DISCUSSION PAPERS IN ECONOMICS

Working Paper No. 09-07

Patenting in the Shadow of Independent Discoveries by Rivals

Tianle Zhang

University of Colorado

updated March 2010

November 2009

Department of Economics

University of Colorado at Boulder
Boulder, Colorado 80309

Patenting in the Shadow of Independent Discoveries by Rivals*

Tianle Zhang
University of Colorado at Boulder
February 2010

Abstract. This paper studies the decision of whether to patent in a dynamic model where firms innovate stochastically and independently. In the model, a firm can choose between patenting and maintaining secrecy to protect a successful innovation. Patenting grants probabilistic protection while secrecy is effective until rivals innovate. We show that (1) firms that innovate early are more inclined to choose secrecy whereas firms that innovate late have a stronger tendency to patent; (2) the incentives to patent increase with the innovation arrival rate; and (3) an increase in the number of firms may cause patenting to occur earlier or later, depending on the strength of patent protection. The socially optimal level of patent protection balances the trade-off between the provision of patenting incentive and the avoidance of unnecessary monopoly. We find that the socially optimal level of patent protection should be lower if the innovation arrival rate is higher or the number of firms is larger.

Key words: Patenting decisions; Patents; Secrecy; Independent discoveries.

JEL Classification: O31, O34

*I thank Yongmin Chen for his advice and encouragement throughout the course of this research. I also thank Martin Byford, Yuk-fai Fong, Byung-Cheol Kim, Jennifer Lamping, Keith Maskus, Anna Rubinchik, Scott Savage, Hale Utar, and participants of 7th International Industrial Organization Conference, European Economic Review Talented Economists Clinic 2 and a seminar at the Hong Kong University of Science & Technology for their helpful comments. All errors are my own.

1 Introduction

An important strategic decision for a firm is how to protect innovations. The firm can apply for patent protection or keep an innovation in secret use. Evidences show that firms often make heterogeneous choices on whether to patent their innovations. In fact, only a small proportion of innovations are patented (Scherer, 1965; Pakes and Griliches, 1980; Mansfield, 1986). Moreover, secrecy is viewed as an increasingly important strategy for appropriating innovations (Levin et al., 1987; Cohen et al., 2000). One question that naturally arises is why some firms choose patents while others adopt secrecy to protect innovations. Moreover, given firms' strategies on whether to patent, what is the socially optimal level of patent protection?

This paper attempts to address these questions. Our analysis is motivated by several observed features concerning innovations and patenting. First, in many situations, multiple firms are capable of independently coming up with identical or similar innovations. As discussed in Varian et al. (2005) and Shapiro (2007), this can happen because innovation firms often share common knowledge bases or find research paths restricted by universal standards. Second, patent protection is probabilistic. Many patent applications are not approved,[1] and as emphasized in Choi (1998) and Lemley and Shapiro (2005), even issued patents can be ruled invalid through litigation.[2] Because of the requirement for full disclosure of innovation information during patenting process, the revealed information, under imperfect patent protection, may be utilized to the benefit of rival firms. Third, a firm that keeps an innovation secret runs the risk of allowing another firm to obtain a patent for the innovation. Under current U.S. patent laws, a later inventor is permitted to obtain a patent for an invention abandoned, suppressed or concealed by previous inventors (e.g. Merges and Duffy, 2007). In addition, U.S. patent laws do not grant prior user rights[3], which means a later inventor has the right to exclude previous inventors that rely on secrecy.[4]

To capture these features, we develop a dynamic model of innovation where multiple

[1] Out of 485,312 patent applications received, only 185,224 (less than 40%) are granted patents in the year of 2008. Data source: U.S. Patent Statistics Chart Calendar Years 1963 - 2008. http://www.uspto.gov/go/taf/us_stat.htm.

[2] Allison and Lemley (1998) report that out of 300 cases of final validity decisions in the data set, the patents were found invalid in 138 cases.

[3] With the exceptions in business methods.

[4] As discussed in Denicolò and Franzoni (2004a), in *Gore v. Garlock (721 F.2d 1540, 1983)*, Garlock Inc. had discovered a process to create a tape of unsintered polytetrafluorethylene filament but decided to keep the process in secret use. However, the process was later rediscovered by W.L. Gore & Associates, Inc. who succeeded in patenting the process. In another case discussed in Marshall (1991), New England Biolabs and Bethesda Research Labs had produced the modified T7 DNA polymerase and offered it for sale but neither of the companies applied for a patent. The patent later was granted to Harvard researchers who threatened the above two Labs with a lawsuit for using the polymerase.

firms stochastically and sequentially discover a technology that is critical to a cost-reduction process or to the development of a new product. Firms that have discovered the technology are referred to as innovators. When a discovery occurs, the innovator decides whether to seek patent protection or to rely on secrecy. We assume patent protection is probabilistic in that it is effective only with some probability. Moreover, we consider a legal environment in which a later innovator can be entitled to the exclusive use of the technology if previous innovators rely on secrecy protection.

Taking into account the uncertainty in patent protection and the threat of independent discoveries by rivals, an innovator's choice between patenting and secrecy becomes less than clear. In particular, by patenting, an innovator who initially seeks to exclude competitors, may actually provide helps by disclosing innovation information if the patent protection is ineffective. As Cohen et al. (2000) report, information disclosure is one of the main reasons that innovation firms do not patent. On the other hand, by adopting secrecy, an innovator with the intention to gain an edge over rivals, may not succeed if rival firms are able to discover the technology independently within a short period of time. As a matter of fact, blocking rivals from obtaining patents on related innovations is often a motive to patent.

We characterize the equilibrium of the model and show how innovators' patenting decisions depend on the timing of discovery (whether the discovery occurs early or late), the nature of the innovation (innovation arrival rate) and market structure (the number of firms). In particular, we show that early innovators are more inclined to choose secrecy while late innovators have a stronger tendency to patent. In other words, patenting incentives increase as more firms innovate. Consequently, given a level of patent protection, in equilibrium early innovators adopt secrecy and only a sufficiently late innovator chooses to patent. We provide a simple condition to identify the critical innovator who chooses to patent. Moreover, we find that firms' incentives to patent are greater if the innovation arrival rate is higher. This result helps explain why firms in hi-tech industries, which are featured by high innovation arrival rates, may choose patenting in spite of weak industry patent protection. Finally, we show that an increase in the number of firms may cause patenting to occur earlier or later, depending on the strength of patent protection. This suggests that an increase in competition need not necessarily promote innovation information disclosure.

Our analysis also sheds light on the important policy issue of socially optimal level of patent protection. In the model, we assume the arrivals of innovations are exogenously determined. Thus, we abstract from the issue of ex-ante innovation incentive. A patent is viewed as a contract or an agreement between the society and the innovator in the sense that certain monopoly power is granted in exchange for innovation information disclosure.[5] In

[5] The distinction of "reward theory" and "contract theory" of patents was first discussed in Denicolò and

particular, a social planner faces the following trade-off in choosing the optimal level of patent protection. For a weak patent protection, early innovators are more likely to adopt secrecy. Thus, the society has to endure markets in which firms have strong market powers, until the time when more firms innovate. To speed up the disclosure of innovation information, a stronger patent protection is necessary, which, however, is associated with a higher chance of monopoly market. We derive the socially optimal level of patent protection and show it should be lower if the innovation arrival rate is higher or the number of firms is larger.

A small body of literature has studied firms' patenting decisions under imperfect patent protection. However, these studies typically assume away the possibility that firms compete for patenting identical or similar innovations (Gallini, 1992, Horstmann, MacDonald and Slivinski, 1985, Anton and Yao, 2004). Departing from the literature, Kultti, Takalo and Toikka (2006, 2007) are the first to consider the situation where multiple firms that innovate independently choose between patenting and secrecy.[6] In their model, firms discover innovations simultaneously and decide whether to patent given the levels of patent and secrecy protection. Our model complements theirs in that we assume independent discoveries occur stochastically and sequentially.

