小特集

Patent Breadth and Patenting of Intermediate Results in Cumulative Innovation Race

Zonglai KOU[¥,*]

Abstract: The paper discusses the effect of patent breadth on the patenting of intermediate results in a cumulative innovation race. Patent breadth is considered as the licensing rate between the patentee of basic innovation and the subsequent innovators. We show that both patent breadth and the strategic first mover advantage due to holding the basic innovation secret determine the final market outcomes. Although increasing patent breadth always encourages public disclosure of intermediate results, the effect of patent breadth on social welfare and technological progress shows a quasi inverted U-shape and thus the optimal patent breadth is an inner solution, sharply different to the existing literature, where the optimal breadth is often corner solution.

Key words: Patent Breadth, Cumulative Innovation, Disclosure.
JEL Classification: L00, O31, O34

1. Introduction

One of the main aims of patent system is to transform private knowledge into public knowledge and hence accelerate technological progress. However, in classical optimal design and patent race literature (Nordhaus, 1969; Loury, 1979; Dasgupta and Stiglitz, 1980; Lee and Wilde, 1980), an innovator discloses its innovation and applies patent protection immediately because it is often assumed that any innovation once finished will become public knowledge. Based on this assumption, the literature derives the optimal patent policy by analyzing the effect of patent protection on social welfare and technological progress. However, empirical studies show that innovators often keep their innovations as trade

secrets (Arundel, 2001, Arundel and Kabla, 1998, Cohen, Nelson and Walsh, 1996, Harabi, 1995, Levin, Klevorick, Nelson and Winter, 1987), which means that innovation per se is not public knowledge; instead whether it is public or private knowledge is determined by firms' rational choice of disclosure. For discrete innovations, which have no subsequent innovations, the public-knowledge assumption is innocuous. However, when innovation is cumulative, whether or not the basic innovation is disclosed becomes crucial because a firm must first accomplish (or acquire it in other way) basic innovation before undertaking subsequent innovations (Green and Scotchmer, 1995). Obviously, public disclosure eliminates wasteful R&D duplication and makes other firms "stand on the shoulders of the giants" to develop subsequent innovations from the viewpoint of social planner. As we show in this paper, however, whether or not the first firm will disclose the basic innovation entails the tradeoff between the costs and benefits doing so, which, in turn, depends on both patent breadth and the characteristics of the innovation process.

In a seminal paper, Green and Scotchmer (1995) point out that the role of patent system in cumulative innovation is to award first-generation patentee some claims to second-generation innovations because the latter are impossible without the former. They view patent breadth as the profit-sharing proportion. However, due to imperfect legal system and asymmetric information, patent breadth in practice is a very vague conception and its determination and implementation depend on discretions of courts. Realizing this, Aoki and Hu (1996) and Aoki and Spiegel (1999) supply a micro-foundation for Green-Scotchmer explanation of patent breadth from the viewpoint of litigation and settlement. They argue that with stronger patent breadth the first-generation innovator will be supported by the court with higher probability in litigation and trial under the doctrine of equivalents and hence the subsequent innovators will be willing to pay a higher licensing rate to get permission to commercialize their products[1].

Following the above literature, we simplify patent breadth in the cumulative innovation as the licensing rate, i.e., that profits-sharing rate between patentees of basic and subsequent innovations. With this definition, we can show intuitively the main considerations that the first firm has concerning the disclosure of the basic innovation in a two stage model and

---

[1] Although economists focus on patent length in the optimal patent design literature traditionally, the "effective patent life" when innovation is cumulative is mainly determined by patent breadth (Merges and Nelson, 1990; O'Donoghue, 1998; O'Donoghue, Scotchmer and Thisse, 1998).

further discuss the effect of patent breadth on the social welfare and the rate of technological progress.

