

Differentiating we get:

$$\frac{\partial \pi_w^{NP/I}}{\partial \mu} = \frac{-\left(\lambda - \mu\right)v}{r\lambda^2}$$

$$\frac{\partial \pi_w^{NP/NIC}}{\partial \mu} = \frac{-3\left(\lambda - \mu\right)v}{r\lambda^2}$$

$$\frac{\partial \pi_w^{NP/NINC}}{\partial \mu} = \frac{-\left(2\lambda - \mu\right)v}{r\lambda^2}.$$

Since $\lambda \geq \mu$ by assumption, the derivative is negative in all three cases.

C Proof of Proposition 2

Setting $\pi_w^P - \pi_w^{NP} = 0$ under the three policy regimes, we find:

$$
\begin{aligned}
\mu^I &= \lambda - \frac{2r\sqrt{c\lambda v}}{v} \\
\mu^{NIC} &= \lambda - \frac{2r\sqrt{c\lambda v}}{3v} \\
\mu^{NINC} &= 2\lambda - \frac{\sqrt{\lambda v \left(cr^2 + \lambda v\right)}}{v}.
\end{aligned}
$$

Clearly, $\mu^{NIC} > \mu^I$. Moreover,

$$
\mu^{NINC} - \mu^{NIC} = \frac{\sqrt{\lambda}\left(3\sqrt{\lambda v} + 2\sqrt{cr^2} - 3\sqrt{\lambda v + cr^2}\right)}{3\sqrt{v}},
$$

which is positive if $3\sqrt{\lambda v} + 2\sqrt{cr^2} > 3\sqrt{\lambda v + cr^2}$. Squaring both sides we get:

$$
9\lambda v + 4cr^2 + 12\sqrt{\lambda v cr^2} \overset{?}{>} 9\lambda v + 9cr^2.
$$

Since $X_0^{P/I} > 0$ implies $\lambda v > 4cr^2$, this inequality always holds. Therefore, $\mu^{NINC} > \mu^{NIC}$. From Lemma 1 it follows that for all values of μ greater than these critical μ values, there will be patenting under the respective policy regimes.

D Proof of Proposition 3

In case of patenting, we have

$$
X_1^{P/I} - X_1^{P/NINC} = \frac{2\lambda v - 3cr^2}{2cr\lambda v}
$$

and

$$
X_1^{P/NINC} - X_1^{P/NIC} = \frac{r}{6\lambda}.
$$

Given the assumption that aggregate investment by the rival firms in the second R&D race is positive, we get $X_1^{P/I} > X_1^{P/NINC} > X_1^{P/NIC}$.

In case of no patenting, we have

$$X_1^{NP/I} - X_1^{NP/NINC} = \frac{\mu v - 3cr^2}{cr\lambda}$$

and

$$X_1^{NP/NINC} - X_1^{NP/NIC} = \frac{(\lambda - \mu)\mu v}{cr\lambda^2} + \frac{r}{3\lambda}.$$

Again, given the assumptions that aggregate investment by the rival firms in the second R&D race is positive and that $\lambda \geq \mu$, we get $X_1^{NP/I} > X_1^{NP/NINC} > X_1^{NP/NIC}$.

E   Proof of Proposition 4

It is necessary to show that $\pi_w^P > \pi_w^{NP}$ implies $W^P > W^{NP}$ under all three regimes. From (3) and (4), it is clear that $W^{P/I} > W^{NP/I}$ and $W^{P/NIC} > W^{NP/NIC}$ if $P\left(X_1^P\right) > P\left(X_1^{NP}\right)$ and $P\left(X_2^P\right) > P\left(X_2^{NP}\right)$ under the respective regimes. Both $P\left(X_1^P\right) > P\left(X_1^{NP}\right)$ and $P\left(X_2^P\right) > P\left(X_2^{NP}\right)$ hold under the condition for $\pi_w^P > \pi_w^{NP}$.

For regime $NINC$, $W^{P/NINC} > W^{NP/NINC}$ if $P\left(X_1^{P/NINC}\right) > P\left(X_1^{NP/NINC}\right)$, $P\left(X_2^{P/NINC}\right) > P\left(X_2^{NP/NINC}\right)$, and $\left(\frac{\lambda X_o^{P/NINC}}{\lambda X_2^{P/NINC} + r}\right) > \left(\frac{\mu X_o^{NP/NINC}}{\lambda x_w^{NP/NINC} + \mu X_o^{NP/NINC} + r}\right)$. Again, all three inequalities hold under the condition for $\pi_w^{P/NINC} > \pi_w^{NP/NINC}$.