Choi (1990) and Erkal (2005) study the decisions to patent in another interesting and important direction. In the framework of cumulative innovation, they examine an innovator's options to patent (and commercialize) the basic version of a product or to keep it as secrecy and continue to develop an improved version. They assume perfect patent protection and emphasize the competition among firms on the developments of two vertically differentiated products. The interest in our model differs from theirs. In particular, we consider the situation of identical innovations (or horizontally similar innovations) and where patent protection is probabilistic.

The rest of the paper is organized as follows. Section 2 describes the model. Equilibrium analysis is conducted in section 3. Section 4 performs comparative statics. Section 5 considers the socially optimal level of patent protection. Section 6 concludes. Proofs not in the text are collected in Appendix.

Franzoni (2004b).

[6] In another paper, Denicolò and Franzoni (2004a) provide an insightful analysis on the optimal patent design in a model with an innovation stage and a duplication stage, and ask the question of whether it is socially desirable to allow a second inventor (an imitator) to patent and whether the patent, if granted, contains the right to exclude the first inventor who adopts secrecy.

2 The Model

Consider an industry with a fixed number, n, of ex-ante identical firms. The firms are about to discover a technology that is crucial to a cost-reduction process or to the development of a new product.[7] The discovery process for each firm is independent and identical, and is determined by a Poisson process with an exogenous arrival rate λ.[8] Our reason for focusing on an exogenous innovation process is threefold. First, in a number of situations, a creative idea is essential for an innovation to occur. Once an idea arrives, it can be turned into an innovation with negligible costs. In addition, ideas are likely to arrive in a stochastic fashion. Thus, our model fits into certain innovation environments.[9] Second, the primary objective of this paper is to understand how firms make patenting decisions. Abstracting from investment choices allows us to disentangle the trade-off in patenting decision in a more transparent way. Third, as we will discuss in section 5, the assumption of exogenous innovation process serves the purpose of separating the function of patents to induce innovation information disclosure from the function to provide ex-ante innovation incentives.

When a discovery occurs, the firm decides whether to patent the technology or to maintain it as secret. To capture the fact that patent protection is probabilistic, we follow Kultti, Takalo and Toikka (2007) and assume that, with probability α, an innovator who applies for patent protection is granted an infinitely lived, perfectly effective property right on the technology; and with probability $1 - \alpha$, patent protection is ineffective, under which the technology becomes public and other firms can access to it. To simplify analysis, we normalize costs associated with patenting to zero.[10] By adopting secrecy, an innovator can use the technology until another innovator successfully obtains effective patent protection. To focus on the effect of multiple innovation discoveries, we assume that the technology information would not leak out if it is kept in secret use.[11]

Firms earn profits in a product market. We do not rely on a specific form of competition. Rather, we assume a general form of profit function that depends only on the number of producing firms. In particular, let π_i be the instantaneous profit for each firm when i firms produce in the product market. We assume π_i is strictly decreasing and convex in i.[12] Three

[7] For convenience, we restrict to one technology. Alternatively, one can think that the firms are about to discover different but similar technologies which are likely to be covered by one patent.

[8] Poisson process has been extensively used in the literature of economics of innovation. See Reinganum (1989) for a survey. Some researchers call λ hit rate or hazard rate.

[9] See Scotchmer (2004) and Erkal and Scotchmer (2009) for discussions on the models of innovation "ideas".

[10] Our model can easily incorporate the case of a positive patenting cost, τ, by scaling down the profit associated with patenting by τ.

[11] Thus, a firm can access to the technology information only if she discovers the technology or another firm applies for patent protection which, however, turns out to be ineffective.

[12] A simple example is Cournot competition with linear market demand and constant marginal production

possible scenarios may appear, each of which determines the number of producing firms and their profits: (1) if patent protection is effective, the patentee earns π_1 and others earn no profit; (2) if patent protection is ineffective, all firms produce and each earns π_n; (3) if i firms discover the technology and all opt for secrecy, each of these i firms earns π_i and others earn zero profit.

We abstract from any issues of asymmetric information and assume whether a firm has discovered the technology is common knowledge. The timing of the model is shown in *Figure 1.* Since firms are ex-ante identical, without loss of generality, we index firms by their ranks in discovery. Let innovator j (or firm j) be the jth firm that discovers the technology where $j \in N$ and $N = \{1, 2, \cdots, n\}$. Time is continuous. Period j is referred to as the time period that begins when innovator j discovers the technology, and ends when innovator $j+1$ discovers the technology. At the beginning of period j, innovator j decides whether to patent if no patent has been granted previously. If innovator j chooses to patent, nature will determine if the patent protection is effective. Alternatively, innovator j can keep the technology as secrecy. In such a case, the model moves on to period $j+1$ in which innovator $j+1$ discovers the technology and decides whether to patent.

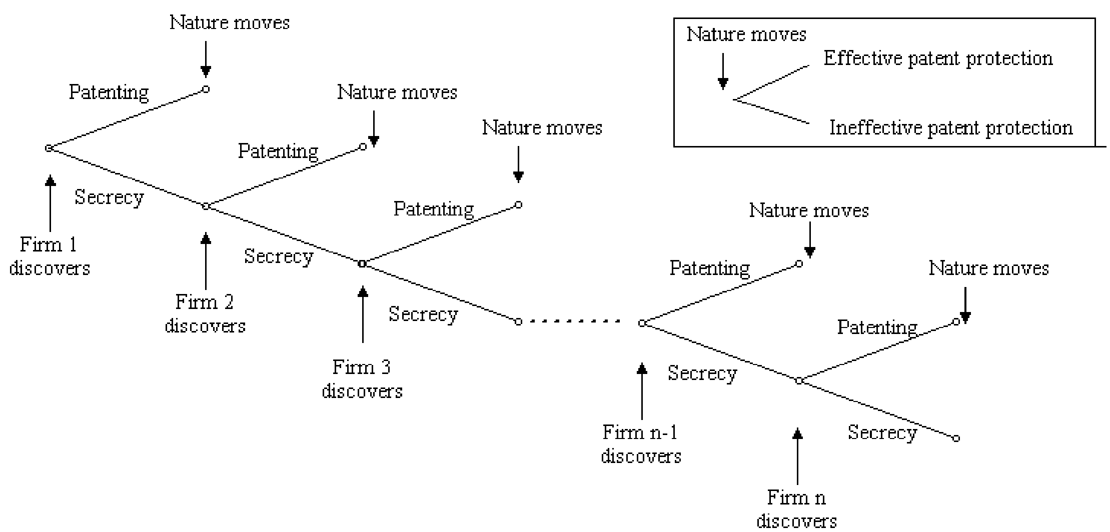

Figure 1 : Timing of the game

The model specifies an n-period dynamic game. Thus, the equilibrium concept is subgame

cost. Assume market demand is $P = a - bQ$ $(a, b > 0)$ where P and Q are market price and quantity, respectively. Let q be the output for each firm in a symmetric equilibrium. Denote c as the marginal cost. Profit maximization and symmetric condition lead to $q = \frac{a-c}{b(n+1)}$ and $\pi_n = \frac{(a-c)^2}{b(n+1)^2}$. We can verify that $\frac{\partial \pi_n}{\partial n} < 0$ and $\frac{\partial^2 \pi_n}{\partial n^2} > 0$.

perfect Nash equilibrium (SPNE). Given no previous patent has been granted, an innovator, taking into account the optimal strategies of subsequent innovators, chooses between patenting and secrecy to maximize expected profit. In equilibrium, innovators' patenting decisions map from N into $\{P, S\}$ where P and S stand for patenting and secrecy respectively.

3 Equilibrium Analysis

In deciding whether to patent, a firm compares the expected profits from the strategies of patenting and secrecy. Since innovator j decides whether to patent at the beginning of period j, the future profit streams should be discounted as present values to that point. Here, we derive some preliminary results that are useful throughout the paper.