On the one hand, by patenting and commercializing the basic innovation, the first firm can get some interim profits. More importantly, the firm can get at least partial claim to subsequent innovations: due to potential infringement, other firms, if they accomplish the subsequent innovations, should pay licensing fee to the first firm to avoid litigation. Conversely, in a first-to-file patent system, if the first firm keeps the basic innovation as secret but other firms patent it before the first firm accomplishes the subsequent innovation, the first firm should instead pay licensing fee to avoid litigation. Therefore, in cumulative innovation, one important motive for a firm to patent is to pre-empt other firm's patenting (Cohen et al., 2000). However, public disclosure, necessary for patenting, means that other firms can "stand on the shoulders of the giants" to undertake immediately the research of subsequent innovations and thus the probability of the first firm winning subsequent innovation will be decreased.

On the other hand, by keeping basic innovation as secret, the firm enjoys a kind of strategic first-mover advantage, i.e., embarking on subsequent innovations before its opponents, so the firm will have a higher probability to win the subsequent innovations than in the case of disclosure. Moreover, if this advantage is strong enough, other firms may exit the market totally due to too pessimistic expectation of the future[2].

We have shown that the costs and benefits of two alternative strategies "disclose" and "hold" concern mainly the licensing rate and first-mover advantage. Intuitively, the first firm's incentive to disclose will be larger when patent breadth is larger and first-mover advantage is smaller.

Dasgupta and David (1994) once discuss the disclosure of intermediate innovation in a very simple two stage race model. The unique Nash equilibrium entails that each firm holds it secret to get first-mover advantage because they omit the potential cost of paying licensing fee to other firm. Grossman and Shapiro (1987) decompose innovation into research stage and development stage and discuss dynamic R&D competition. Although they also analyze how

---

[2] However, if there is uncertainty of innovation hazard rate, the fact that the firm has completed basic innovation (even though its details are not disclosed) will be good news for other firms and increase their confidence to continue it. After all, "if others can do, why not me?" (Choi, 1991). In this article, we do not consider this case.

the firms patent intermediate results, they focus on how firms' R&D efforts are dynamically adjusted according to their relative positions in innovation process. That is, they do not discuss how their incentives to patent intermediate results change with the degree of patent protection. On the contrary, this is the main aim of this paper.

In the paper, we show that strong patent breadth is helpful to information disclosure. However, the effect of patent breadth on social welfare and technological progress show a quasi-inverted U shape. Indeed, there is an optimal patent breadth zone. Besides, if firms can make ex ante licensing arrangement social welfare will be higher. Finally, we prove that if only two firms have the same expectation of the results of litigation, then no trial occurs and the determination and implementation of patent breadth is costless.

The rest of the paper is arranged as follows: in section 2, we construct a two-stage two-firm cumulative innovation race model, in which subsection 2.1 and 2.2 discuss how the strategic firm-mover advantage and patent breadth affect firm 1 and 2's choices, respectively, subsection 2.3 studies the effect of patent breadth on social welfare and technological progress, subsection 2.4 discusses the effect of the pattern of licensing arrangement between firms, and subsection 2.5 analyzes the determination and implementation of patent breadth. Finally, section 3 concludes.

2. Model

We analyze the patenting of intermediate innovation results in a continuous-time model in the framework of Grossman and Shapiro (1987). A firm needs to accomplish two stages, i.e., "research" and "development" to obtain a commercially useful product. We denote "research" as the first stage or innovation 1 and "development" as the second stage or innovation 2. The results of both "research" and "development" are patentable[3]. Compared to the "final product" from "development", the result of "research" can be seen as an intermediate result. The cumulativeness of innovation entails that a firm can not do "development" unless it has obtained the result of "research" by accomplishing it by himself or by learning it due to the other firm's public disclosure. For simplicity, like Grossman and Shapiro (1987) and Schankerman and Scotchmer (2000), we assume that innovation 1 is a kind of research tool, say a kind of algorithm or some gene fragments, which have no intrinsic

---

[3] In practice, an innovation should satisfy utility, novelty, and non-obviousness to apply patent.

commercial value, but is necessary for the subsequent innovations, say, some applied software or biomedicines. As a result, innovation 2, if accomplished by other firm, may infringe the research tool and can not be commercialized by patent injunction without paying licensing fee to the patentee of innovation 1. Let k denote the licensing rate or patent breadth, which is determined by the social planner through controlling the patent claim scope in application and through the doctrine of equivalents (discussed at the end of the paper). We also assume that licensing arrangement takes place ex post, that is, after innovation 2 has been accomplished. Of course, if both innovations are accomplished by one firm, no licensing problem arises.