F   Proof of Proposition 5

We have:

$$W^{NP/NIC} - W^{NP/I} = \frac{2s}{r}\left[P\left(X_1^{NP/NIC}\right)P\left(X_2^{NP/NIC}\right) - P\left(X_1^{NP/I}\right)P\left(X_2^{NP/I}\right)\right]$$

which is positive assuming $X_o^{NP/I} > 0$.

From the expressions for $W^{NP/NINC}$ and $W^{NP/NIC}$ it is clear that when $s = 0$, $W^{NP/NINC} >$

$W^{NP/NIC} = 0$. Consider the expression for $\frac{W^{NP/NINC}}{W^{NP/NIC}}$.

$$\frac{W^{NP/NINC}}{W^{NP/NIC}} = \frac{P\left(X_1^{NP/NINC}\right)P\left(X_2^{NP/NINC}\right)}{P\left(X_1^{NP/NIC}\right)P\left(X_2^{NP/NIC}\right)} +$$

$$\frac{P\left(X_1^{NP/NINC}\right)\left(\frac{\mu X_o^{NP/NINC}}{\lambda x_w^{NP/NINC}+\mu X_o^{NP/NINC}+r}\right)v}{P\left(X_1^{NP/NIC}\right)P\left(X_2^{NP/NIC}\right)2s}.$$

Taking the derivative with respect to s, we have $\frac{\partial\left(W^{NP/NINC}/W^{NP/NIC}\right)}{\partial s} < 0$. As $s \to \infty$, $\frac{W^{NP/NINC}}{W^{NP/NIC}} \to \frac{P\left(X_1^{NP/NINC}\right)P\left(X_2^{NP/NINC}\right)}{P\left(X_1^{NP/NIC}\right)P\left(X_2^{NP/NIC}\right)}$. This term is < 1 for $\mu^I \geq \mu$, so we have $W^{NP/NINC} < W^{NP/NIC}$.

G   Proof of Proposition 6

We need to show that $W^{P/I} > W^{NP/NIC}$ for $\mu^{NIC} > \mu > \mu^I$ and $W^{P/I} > W^{P/NIC}$ for $\mu > \mu^{NIC}$. For the latter case we have:

$$W^{P/I} - W^{P/NIC} = \frac{4crs\left(3\lambda v - 4cr^2\right)\left(9\lambda^2 v^2 - 13\lambda vcr^2 + 2c^2 r^4\right)}{\lambda v\left(9\lambda v - 2cr^2\right)\left(3\lambda v - 2cr^2\right)\left(3\lambda v + 2cr^2\right)}$$

which is positive assuming $X_o^{P/I} > 0$. For $\mu^{NIC} > \mu > \mu^I$, we use numerical analysis to show $W^{P/I} > W^{NP/NIC}$.

H   Proof of Proposition 7

The proof is similar to the proof of Proposition 5 and is not repeated here.

I   Proof of Proposition 8

Note that s^{high} is independent of μ since it is the critical s value for $\mu \geq \mu^{NNNC}$. We have:

$$s^{high} = \frac{P\left(X_1^{P/NINC}\right)\left(\frac{\lambda X_o^{P/NINC}}{\lambda X_2^{P/NINC}+r}\right)v}{P\left(X_1^{P/I}\right)\left[1 + P\left(X_2^{P/I}\right)\right] - P\left(X_1^{P/NINC}\right)\left[1 + P\left(X_2^{P/NINC}\right)\right]}$$

and

$$s^{int} = \frac{P\left(X_1^{NP/NINC}\right)\left(\frac{\mu X_o^{NP/NINC}}{\lambda x_w^{NP/NINC} + \mu X_o^{NP/NINC} + r}\right) v}{P\left(X_1^{P/I}\right)\left[1 + P\left(X_2^{P/I}\right)\right] - 2P\left(X_1^{NP/NINC}\right)P\left(X_2^{NP/NINC}\right)}.$$

It suffices to show that the numerator of s^{high} is greater than the numerator of s^{int}, and $P\left(X_1^{P/NINC}\right)\left[1 + P\left(X_2^{P/NINC}\right)\right] > 2P\left(X_1^{NP/NINC}\right)P\left(X_2^{NP/NINC}\right)$. Using the fact that $X_o^{NP/NINC} > 0$ and $\mu > \mu^I$, we can show that both inequalities hold.