3.1 Preliminaries

(I) First, we calculate the present value for innovator j if she receives a stream of profit π through the entire period j. Let T_j denote the time length of period j. Note that T_j is distributed as a Poisson process with industry arrival rate $\Gamma(j) = (n-j)\lambda$.[13] Thus, it has probability density function $\Gamma e^{-\Gamma T_j}$. For a profit stream π through the entire period j, the present value of such a profit stream with a fixed time length T is

$$\int_0^T \pi e^{-rt} dt = \frac{1-e^{-rT}}{r}\pi.$$

Thus, the present value of the profit stream with a random time length T_j is:

$$\begin{aligned} & \int_0^\infty \left(\int_0^{T_j} \pi e^{-rt} dt\right)(n-j)\lambda e^{-(n-j)\lambda T_j} dT_j \\ = & \int_0^\infty \frac{1-e^{-rT_j}}{r}\pi(n-j)\lambda e^{-(n-j)\lambda T_j} dT_j \\ = & \frac{1}{r+(n-j)\lambda}\pi = \frac{1}{r}\theta_{n-j}\pi \end{aligned} \tag{1}$$

where θ_{n-j} is defined as

$$\theta_{n-j} = \frac{r}{r+(n-j)\lambda}. \tag{2}$$

(II) Second, we calculate the present value for innovator j if she receives an instantaneous profit π at the beginning of period $j+1$. The present value of such an instantaneous profit

[13] The sum of independent Poisson processes is a Poisson process with the arrival rate equaling the sum of individual arrival rates. For detailed derivations, see Theorem 5.1 in Taylor and Karlin (1984). When jth firm discovers the technology, $n-j$ firms remains to innovate. Thus, the industry arrival rate is $(n-j)\lambda$.

with a fixed time length T is πe^{-rT}. Thus, the present value of the instantaneous profit with a random time length T_j is

$$\begin{aligned} & \int_0^\infty \pi e^{-rT_j}(n-j)\lambda e^{-(n-j)\lambda T_j} dT_j \\ = & \frac{(n-j)\lambda}{r+(n-j)\lambda}\pi = (1-\theta_{n-j})\,\pi. \end{aligned} \tag{3}$$

(III) Third, from (1) and (3), we can show that if innovator j receives a stream of profit π in period h $(h > j)$, the present value of the profit stream is

$$\frac{1}{r}\theta_{n-h}(1-\theta_{n-h+1})(1-\theta_{n-h+2})\cdots(1-\theta_{n-j})\pi. \tag{4}$$

To see (4), note that, by (1), the present value at the beginning of period h for a stream of profit π in period h is $\frac{1}{r}\theta_{n-h}\pi$. By (3), multiplying $\frac{1}{r}\theta_{n-h}\pi$ by $(1-\theta_{n-h+1})$ gives the present value at the beginning of period $h-1$. Applying the same logic repeatedly, we can show (4) is the present value at the beginning of period j for a stream of profit π in period h.

3.2 Expected Profit from Patenting

Conditional on that all previous innovators adopt secrecy, we consider innovator j's expected profit if she chooses to patent. With probability α, she is awarded with effective patent protection, reaping monopoly profit π_1. With probability $1-\alpha$, patent protection is ineffective, and innovator j earns profit π_n. Hence, the expected payoff for innovator j to patent is:

$$\begin{aligned} \Pi_p &= \int_0^\infty [\alpha\pi_1 + (1-\alpha)\pi_n]e^{-rt}dt \\ &= \frac{1}{r}\left[\alpha\pi_1 + (1-\alpha)\pi_n\right]. \end{aligned} \tag{5}$$

Note that Π_p is invariant with the rank in discovery and does not depend on the patenting strategies of subsequent innovators. This is because once a firm chooses to patent, the uncertainty of patent protection fully reveals.

3.3 Expected Profit from Secrecy

If innovator j adopts secrecy, the expected profit, unlike in the case of patenting, does depend on the strategies of subsequent innovators. Let $\Pi_s(j|h)$ $(h > j)$ denote the expected payoff for innovator j to adopt secrecy, conditional on that innovator h chooses to patent. By (4),

$$\Pi_s(j|h) = \sum_{i=j}^{h-1} \frac{1}{r}\theta_{n-i}(1-\theta_{n-i+1})\cdots(1-\theta_{n-j})\pi_i + \frac{1}{r}(1-\theta_{n-h+1})\cdots(1-\theta_{n-j})(1-\alpha)\pi_n. \quad (6)$$

The first term (the summation term) is the expected profit associated with secrecy protection from period j through period $h-1$. The second term represents the expected profit from period h and subsequent periods. Given innovator h chooses to patent, innovator j can earn π_n in or after period h only if the patent protection is ineffective, which occurs with probability $(1-\alpha)$.

Let

$$\Pi_s(j) = \Pi_s(j|j+1)$$

which denotes the expected profit if innovator j opts for secrecy given innovator $j+1$ chooses to patent. By (6), we have

$$\Pi_s(j) = \frac{1}{r}\theta_{n-j}\pi_j + \frac{1}{r}(1-\theta_{n-j})(1-\alpha)\pi_n. \quad (7)$$

3.4 Equilibrium Choice of Patenting or Secrecy

To avoid mixed strategies, we assume that a firm chooses to adopt secrecy if patenting and secrecy yield the same expected profit. To pin down the equilibrium, we consider an innovator's patenting decision conditional on that the next innovator chooses to patent and to adopt secrecy, respectively. Results are summarized in Lemma 1 and Lemma 2.

To begin, consider innovator j's patenting decision if the next innovator chooses to patent. Clearly, innovator j chooses to patent if

$$\Pi_p > \Pi_s(j). \quad (8)$$

Define

$$\alpha_j = \frac{\pi_j - \pi_n}{\frac{\pi_1}{\theta_{n-j}} - \pi_n} \qquad for \quad j = 1, ..., n. \quad (9)$$

By (5) and (7), (8) becomes

$$\alpha > \alpha_j \quad (10)$$

Thus, α_j can be interpreted as the incentive for innovator j to patent if the next innovator chooses to patent. A smaller α_j implies a higher incentive to patent.

Lemma 1 *α_j strictly decreases with j.*

According to Lemma 1, conditional on that the next innovator chooses to patent, a later innovator has a higher incentive to patent. The intuition is as follows. As more rivals discover the technology, an innovator makes less profit from secrecy. However, patenting provides invariant expected profit for all firms regardless of the timing of discovery. As a result, early innovators are more inclined to choose secrecy while later innovators have stronger incentives to patent.

Next, we turn to describing innovator j's strategy if innovator $j+1's$ optimal strategy is to adopt secrecy.

Lemma 2 *If choosing secrecy over patent is optimal for innovator $j+1$, the same strategy has to be optimal for innovator j.*

The key to understanding this result is that innovator j earns more profit than innovator $j+1$ does when they both adopt secrecy, regardless of the strategies of subsequent innovators. Moreover, both innovators receive identical profits from patenting as in (5). Therefore, if innovator $j+1$ finds it optimal to choose secrecy over patenting, innovator j should find the same strategy to be optimal as well.

One immediate result that follows from Lemma 2 is that if innovator $j+1$ optimally chooses secrecy, all previous innovators will optimally choose secrecy as well in equilibrium. With Lemma 1 and Lemma 2, we are in a position to characterize the equilibrium of the model.

Proposition 1 *Let $\alpha_0 = 1$. Given the level of patent protection α, there exists a unique $m \in N$ such that $\alpha_m < \alpha \leq \alpha_{m-1}$. In equilibrium, innovator m chooses to patent while previous innovators (if any) adopt secrecy.*

Proof. By Lemma 1, $\{\alpha_j\}_{j\in N}$ is a strictly decreasing sequence and thus, divides $[0,1]$ into non-overlapping intervals. It follows that, given $0 \leq \alpha \leq 1$, a unique $m \in N$ exists such that $\alpha_m < \alpha \leq \alpha_{m-1}$.