Suppose that arrival of innovation idea follows Poisson process and the hazard rate parameters for "research" and "development" are λ and μ respectively, exogenously given and the same for two firms. In each stage, the cost flow of innovation is constant c. A firm incurs no costs any more if it exits[4,5]. Furthermore, to focus on the patenting of intermediate results in cumulative innovation race, we assume, without loss of generality, that initially firm 1 has accomplished innovation 1 (in stage 2) but firm 2 is still in the first stage of "research". Timing is showed in Fig. 1[6].

Firm 1 either discloses innovation 1 to apply patent protection or holds it as secret. In each contingency, firm 2 can either exit or continue the innovation race. When firm 1 discloses, if

---

[4] Poisson process for the arrival of innovation is an ordinary assumption, see Dasgupta and Stiglitz (1980), Loury (1979), Lee and Wilde (1980), O'Donoghue et al (1998). This assumption greatly simplifies analysis. But its main drawback is assume away learning and accumulation of knowledge (Reinganum, 1989).

[5] In the ordinary R&D race models, be innovation cost once for all investment (Dasgupta and Stiglitz, 1980, Loury, 1979) or flow cost (Lee and Wilde, 1980), hazard rate is assumed to be an increasing function of innovation investment. By this assumption, one can analyze the effect of patent race on R&D incentive and social welfare. Along this route, Denicolo (2000) has given an excellent analysis of multistage cumulative innovation race. However, just as pointed out in introduction, the existing literature all omit the issue of disclosure. We assume in this paper exogenous λ and μ to isolate the disclosure effect from incentive effect in multiple innovation race. Moreover, according to analysis hereafter, λ and μ represent the difficulty of "Research" and "Development" respectively and hence determines the strength of strategic first move advantage and their different combinations may characterize different industries.

[6] From the point of view of game theory, which firm first accomplishes prior innovation is chosen by pseudo player Nature. If both the firms have been in stage 1 till time t and neither accomplishes prior innovation in the next duration dt, then according to the memorilessness of Poisson process, the process beginning at $t+$dt is the same as the process beginning at t. Obviously, because neither firm has accomplished innovation 1, there is no problem of disclosure.
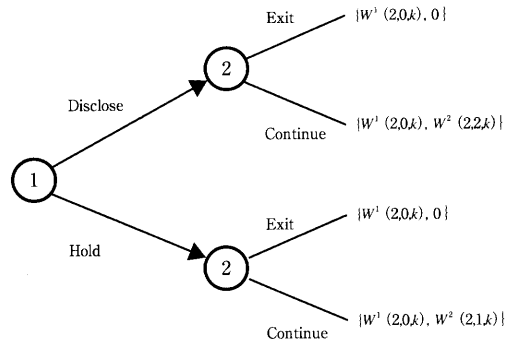
Fig. 1   Game tree

firm 2 exits, only firm 1 stays in the market in stage 2; if firm 2 continues, it can immediately enter stage 2 and start the development of innovation because of public disclosure of innovation 1. When firm 1 holds innovation 1 secret, if firm 2 exits, then only firm 1 stays in the market; if firm 2 continues, it is still in stage 1, undertaking the research of innovation 1[7]. Finally, if both firms are in stage 2, each firm will immediately patent innovation 2 on accomplishing it because of no value of keeping it secret.

For convenience of exposition, we define payoff function $W^i(p,q,k)$, $i = L, F$; $p, q = 0, 1, 2$, where L (Leader) represents the firm possessing the patent of innovation 1, and F (Follower) its rival. Thus, $W^i(p,q,k)$ represents the profits of firm i $(= L, F)$ when firm 1 is in stage p and firm 2 in stage q. Especially, we say that a firm is in stage 0 when it exits. If follower first accomplishes innovation 2, then due to (potential) infringement, it should pay licensing fee to Leader, and k is the ex post licensing rate. Similarly, we define payoff function $W^i(p,q,k)$ as firm j's profits when firm 1 is in stage p and firm 2 in stage q, where $j = 1, 2$; $p, q = 0, 1, 2$. In Fig. 1, the first terms in the brackets denote firm 1's payoffs and the second the firm 2's payoffs.

---

[7] Consider two alternative "secret-keeping" strategies. Strategy A: firm 1 just tells that it has accomplished innovation 1 but does not disclose concrete contents, which is exactly the strategy taken in this paper. Strategy B: firm 1 even keeps secret the news that it has accomplished innovation 1. It's easy to show that A always dominates B. Note that the best result the firm 1 can expect by secret-keeping is exclude firm 2. By telling "I have accomplished innovation 1", firm 1 can make firm 2 having a more pessimistic expectation; however, by taking strategy B, firm 1 can not affect firm 2's expectation of future while still taking the risk of paying licensing fee if firm 2 first patents the basic innovation. Of course, we omit the information effect pointed by Choi (1991). Furthermore, to avoid firm 1 sending false signal, we suppose the cost to verify whether firm 1 has really accomplished prior innovation (but not the contents) is very small.

If a firm invests in innovation 2, the cost flow is c and the total expected social revenue is $V\mu$. Therefore, $V\mu/c$ is the simultaneous social cost-benefits ratio of innovation 2. To make sense, we suppose $\theta = V\mu/c < 1$.

2.1 Firm 2's decision: continue or exit

On completing innovation 1, firm 1 will patent it and become leader by disclosure. Things become somewhat complex if firm 1 holds it secret and firm 2 continues. If firm 1 accomplishes innovation 2 before firm 2 accomplishes innovation 1, then firm 1 will patent both innovation 1 and 2 immediately. But if firm 2 accomplishes innovation 1 before firm 1 accomplishes innovation 2, firm 2 will necessarily patent innovation 1 and become leader instead.

2.1.1 Firm 1 discloses and becomes leader

If firm 2 continues, it immediately enters stage 2, in which case, firm 1's payoff is

$$W^1(2,2,k) = W^L(2,2,k) = \int_0^\infty e^{-(2\mu+r)t} [\mu V - c + \mu k V]\, dt = \frac{\mu V[1-\theta+k]}{2\mu+r}. \tag{1}$$

The explanation of equation (1): ecause both firms are in stage 2 and innovation follows Poisson process with parameter μ, then till time t the probability that no one accomplishes innovation 2 is $\exp(-2\mu t)$. If innovation 2 has not been accomplished, then in the next duration dt, firm 1 has to pay cost $c\,dt$ with probability $\mu\,dt$ accomplishing innovation 2 and getting expected innovation benefits $\mu V\,dt$. Meanwhile, firm 2 also have probability $\mu\,dt$ to accomplish innovation 2, in which case, firm 1 gets licensing revenue $\mu k V\,dt$. Note that r is time preference rate and then by $\exp(-rt)$ we discount the current profits during $[t, t+dt]$ to time zero.

Note that firm 2 has to pay kV (current value) to firm 1 on accomplishing innovation 2, its payoff in this case is

$$W^2(2,2,k) = W^F(2,2,k) = \int_0^\infty e^{-(2\mu+r)t} [\mu(1-k)V - c]\, dt = \frac{\mu V[1-\theta-k]}{2\mu+r}. \tag{2}$$

If firm 2 exits, it gets $W^2(2,0,k) = 0$. Then firm 1's payoff is

$$W^1(2,0,k) = W^L(2,0,k) = \int_0^\infty e^{-(\mu+r)t} [\mu V - c]\, dt = \frac{\mu V[1-\theta]}{\mu+r}. \tag{3}$$

Define $k_0 \equiv 1\text{-}\theta$, we get the following proposition 1.