To show the second half of the proposition, we use backward induction. Consider the choice of the last innovator (innovator n). Comparing the profit from patenting, Π_p, to that from secrecy, $\Pi_s(n)$, we find that $\Pi_s(n) = \frac{1}{r}\pi_n < \frac{1}{r}[\alpha\pi_1 + (1-\alpha)\pi_n] = \Pi_p$. That is, the last innovator will choose to patent. Intuitively, this is because secrecy does not provide extra benefit since all other firms have discovered the technology. Given that innovator n will patent, we consider the choice of innovator $n-1$. If $m = n$, then $\alpha < \alpha_{n-1}$, which implies that innovator $n-1$ will choose secrecy. By Lemma 2, innovator j $(j < m)$, if any, opts for secrecy. If $m < n$, we have $\alpha > \alpha_{n-1}$. Thus, innovator $n-1$ chooses to patent. Since $\{\alpha_j\}$ is strictly decreasing with j, we can show that, for $\alpha > \alpha_m$, innovator j $(j \geq m)$ chooses to patent. In addition, since $\alpha \leq \alpha_{m-1}$ it follows that innovator $m-1$ chooses secrecy over patenting. By Lemma 2, it is straightforward to show innovator j $(j < m)$ opts for secrecy. ■

Proposition 1 provides a simple characterization of the equilibrium. Depending on the strength of patent protection, the innovation arrival rate, market structure and the timing of discovery, firms may choose different means to protect innovations. Two scenarios may occur in equilibrium. First, the first innovator chooses to patent. Second, it is possible that firms that innovate early opt for secrecy while only a sufficiently late innovator chooses to patent.

The following example illustrates Proposition 1.

Example 1 *Let $n = 3$, $\lambda = 0.1$, $r = 0.2$. Moreover, we assume linear market demand, $P = a - bQ$, and constant marginal cost, c. By (9), we compute that $\alpha_1 = 0.43$, $\alpha_2 = 0.16$, and $\alpha_3 = 0$. Therefore, if $\alpha > \alpha_1$, in equilibrium, innovator 1 patents. If $\alpha_1 \geq \alpha > \alpha_2$, in equilibrium, innovator 1 adopts secrecy while innovator 2 patents. If $\alpha_2 \geq \alpha > \alpha_3$ in equilibrium, innovator 1 and innovator 2 adopt secrecy while innovator 3 patents.*

We end this section by showing how innovators' expected profits depend on the timing of discoveries.

Proposition 2 *Innovators expected profits decrease with their ranks of discovery.*

Proof. Suppose firm m patents. By Proposition 1, firm i $(i < m)$ opts for secrecy. By Lemma 2, $\Pi_s(j|m) > \Pi_s(j+1|m)$. Hence, the expected profit decreases with j when $j < m$. In addition, Π_p is the expected profit for firm m. Firm $m-1$ opts for secrecy, which implies $\Pi_s(m-1) > \Pi_p$. Finally, firm j $(j > m)$ earns

$$\frac{1}{r}(1-\alpha)\pi_n < \frac{1}{r}\left[\alpha\pi_1 + (1-\alpha)\,\pi_n\right] = \Pi_p.$$

This completes the proof. ■

4 Comparative Statics

In this section, we examine how changes in the strength of patent protection, the innovation arrival rate and the number of firms affect the incentive to patent and the timing of patenting. Proposition 1 shows that there is a unique $m = m(\alpha, \lambda, n)$ such that innovator m patents and previous innovators (if any) opt for secrecy. Define $\rho(\alpha, \lambda, n)$ as the proportion of firms that adopt secrecy:

$$\rho(\alpha, \lambda, n) = \frac{m(\alpha, \lambda, n) - 1}{n} \tag{11}$$

Since the industry innovation arrival rate during period i is $(n-i)\lambda$, the expected length of period i is

$$T_i(n, \lambda) = \frac{1}{(n-i)\lambda}.$$

Define $T(\alpha, \lambda, n)$ as the expected time when patenting occurs:

$$T(\alpha, \lambda, n) = \sum_{i=1}^{m(\alpha,\lambda,n)-1} T_i(n, \lambda). \tag{12}$$

We first show the effect of a change in the level of patent protection α.

Proposition 3 *$m(\alpha, \lambda, n)$, $\rho(\alpha, \lambda, n)$ and $T(\alpha, \lambda, n)$ decrease with α.*

The intuition is straightforward. Strengthening patent protection directly increases the profit from patenting. At the same time, it reduces the profit from secrecy because subsequent innovators have greater chances of obtaining effective patent protection. Therefore, a higher α encourages firms to choose patenting and thus, advances the timing of patenting.

We next study the effect of a change in the innovation arrival rate λ.

Proposition 4 *$m(\alpha, \lambda, n)$, $\rho(\alpha, \lambda, n)$ and $T(\alpha, \lambda, n)$ decrease with λ.*

An increase in the innovation arrival rate does not affect firms' profits from patenting. However, it shortens the length during which an innovator enjoys profit from secrecy because the discoveries by rival firms arrive more quickly. Thus, profit from secrecy decreases with λ. As a result, innovators have more incentive to patent and thus, patenting occurs earlier.

The result that firms prefer patenting under a larger λ may help explain why firms in hi-tech industries find patenting attractive in spite of relatively weak industry patent protection. This is because independent discoveries are likely to happen frequently in hi-tech industries. Expecting that rivals will discover the technology soon, firms find secrecy protection has little value and, as a consequence, choose to patent even if the patent protection is weak.

Define

$$\tilde{\lambda}(\alpha) = \frac{1}{n-1} \frac{1-\alpha}{\alpha} r \left(\frac{\pi_1 - \pi_n}{\pi_1} \right). \tag{13}$$

When $\lambda > \tilde{\lambda}$, by (9) and (2),

$$\begin{aligned}
\alpha_1 &= \frac{\pi_1 - \pi_n}{\frac{\pi_1}{\frac{r}{r+(n-1)\lambda}} - \pi_n} = \frac{r(\pi_1 - \pi_n)}{(n-1)\pi_1\lambda + r(\pi_1 - \pi_n)} \\
&< \frac{r(\pi_1 - \pi_n)}{(n-1)\pi_1 \frac{1}{n-1}\frac{1-\alpha}{\alpha} r\left(\frac{\pi_1 - \pi_n}{\pi_1}\right) + r(\pi_1 - \pi_n)} \\
&= \alpha.
\end{aligned}$$

By Proposition 1, it follows that the first innovator chooses to patent.

Corollary 1 *Given patent protection level α, patenting is a dominant strategy if $\lambda > \tilde{\lambda}(\alpha)$ where $\tilde{\lambda}(\alpha)$ is defined as in (13).*

In other words, given a patent protection level, there always exists a sufficiently large λ such that the first innovator applies for patent protection. As a special case, when discoveries occur almost simultaneously, that is, $\lambda \to \infty$, patenting is a dominant strategy.

Finally, we examine how market structure affects the incentives to patent and the timing of patenting.

Proposition 5 *There exists a $\tilde{j}$ such that as the number of firms increases, firm j's incentive to patent is higher (lower) if $j < \tilde{j}$ ($j > \tilde{j}$). Consequently, there exists an $\tilde{\alpha}$ such that as the number of firms increases, patenting advances (delays) when $\alpha > \tilde{\alpha}$ ($\alpha < \tilde{\alpha}$).*

The key to understanding Proposition 5 is that an increase in the number of firms has effects on the profit from both patenting and secrecy. On one hand, an increase in the number of firms decreases the profit from patenting since more firms produce in output market if patent protection is ineffective. We refer to this effect as the *free ride effect*. Note that the magnitude of the *free ride effect* is the same for early and late innovators. On the other hand, an increase in the number of firms also decreases the profit from secrecy because more firms are in innovation race, causing the next innovation to occur sooner. We label this effect as the *racing effect*. However, compared to late innovators, early innovators are affected more significantly as they face more potential competitors that race for discoveries. In other words, the *racing effect* is more prominent for early innovators. As shown in *Figure 2*, for early (late) innovators, incentives to patent increase (decrease) as the thresholds, above which innovators will patent, are lower (higher).

Obviously, whether the timing of patenting advances or delays depends on whether an early innovator or a late innovator patents in the equilibrium which, in turn, is determined by the level of patent protection. When patent protection is strong, an early innovator patents in equilibrium. Since a higher number of firms increases the patenting incentive of early innovators, it causes patenting to occur earlier. When patent protection is weak, a late innovator patents in equilibrium. In this case, an increase in the number of firms lowers the late innovator's incentive to patent which delays the timing of patenting.