Proposition 1: Given firm 1 discloses, then (i) if $k > k_0$, then $W^2(2,2,k) < 0$, firm 2 exits and the market outcome is (disclose, exit), with firm 1 getting $W^1(2,0,k)$ and firm 2 zero. (ii) if $k < k_0$, then $W^2(2,2,k) > 0$, firm 2 continues and the market outcome is (disclose, continue), with firm 1 getting $W^1(2,2,k)$ and firm 2 $W^2(2,2,k)$.

2.1.2 Firm 1 holds innovation 1 secret and Leadership is uncertain

First suppose that firm 2 continues. Noting that firm 1 is in stage 2 and firm 2 in stage 1, the asset value of firm 2, i.e., $W^2(2,1,k)$, is determined by Bellman equation

$$rW^2(2,1,k)\,dt = -c\,dt + \lambda dt\,(W^L(2,2,k) - W^2(2,1,k)) + \mu dt\,(0 - W^2(2,1,k)). \tag{4}$$

The left side of the equation represents the rate of return of asset $W^2(2,1,k)$ during $dt\rightarrow 0$. Let's see the right side. The first term $-c\,dt$ is dividends. The second term means that, during dt, firm 2 has probability λdt to accomplish innovation 1 before firm 1 accomplishes innovation 2 and becomes leader in a first-to-file system, in which case, if firm 1 first accomplishes innovation 2 later it should instead pay licensing fee to firm 2 (then leader), therefore, the value asset of firm 2 becomes $W^L(2,2,k)$ but not $W^F(2,2,k)$! The third term just says that, during dt, firm 1 has probability μdt to accomplish innovation 2 and immediately patent both innovations and game is over, i.e., firm 2's asset value decreases to zero. Divided by dt at both side of equation (4), we obtains

$$W^2(2,1,k) = \frac{\lambda W^L(2,2,k) - c}{\lambda + \mu + r} = \frac{\mu V[\lambda(1-\theta+k) - (2\mu + r)\theta]}{(2\mu + r)(\lambda + \mu + r)}. \tag{5}$$

Noting that firm 1's asset value becomes $W^F(2,2,k)$ but not $W^L(2,2,k)$ when firm 2 patents innovation 1, firm 1's payoff $W^1(2,1,k)$ is given by equation

$$rW^1(2,1,k) = -c + \lambda(W^F(2,2,k) - W^1(2,1,k)) + \mu(V - W^1(2,1,k)). \tag{6}$$

Rearranging the above equation, we obtain

$$W^1(2,1,k) = \frac{\lambda W^F(2,2,k) + \mu V - c}{\lambda + \mu + r} = \frac{\mu V[\lambda(1-\theta)(2\mu + \lambda + r) - \lambda k]}{(2\mu + r)(\lambda + \mu + r)}. \tag{7}$$

Defining $m = \dfrac{(2\mu + \lambda + r)\theta}{\lambda} - 1 = (2 + \dfrac{\lambda + r}{\mu})\dfrac{c}{\lambda V} - 1$, we obtain the following proposition.

Proposition 2: Given firm 1 holds innovation 1 secret, then (i) if $k < m$, then $W^2(2,1,k) < 0$, firm 2 exits and market outcome is (hold, exit), with firm 1 getting $W^1(2,0,k)$ and firm 2 gets zero; (ii) if $k > m$, then $W^2(2,1,k) > 0$, firm 2 continues and market outcome is (hold, continue), with firm 1 getting $W^1(2,1,k)$ and firm 2 $W^2(2,1,k) > 0$.

Now we show that m just measures the magnitude of first-mover advantage. By definition, m depends only on pure technical parameters λ, μ and θ, irrelevant to k. Note first that m will be very big, even bigger than 1, if λ is very small. Thus, given patent breadth or licensing rate $k \in [0,1]$, if λ is very small there must be $k < m$ such that, if only firm 1 takes "hold" strategy, firm 2 will exit because of too pessimistic expectation of winning innovation 1 and hence innovation 2 (or getting some licensing revenue). Second, given c/V, m is larger if μ is smaller (of course, still $\theta = c/\mu V < 1$), that is, innovation 1 is more difficult to accomplish, then firm 2 will be more likely to exit due to a more pessimistic expectation.

2.2 Firm 1' decision: "disclose" or "hold"

Back to the beginning of the game, firm 1 compares the corresponding expected revenue of "disclose" and "hold" to take whichever strategy.

Proposition 3: At the beginning of the game, whichever strategy of "disclose" or "hold" firm 1 takes, it is time consistent.

Obviously, if firm 1 "discloses", this strategy is irreversible. Then we need to show that if firm 1 "holds" it will never turn to "disclose" after a while unless it accomplishes innovation 2. Use "X" to denote the event that firm 1 has accomplished innovation 2 and "Y" the event that firm 2 has accomplished innovation 1. If neither X or Y happens during interval $[t_1,t_2]$, then the process thereafter, i.e., starting from t_2, is totally the same as that starting from time t_1 according to memorilessness of Poisson process. By proposition 3, we only need discuss firm 1's choice on its accomplishing innovation 1.

First, note that the best outcome for firm 1 to take "hold" strategy is exclude firm 2 and gets $W^1(2,0,k) = W^L(2,0,k)$. Note further that "hold" is a risky strategy for firm 1 because if firm 2 does not exit there must be some probability that firm 2 will accomplish innovation 1 before firm 1 accomplishes innovation 2, in which case, firm 1's payoff must be less than $W^1(2,0,k)$ in a first-to-file system. However, by proposition 1, if $k > k_0$ firm 2 will exit if only firm 1 discloses. Therefore, the final market outcome is (disclose, exit) when $k > k_0$.

Second, define $k_1 = (1-\theta)\mu/(\mu+r) < k_0$. If $k \in [k_1,k_0]$, firm 1 gets $W^1(2,0,k)$ by "hold" strategy, less than $W^1(2,2,k) = W^L(2,2,k)$, the payoff by "disclose" strategy. Meanwhile, firm

2 does not exit because $W^F(2,2,k) > 0$. Therefore, the final market outcome is (disclose, continue) if $k \in [k_1, k_0]$.

If $k < k_1$, firm 1 has incentive to exclude firm 2 because $W^L(2,2,k) = W^1(2,0,k)$. But by proposition 2 the effectiveness of "hold" strategy depends on k and m. If $k < m$ firm 2 exits if firm 1 "holds" and the final outcome is (hold, exit); If $k > m$ firm 2 ever continues (noting $k < k1$ now). Therefore, firm 1 determines whichever strategy to take by comparing the "disclose" payoff $W^L(2,2,k)$ "hold" payoff $W^1(2,1,k)$. Define further that

$$k_2 = \frac{\mu(1-\theta)}{\mu+r} - \frac{2\lambda}{\mu+r} = k_1 - \frac{2\lambda}{\mu+r} < k_1.$$

By easy computation we can find: Given $k > m$, if $k > k_2$, then $W^1(2,1,k) < W^L(2,2,k)$, i.e., firm 1 can get a higher profit from "disclose", and hence the final outcome is (disclose, continue); However, if $k < k_2$, $W^1(2,1,k) > W^L(2,2,k)$, i.e., firm 1 can get a higher profit from "hold" and hence the final outcome is (hold, continue).

Referring to Fig. 2 through Fig. 5, in which the thick solid lines represent firm 1's payoffs corresponding to different patent breadth k under different conditions (different m), we conclude above analysis as the following proposition.

Proposition 4: When $m > k_1$, the equilibrium is the following: (i) (Disclose, Exit), if $k > k_0$; (ii) (Disclose, Continue), if $k_1 < k < k_0$; (iii) (Hold, Exit), if $k < k_1$.