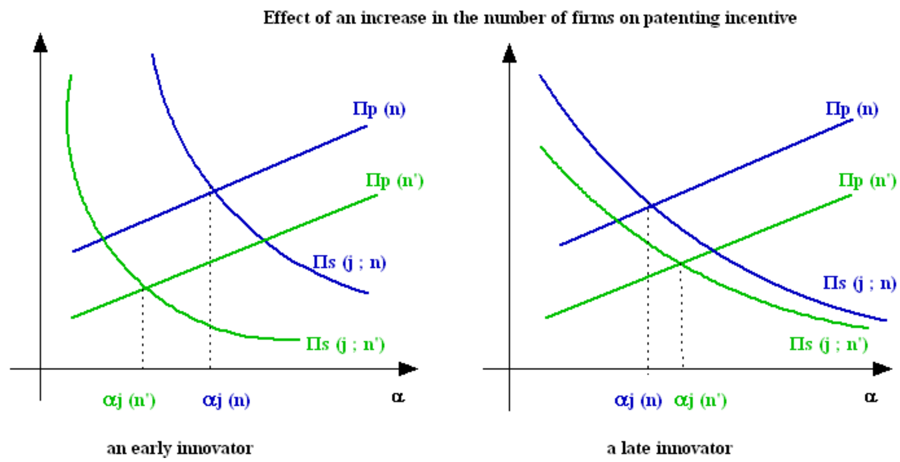

Figure 2

Proposition 5 has important implications. An increase in competition does not necessarily increase the incentive to patent. More importantly, when the patent protection is initially weak, it delays the timing of patenting and thus, innovation information disclosure.

5 Socially Optimal Patent Protection

When an innovator considers the choice between patenting and secrecy, she does not internalize the effects on consumer surplus and the profits of other firms. This section addresses the following question: given firms' patenting strategies, what is the optimal level of patent protection that maximizes social welfare?

There are two distinct perspectives on the function of patents. First, patents are considered as rewards for innovators. The idea is as follows. Without patent protection, an innovation can be easily imitated. Thus, firms may not be able to capture enough profits to cover the costs associated with the innovation. Expecting this, firms may not invest. As a result, the innovation would not occur at the first place. Under this view, the main goal of patents is to provide sufficient ex-ante innovation incentive.

In this paper, we focus on another function of patents. In particular, following Denicolò and Franzoni (2004b), we view a patent as a contract or an agreement between society and an innovator in the sense that the society gives some exclusive right to the innovator in exchange for innovation information disclosure. Under this view, the main function of patents is to induce information disclosure after innovations take place. In our model, innovations occur following a random process with an exogenous arrival rate. This structure allows us to abstract from patents' role of providing ex-ante innovation incentive and to focus on patents' function to induce innovation information disclosure.

On one hand, innovation information disclosure benefits the society because, with some probability, other firms may utilize the innovation. This leads to a more competitive product market, and thus, increases the social welfare. On the other hand, a firm would not disclose information unless she receives sufficient compensation. This implies that a sufficiently strong patent protection has to be granted to trigger information disclosure. However, a stronger patent protection is associated with a higher chance of monopoly which reduces social welfare. Therefore, an optimal patent protection has to strike the balance of the provision of patenting incentive and the avoidance of unnecessary monopoly.

Let S_k be the instantaneous social welfare when k firms produce in the product market.[14] We assume S_k strictly increases with k. Suppose that, given α, innovator m patents in equilibrium. Define total social welfare, $TS(\alpha)$, as the sum of discounted instantaneous social welfare:

$$\begin{aligned} TS(\alpha) &= \frac{1}{r}\theta_{n-1}S_1 \\ &+\frac{1}{r}\theta_{n-2}(1-\theta_{n-1})S_2 \\ &+\cdots \\ &+\frac{1}{r}\theta_{n-m+1}(1-\theta_{n-m+2})\cdots(1-\theta_{n-1})S_{m-1} \\ &+\frac{1}{r}(1-\theta_{n-m+1})(1-\theta_{n-m+2})\cdots(1-\theta_{n-1})[\alpha S_1+(1-\alpha)S_n]. \end{aligned} \tag{14}$$

The first $m-1$ lines are the discounted social welfare in the first $m-1$ periods when innovators opt for secrecy while the last line is the discounted social welfare when innovator m patents. A social planner's objective function is

$$\max_{\alpha\in[0,1]} TS(\alpha). \tag{15}$$

Consider an increase in the level of patent protection from α_1 to α_2 such that $\alpha_{m-1} < \alpha_1 < \alpha_2 \leq \alpha_m$. In words, the change in α is sufficiently small such that m does not alter in equilibrium. The only effect of the increase in α is a higher chance of monopoly for innovator

[14] The instantaneous social welfare is defined as the sum of consumer surplus and producer surplus.

m. As a result, total social welfare decreases. This leads to the following lemma.

Lemma 3 *Total social welfare can be increased if a reduction of α results in the same m in equilibrium.*

According to Lemma 3, we can greatly cut down the set of possible α leading to total social welfare maximization. In particular, total social welfare maximization has to occur when α is only sufficient to induce a switch in m. In other words, the optimal level of patent protection has to occur at one of the $\{\alpha_j\}_{j\in N}$. Hence, we only need to compare n possible equilibrium outcomes. The social planner's problem is simplified to

$$\max_{\alpha\in\Omega(\alpha)} TS(\alpha) \quad \text{where} \quad \Omega(\alpha)=\{\alpha_j,\ j\in N\}. \tag{16}$$

The social planner chooses an α_j to maximize total social welfare. However, each α_j is uniquely associated with an m. Therefore, it is as if the social planner chooses m to maximize total social welfare. The next proposition states that socially optimal level of patent protection should induce the first innovator to patent.

Proposition 6 *The socially optimal level of patent protection α^* should be such that it is only sufficient to induce the first innovator to patent. That is, $\alpha^*=\alpha_1$.*

For weak patent protection, early innovators are more likely to adopt secrecy. Thus, the society has to endure markets in which firms have strong market powers until more firms innovate. To speed up the disclosure of innovation information, a stronger patent protection is necessary which, however, increases the chances of monopoly. We find that the socially optimal patent protection should be such that it is just sufficient to induce the first innovator to patent.

To better illustrate the intuition behind the proposition, consider the case of two firms. From Lemma 3, the socially optimal patent protection is either α_1, which induces the first innovator to patent, or an arbitrarily small patent protection such that in equilibrium, the first innovator opts for secrecy and the second innovator chooses to patent. In the former case, the social welfare is $\frac{1}{r}\left[\alpha_1 S_1+(1-\alpha_1)S_2\right]$, a randomization between monopoly and duopoly markets. In the latter case, social welfare is $\frac{1}{r}\left[\theta_1 S_1+(1-\theta_1)S_2\right]$, a monopoly market followed by a duopoly market. Thus, inducing the first innovator to patent improves social welfare if and only if $\alpha_1<\theta_1$. Note that profit from patenting is $\frac{1}{r}\left[\alpha_1\pi_1+(1-\alpha_1)\pi_2\right]$ and profit from secrecy is $\frac{1}{r}\left[\theta_1\pi_1+(1-\theta_1)(1-\alpha_1)\pi_2\right]<\frac{1}{r}\left[\theta_1\pi_1+(1-\theta_1)\pi_2\right]$. Thus, the first innovator demands a level of patent protection that is less than θ_1 to choose patenting over secrecy. It follows that granting a level of patent protection that induces the first innovator to patent yields a higher social welfare.

From (2), we have $\theta_1 = \frac{r}{r+\lambda} < 1$. Hence, from (9), $\alpha_1 < 1$. We have the following corollary.

Corollary 2 *Full patent protection ($\alpha = 1$) is never socially optimal when $n > 1$.*

A monopoly firm in an innovation market demands full patent protection in exchange for revealing innovation information because the firm does not face any potential threat. The situation changes in an oligopoly market. If the first innovator opts for secrecy it could potentially be excluded by a later innovator that obtains a patent. Thus, the level of patent protection that would induce the first innovator to disclose the information should be less than full protection.