When $k_2 < m < k_1$, the equilibrium is: (i) (Disclose, Exit), if $k > k_0$; (ii) (Disclose, Continue), if $m < k < k_0$; (iii) (Hold, Exit), if $k < m$.

When $0 < m < k_2$, the equilibrium is: (i) (Disclose, Exit), if $k > k_0$; (ii) (Disclose, Continue) if $m < k < k_0$; (iii) (Hold, Continue), if $m < k < k_2$; (iv) (Hold, Exit), if $0 < k < m$.
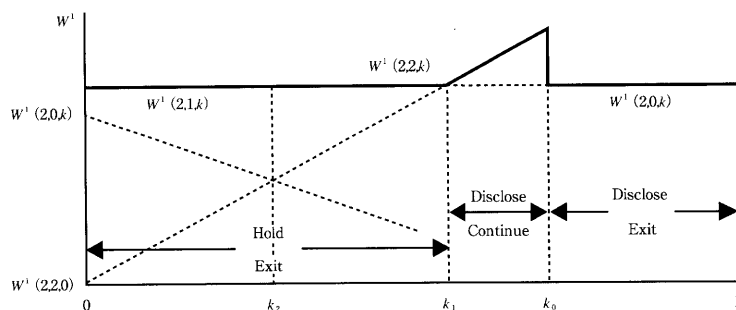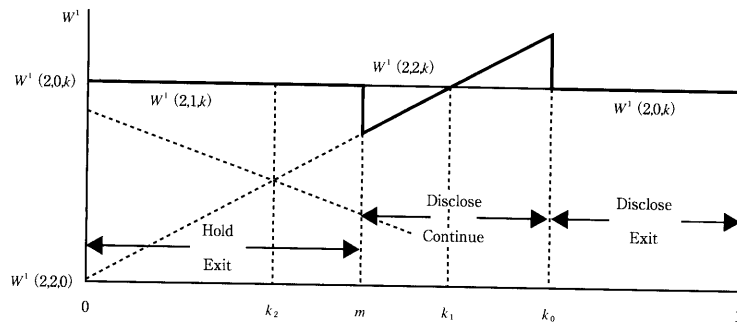

Fig. 2   the case of $m > k_1$
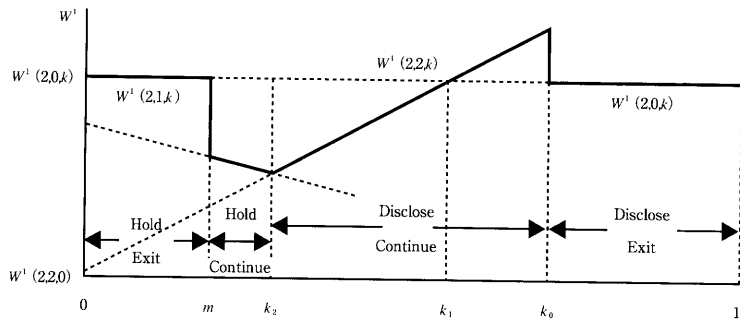
Fig. 3   the case of $k_2 < m < k_1$
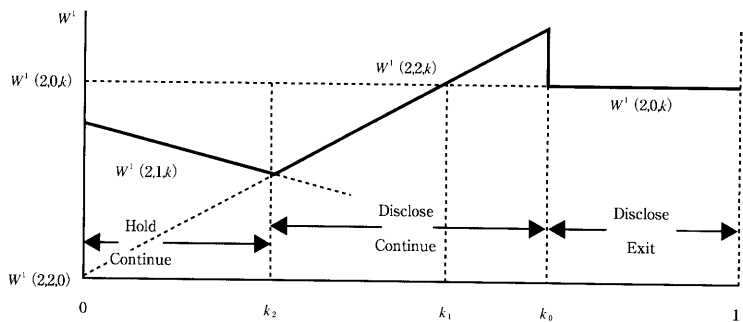

Fig. 4   the case of $m < k_2$

Fig. 5   the case of $m \leq 0$

When $m \leq 0$, the equilibrium is: (i) (Disclose, Exit), if $k > k_0$; (ii) (Disclose, Continue), if $k_2 < m < k_0$; (iii) (Hold, Continue), if $0 < k < k_2$.