Finally, by (9) and (2), it is straightforward to show the following proposition.

Proposition 7 *α^* decreases, respectively, with λ and n.*

Proposition 7 implies that the optimal level of patent protection should be varying with the nature of an innovation and industry market structure. In particular, the optimal level of patent protection should be lower if the innovation arrival rate is higher or the number of firms is larger. The intuition behind these results is easy to see. For a higher innovation arrival rate or a larger number of firms, by Proposition 4 and Proposition 5, the first innovator receives less profit from secrecy and thus, demands a lower level of patent protection in exchange for the disclosure of innovation information.

6 Conclusion

The heterogeneous choices to patent or to maintain secrecy are well documented. We show how these choices may arise as a market equilibrium by developing a dynamic model in which innovations occur stochastically and independently. We further show how innovators' incentives to patent depend on the nature of innovation and market structure. Focusing on patents' function to induce innovation information disclosure, we find that the optimal level of patent protection should be lower when the innovation arrival rate is higher or the number of firms is larger.

There are a number of directions for future research. First, to capture certain innovation environments and to focus on firms' patenting decisions, we make an assumption that innovations arrive exogenously. It would be desirable to extend to a model with endogenous innovation arrival rates and to examine how patenting decisions and investment decisions are jointly determined. Another direction for future research would be to examine how firms' patenting decisions depend on the nature of innovations and market structure in the framework of cumulative innovation.[15] Finally, it would be interesting to extend our model to a vertical industry structure and to see how the presence of vertical integration affects an upstream innovation firm's incentive to patent.

Appendix

Proof of Lemma 1

Proof. By (9), for any $i = 1, ..., n-1$,

$$\begin{aligned} \alpha_j - \alpha_{j+1} &= \frac{\pi_j - \pi_n}{\frac{\pi_1}{\theta_{n-j}} - \pi_n} - \frac{\pi_{j+1} - \pi_n}{\frac{\pi_1}{\theta_{n-j-1}} - \pi_n} \\ &= \frac{(\pi_j - \pi_n)\left(\frac{\pi_1}{\theta_{n-j-1}} - \pi_n\right) - (\pi_{j+1} - \pi_n)\left(\frac{\pi_1}{\theta_{n-j}} - \pi_n\right)}{\left(\frac{\pi_1}{\theta_{n-j}} - \pi_n\right)\left(\frac{\pi_1}{\theta_{n-j-1}} - \pi_n\right)}. \end{aligned}$$

By (2), $\theta_{n-j} < 1$ and $\theta_{n-j-1} = \frac{r}{r+(n-j-1)\lambda} < 1$. It follows that $\frac{\pi_1}{\theta_{n-j}} - \pi_n > 0$ and $\frac{\pi_1}{\theta_{n-j-1}} - \pi_n > 0$ because π_i is strictly decreasing in i. Hence,

$$sign\ (\alpha_j - \alpha_{j+1}) = sign\ \left[(\pi_j - \pi_n)\left(\frac{\pi_1}{\theta_{n-j-1}} - \pi_n\right) - (\pi_{j+1} - \pi_n)\left(\frac{\pi_1}{\theta_{n-j}} - \pi_n\right)\right]$$

Note that, by (2),

$$\begin{aligned} &(\pi_j - \pi_n)\left(\frac{\pi_1}{\theta_{n-j-1}} - \pi_n\right) - (\pi_{j+1} - \pi_n)\left(\frac{\pi_1}{\theta_{n-j}} - \pi_n\right) \\ = &(\pi_j - \pi_n)\left(\frac{r + (n-j-1)\lambda}{r}\pi_1 - \pi_n\right) - (\pi_{j+1} - \pi_n)\left(\frac{r + (n-j)\lambda}{r}\pi_1 - \pi_n\right) \\ = &\frac{\lambda}{r}\pi_1\left[(n-j-1)\pi_j - (n-j)\pi_{j+1} + \pi_n\right] + (\pi_j - \pi_{j+1})(\pi_1 - \pi_n). \end{aligned} \tag{17}$$

Since π_i is strictly decreasing and convex in j we have $\left(\frac{n-j-1}{n-j}\right)\pi_j + \left(\frac{1}{n-j}\right)\pi_n > \pi_{j+1}$. Rearranging the inequality gives $(n-j-1)\pi_j - (n-j)\pi_{j+1} + \pi_n > 0$.Therefore, the expressions in (17) is positive. Thus, $sign\ (\alpha_j - \alpha_{j+1}) > 0$. It follows that $\alpha_j > \alpha_{j+1}$ for $i = 1, ..., n-1$.That is, $\{\alpha_j\}$ is a decreasing sequence. ■

[15] See Green and Scotchmer (1995) and Erkal (2005) for models of cumulative innovation.

Proof of Lemma 2

Proof. Since the game in the model assumes complete information, all firms correctly expect the strategies of subsequent firms. Suppose it is expected that firm h ($h > j+1$) will patent when she discovers the technology. By (6), the expected profit associated with secrecy for firm j is

$$\Pi_s(j|h) = \sum_{i=j}^{h-1} \frac{1}{r}\theta_{n-i}(1-\theta_{n-i+1})\cdots(1-\theta_{n-j})\pi_i + \frac{1}{r}(1-\theta_{n-h+1})\cdots(1-\theta_{n-j})(1-\alpha)\pi_n \ ,$$

and the expected profit from secrecy for firm $j+1$ is

$$\Pi_s(j+1|h) = \sum_{i=j+1}^{h-1} \frac{1}{r}\theta_{n-i}(1-\theta_{n-i+1})\cdots(1-\theta_{n-j-1})\pi_i + \frac{1}{r}(1-\theta_{n-h+1})\cdots(1-\theta_{n-j-1})(1-\alpha)\pi_n.$$

To compare $\Pi_s(j|h)$ and $\Pi_s(j+1|h)$, we define two auxiliary variables:

$$\beta_i = \frac{1}{r}\theta_{n-i}(1-\theta_{n-i+1})\cdots(1-\theta_{n-j}) \qquad \text{and} \qquad \delta_i = \frac{1}{r}\theta_{n-i}(1-\theta_{n-i+1})\cdots(1-\theta_{n-j-1}).$$

The expected profits from secrecy for firm j and firm $j+1$ become respectively

$$\Pi_s(j|h) = \sum_{i=j}^{h-1} \beta_i\pi_i + (\frac{1}{r} - \sum_{i=j}^{h-1}\beta_i)(1-\alpha)\pi_n$$

and

$$\Pi_s(j+1|h) = \sum_{i=j+1}^{h-1} \delta_i\pi_i + (\frac{1}{r} - \sum_{i=j+1}^{h-1}\delta_i)(1-\alpha)\pi_n.$$

Note that

$$\begin{aligned}\sum_{i=j}^{h-1}\beta_i &= \frac{1}{r}\theta_{n-j} + \frac{1}{r}\theta_{n-j-1}(1-\theta_{n-i}) + \cdots + \frac{1}{r}\theta_{n-h+1}(1-\theta_{n-h+2})\cdots(1-\theta_{n-j}) \\ &\quad + \frac{1}{r}(1-\theta_{n-h+1})\cdots(1-\theta_{n-j}) - \frac{1}{r}(1-\theta_{n-h+1})\cdots(1-\theta_{n-j}) \\ &= \frac{1}{r} - \frac{1}{r}(1-\theta_{n-h+1})\cdots(1-\theta_{n-j}). \qquad (18)\end{aligned}$$

Similarly, we have

$$\sum_{i=j+1}^{h-1}\delta_i = \frac{1}{r} - \frac{1}{r}(1-\theta_{n-h+1})\cdots(1-\theta_{n-j-1}). \qquad (19)$$

Substituting (18) into $\Pi_s(j|h)$ and (19) into $\Pi_s(j+1|h)$ and taking difference give

$$\begin{aligned}
&\Pi_s(j|h) - \Pi_s(j+1|h) \\
= \; & \sum_{i=j}^{h-1} \beta_i \pi_i + (\frac{1}{r} - \sum_{i=j}^{h-1} \beta_i)(1-\alpha)\pi_n - \sum_{i=j+1}^{h-1} \delta_i \pi_i - (\frac{1}{r} - \sum_{i=j+1}^{h-1} \delta_i)(1-\alpha)\pi_n \\
> \; & \left(\sum_{i=j}^{h-1} \beta_i - \sum_{i=j+1}^{h-1} \delta_i \right) \pi_{h-1} - (\sum_{i=j}^{h-1} \beta_i - \sum_{i=j+1}^{h-1} \delta_i)(1-\alpha)\pi_n \\
= \; & (\sum_{i=j}^{h-1} \beta_i - \sum_{i=j+1}^{h-1} \delta_i)[\pi_{h-1} - (1-\alpha)\pi_n] > 0.
\end{aligned}$$

The last inequality holds because $\sum_{i=j}^{h-1} \beta_i > \sum_{i=j+1}^{h-1} \delta_i$ which follows by (18) and (19). Thus, $\Pi_s(j|h) > \Pi_s(j+1|h)$.

Finally, given that firm $j+1$ optimally opts for secrecy, we have $\Pi_s(j+1|h) > \Pi_p$. It follows that $\Pi_s(j|h) > \Pi_p$. That is, firm j opts for secrecy. ■

Proof of Proposition 3

Proof. Suppose $\hat{\alpha} > \alpha$. Let $m = m(\alpha, \lambda, n)$ and $\hat{m} = m(\hat{\alpha}, \lambda, n)$. In equilibrium, we have $\alpha_m < \alpha \leq \alpha_{m-1}$. It follows that $\alpha_m \leq \hat{\alpha}$. Since a change in α does not change α_j and note that $\hat{\alpha} \in (\alpha_{\hat{m}}, \alpha_{\hat{m}-1}]$, it follows that $\alpha_m \leq \alpha_{\hat{m}}$. By Lemma 1, $\hat{m} \leq m$. Hence, $\hat{\rho} \leq \rho$. Let $T = T(\alpha, \lambda, n)$ and $\hat{T} = T(\hat{\alpha}, \lambda, n)$. When $\hat{m} \leq m$, we have $\hat{T} \leq T$. ■

Proof of Proposition 4

Proof. We first show that α_j decreases with λ. From equation (9), we have α_j increases in θ_{n-j}. By (2), θ_{n-j} decreases in λ. Therefore, α_j decreases in λ. Hence, given that $\hat{\lambda} > \lambda$, we have $\alpha_j(\hat{\lambda}) < \alpha_j(\lambda)$. Given λ, in equilibrium, $\alpha_m(\lambda) < \alpha \leq \alpha_{m-1}(\lambda)$. Hence, $\alpha > \alpha_m(\hat{\lambda})$. Let $\hat{m} = m(\alpha, \hat{\lambda}, n)$ and $\hat{\rho} = \rho(\alpha, \hat{\lambda}, n)$. Given $\hat{\lambda}$, we have $\hat{m} \leq m$. It follows that $\hat{\rho} \leq \rho$. Let $\hat{T} = T(\alpha, \hat{\lambda}, n)$. When $\hat{m} \leq m$ and $\hat{\lambda} > \lambda$, we have $\hat{T} \leq T$. ■

Proof of Proposition 5

Proof. Step 1: We show that, for each α_j there exists a cutoff value,

$$\lambda_j = \frac{r\left(\pi_1 - \pi_j\right)\left(\pi_n - \pi_{n+1}\right)}{\pi_1\left[\pi_j - \pi_n - (n-j)\left(\pi_n - \pi_{n+1}\right)\right]} \tag{20}$$

such that α_j increases with n when $\lambda < \lambda_j$ but decreases with n when $\lambda > \lambda_j$.

To see this, we take the difference of $\alpha_j(n)$ and $\alpha_j(n+1)$. By (9),

$$\alpha_j(n) - \alpha_j(n+1) = \frac{(\pi_j - \pi_n)\left(\frac{\pi_1}{\theta_{n+1-j}} - \pi_{n+1}\right) - (\pi_j - \pi_{n+1})\left(\frac{\pi_1}{\theta_{n-j}} - \pi_n\right)}{\left(\frac{\pi_1}{\theta_{n-j}} - \pi_n\right)\left(\frac{\pi_1}{\theta_{n+1-j}} - \pi_{n+1}\right)}.$$

Clearly, the denominator of the right hand side of the equation is postive since $\theta_{n-j} < 1$. Substituting (2) into the numerator of right-hand side of the equation and rearranging terms, we have

$$sign[\alpha_j(n) - \alpha_j(n+1)] = sign[(\pi_j - \pi_n) - (n-j)(\pi_n - \pi_{n+1})]\frac{\pi_1}{r}\lambda - (\pi_1 - \pi_j)(\pi_n - \pi_{n+1}).$$

Define λ_j as in (20). If $\lambda > \lambda_j$, $[(\pi_j - \pi_n) - (n-j)(\pi_n - \pi_{n+1})]\frac{\pi_1}{r}\lambda - (\pi_1 - \pi_j)(\pi_n - \pi_{n+1}) > 0$, which implies $\alpha_j(n) > \alpha_j(n+1)$. If $\lambda < \lambda_j$, we have $[(\pi_j - \pi_n) - (n-j)(\pi_n - \pi_{n+1})]\frac{\pi_1}{r}\lambda - (\pi_1 - \pi_j)(\pi_n - \pi_{n+1}) < 0$, which implies $\alpha_j(n) < \alpha_j(n+1)$.

Step 2: We show that λ_j increases with j.

It is straightforward to show that $\lambda_1 = 0$. To see $\{\lambda_j\}$ increases in j, note that

$$\begin{aligned} \lambda_j - \lambda_{j+1} &= \frac{r(\pi_1 - \pi_j)(\pi_n - \pi_{n+1})}{\pi_1[\pi_j - \pi_n - (n-j)(\pi_n - \pi_{n+1})]} - \frac{r(\pi_1 - \pi_{j+1})(\pi_n - \pi_{n+1})}{\pi_1[\pi_{j+1} - \pi_n - (n-j-1)(\pi_n - \pi_{n+1})]} \\ &= \zeta \cdot \{(\pi_1 - \pi_j)[\pi_{j+1} - \pi_n - (n-j-1)(\pi_n - \pi_{n+1})] - (\pi_1 - \pi_{j+1})[\pi_j - \pi_n - (n-j)(\pi_n - \pi_{n+1})]\} \end{aligned}$$

where $\zeta = \frac{r(\pi_n - \pi_{n+1})}{\pi_1[\pi_j - \pi_n - (n-j)(\pi_n - \pi_{n+1})][\pi_{j+1} - \pi_n - (n-j-1)(\pi_n - \pi_{n+1})]} > 0$. Thus,

$$\begin{aligned} & sign\ (\lambda_j - \lambda_{j+1}) \\ =\ & sign\{-(\pi_1 - \pi_j)[\pi_j - \pi_{j+1} - (\pi_n - \pi_{n+1})] - (\pi_j - \pi_{j+1})[(\pi_j - \pi_n) - (n-j)(\pi_n - \pi_{n+1})]\}. \end{aligned}$$

However, $\pi_j - \pi_n - (\pi_{j+1} - \pi_{n+1}) = \pi_j - \pi_{j+1} - (\pi_n - \pi_{n+1}) > 0$ and $(\pi_j - \pi_n) = \pi_j - \pi_{j+1} + \pi_{j+1} - \pi_{j+2} + \cdots + \pi_{n-1} - \pi_n > (n-j)(\pi_{n-1} - \pi_n) > (n-j)(\pi_n - \pi_{n+1})$. Therefore, $sign\ (\lambda_j - \lambda_{j+1}) < 0$. That is, $\{\lambda_j\}$ increases in j.

Step 3: Since λ_j increases in j, as the number of firms increases, for any given λ, there exists a $k(\lambda)$ such that $\alpha_j(n) > \alpha_j(n+1)$ for all $j < k$ and $\alpha_j(n) < \alpha_j(n+1)$ for all $j \geq k$. Define $\tilde{\alpha} = \alpha_k(n+1)$. Then, if $\alpha > \tilde{\alpha}$, we have $m(n) \geq m(n+1)$ which implies $T(n) \geq T(n+1)$. If $\alpha < \tilde{\alpha}$, we have $m(n) < m(n+1)$ which means $T(n) < T(n+1)$. ∎

Proof of Lemma 3

Proof. Suppose that $\hat{\alpha} > \alpha$ but they lead to same equilibrium m. By (14),

$$\begin{aligned} TS(\alpha) - TS(\hat{\alpha}) &= \frac{1}{r}(1-\theta_{n-m+1})(1-\theta_{n-m+2})\cdots(1-\theta_{n-1})[\alpha S_1 + (1-\alpha)S_n] \\ &\quad -\frac{1}{r}(1-\theta_{n-m+1})(1-\theta_{n-m+2})\cdots(1-\theta_{n-1})[\hat{\alpha} S_1 + (1-\hat{\alpha})S_n] \\ &= \frac{1}{r}(1-\theta_{n-m+1})(1-\theta_{n-m+2})\cdots(1-\theta_{n-1})(\alpha-\hat{\alpha})(S_1 - S_n) > 0. \end{aligned}$$

Therefore, total social welfare can be increased by reducing $\hat{\alpha}$ to α. ■

Proof of Proposition 6

Proof. From Proposition 1, firm 1 patents when $\alpha > \alpha_1$. When $\alpha = \alpha_1$, we have

$$TS(\alpha_1) = \frac{1}{r}[\alpha_1 S_1 + (1-\alpha_1)S_n].$$

Next consider $\alpha = \alpha_j$, $j > 1$. Note that $S_1 < S_2 < \cdots < S_n$, thus, $\alpha S_1 + (1-\alpha)S_n < S_n$.

$$\begin{aligned} TS(\alpha_j) &< \frac{1}{r}\theta_{n-1}S_1 + \frac{1}{r}\theta_{n-2}(1-\theta_{n-1})S_n + ... + \frac{1}{r}\theta_{n-m+1}(1-\theta_{n-m+2})\cdots(1-\theta_{n-1})S_n \\ &\quad +\frac{1}{r}(1-\theta_{n-m+1})(1-\theta_{n-m+2})\cdots(1-\theta_{n-1})S_n \\ &= \frac{1}{r}[\theta_{n-1}S_1 + (1-\theta_{n-1})S_n]. \end{aligned}$$

By (2) and $0 < \theta_{n-1} < 1$, $\alpha_1 = \frac{\pi_1 - \pi_n}{\frac{\pi_1}{\theta_{n-1}} - \pi_n} = \frac{\pi_1 - \pi_n}{\pi_1 - \theta_{n-1}\pi_n}\theta_{n-1} < \theta_{n-1}$.Therefore, $TS(\alpha_1) > TS(\alpha_j)$ for $j > 1$.This completes the proof. ■

References

[1] Allison, J. and Lemley, M. "Empirical analysis of the validity of litigated patents". *American Intellectual Property Law Association Quarterly Journal*, 26(3):185-275, (1998).

[2] Anton, J. and Yao, D. "Little patents and big secrets: managing intellectual property", *RAND Journal of Economics*, 35, 1-22, (2004)

[3] Aoki, R. and Spiegel, Y. "Pre-grant patent publication and cumulative innovation" *International Journal of Industrial Organization*, 27, 333-345, (2009)

[4] Bessen, J. and Maskin, E., "Sequential innovation, patents, and imitation" *RAND Journal of Economics,* 40, 611-635, (2009)

[5] Choi, J. "Market structure, incentive to patent and the pace of innovation" *Economics Letters*, 34, 277-283, (1990)

[6] Choi, J. "Patenting litigation as an information transmission mechanism" *American Economic Review*, 88, 1249-1263. (1998)

[7] Cohen, W., Nelson, R., and Walsh, J. "Protecting their intellectual assets: appropriability conditions and why U.S. manufacturing firms patent (or not)" *NBER working paper* W7552. (2000)

[8] Denicolò, V. and Franzoni, L., "Patenting, Secrets, and The First-Inventor Defense" *Journal of Economics and Management Strategy,* 13, 517-538, (2004)

[9] Denicolò, V. and Franzoni, L., "The contract theory of patents" *International Review of Law and Economics*, 23, 365–380, (2004)

[10] Erkal, N. "The decision to patent, cumulative innovation, and optimal policy" *International Journal of Industrial Organization*, 23, 535-562, (2005)

[11] Erkal, N. and Scotchmer, S., "Scarcity of Ideas and R&D Options: Use it, Lose it or Bank it," *NBER Working Paper* No. 14940. (2009)

[12] Farrell, J. and Shapiro, C. "How Strong Are Weak Patents?" *American Economic Review*, 98(4): 1347–69.(2008)

[13] Gallini, T. "Patent policy and costly imitation". *RAND Journal of Economics*, 23 (1):52-63, (1992).

[14] Green, J. and Scotchmer, S., "On the division of profit in sequential innovation", *Rand Journal of Economics,* 26, 20-33, (1995)

[15] Horowitz, A. and Lai, E., "Patent Length and The Rate of Innovation" *International Economic Review*, 37 (4) (1996)

[16] Horstmann, I., MacDonald, G., and Slivinski, A. "Patents as information transfer mechanisms: to patent or (maybe) not to patent" *Journal of Political Economy*, Vol, 93. 837-58.(1985)

[17] Kultti, K., Takalo, T., and Toikka, J. "Simultaneous Model of Innovation, Secrecy, and Patent Policy" *American Economic Review*, 96, 82-86 (2006)

[18] Kultti, K., Takalo, T., and Toikka, J. "Secrecy versus patenting" *RAND Journal of Economics*, 38, 22-42.(2007)

[19] Lemley, M. and Shapiro, C. "Probabilistic Patents" *Journal of Economic Perspectives*, 19, 75-98.(2005)

[20] Lerner, J. "Patenting under the shadow of competitors" *Journal of Law and Economics*, 38, 463-95. (1995)

[21] Levin, R., Klevorick, R., Nelson, R., and Winter, S. "Appropriating the returns from industrial research and development" *Brookings Papers on Economic Activity*, 3, 783-820.(1987)

[22] Loury, G. "Market Structure and Innovation" Quarterly Journal of Economics, 93, 395-410 (1979)

[23] Marshall, E. , "The Patent Game: Raising the Ante" *Science* 253: 20-24 (1991)

[24] Maskus, K., "Reforming US Policy: Getting the Incentives Right", CSR No. 19, *Council on Foreign Relations*, (2006)

[25] Mansfield, E."Patents and innovation: An empirical study". *Management Science*, 32(2):173-181, (1986).

[26] Merges, R. and Duffy, J. *Patent Law and Policy: Cases and Materials*, LexisNexis, (2007)

[27] Pakes, A. and Griliches, Z., "Patents and R & D at the firm level: A first look." *Economics Letters*, 5, no 4, 377-81. (1980)

[28] Reinganum, J. "The Timing of Innovation: Research, Development, and Diffusion" *Handbook of Industrial Organization*, Vol. 1, R. Schmalensee and R.D. Willig eds.(1989)

[29] Scherer, F. , "Firm size, market structure, opportunity, and the output of patented inventions." *American Economic Review*, 55, 1097-1125. (1965)

[30] Scherer, F., "Market structure and employment of scientists and engineers" *American Economic Review*, 57, 542-31. (1967)

[31] Schuett, F., "Inventors and Impostors: An Economic Analysis of Patent Examination" *Working Paper* (2008)

[32] Scotchmer, S. *Innovation and Incentives*. The MIT Press, (2004).

[33] Shapiro, C. "Patent reform: Aligning reward and contribution". *NBER working papers*, National Bureau of Economic Research, Inc, May (2007).

[34] Taylor, H. and Karlin, S., , *An Introduction to Stochastic Modeling*, Academic Press, (1984)

[35] Varian, H., Farrell, J. and Shapiro, C. *The Economics of Information Technology: An Introduction.* Cambridge University Press, (2005).