2.3 Patent Breadth, Social Welfare and the Pace of Technological Progress

Proposition 4 shows that final outcomes vary with m and k, but larger breadth always encourages firm 1 to discose. In what follows, we discuss the effect of patent breadth on social welfare W. Like Loury (1979), social welfare in this paper is equal to the total expected profits of the two firms because we do not consider consumer surplus. For convenience of

23

exposition, we define the social welfare under (Disclose, Exit) as W^{DE}, (Hold, Exit) as W^{HE}, (Disclose, Continue) as W^{DC}, and (Hold, Continue) as W^{HC}, with superscript D for Disclose, H for Hold, E for Exit and C for Continue. It is easy to show that

$$W^{DE} = W^{HE} = W^1(2,0,k) = \mu V(1-\theta)/(\mu+r), \tag{8}$$

$$W^{DC} = W^1(2,2,k) + W^2(2,2,k) = 2\mu V(1-\theta)/(2\mu+r), \tag{9}$$

$$W^{HC} = W^1(2,1,k) + W^2(2,1,k) = \frac{\mu V[(1-2\theta)(2\mu+r+\lambda)+\lambda]}{(\lambda+\mu+r)(2\mu+r)}. \tag{10}$$

Proposition 5: $W^{DC} > W^{HC}$; $W^{DC} > W^{DE} = W^{HE}$.

Although k does not appear in (8) through (10), its magnitude determines whichever expression for social welfare to take. Although increasing k always promotes disclosure, the effect of k on social welfare shows a quasi inverted-U shape. If k is very small, firm 1 will take "Hold" strategy and social welfare is either W^{HC} if firm 2 continues or W^{HE} if firm 2 exits. If k is large but still less than k_0, firm 1 "Discloses", firm 2 continues, and social welfare is W^{DC}. Finally, if k is too big, i.e., $k > k_0$, firm 2 will exit even if firm 1 discloses, reducing social welfare from W^{DC} to W^{DE}.

Obviously, social welfare is maximized when firm 1 discloses and firm 2 continues. First, because two firms' successes of innovations are independent under Poisson process assumption, the probability that neither firm accomplishes innovation is less. Second, disclosure reduces wasteful replication. If firm 2 continues when firm 1 takes "hold" strategy, then the private expected payoff of firm 2 doing so must be positive. However, from the point of view of social planner, it is wasteful for firm 2 to continue to invest in innovation 1 because this does not increase social knowledge stock given firm 1 has accomplished innovation 1.

Now we begin to study the effect of disclosure on accelerating the speed of technological progress, which is considered as one of the core aims while making and revising patent policies in many countries (Horowitz and Lai. 1996). Define τ_{20} as the average time for society to accomplish innovation 2 when firm 1 is in stage 2 and firm 2 exits and similarly τ_{22} and τ_{21}. According to Poisson process, we obtain

$$\tau_{20} = \int_0^{\infty} te^{-ut}\mu dt = 1/\mu \tag{11}$$

$$\tau_{22} = \int_0^\infty te^{-2ut}2\mu dt = 1/2\mu \tag{12}$$

$$\tau_{21} = \int_0^\infty te^{-ut}\{e^{-\lambda t}\mu + [\int_0^t e^{-\lambda s}\lambda e^{-\mu(t-s)}ds\,]2\mu\}\,dt = \frac{2\mu+\lambda}{2\mu(\mu+\lambda)} \tag{13}$$

Proposition 6: $\tau_{20} > \tau_{21} > \tau_{22}$, i.e., disclosure will accelerate technological progress.

We already know that patent breadth has a quasi-U shaped effect on social welfare and the speed of technological progress. Based on proposition 4 and 5, we can determine the optimal zone of patent breadth, as shown in Fig. 6.
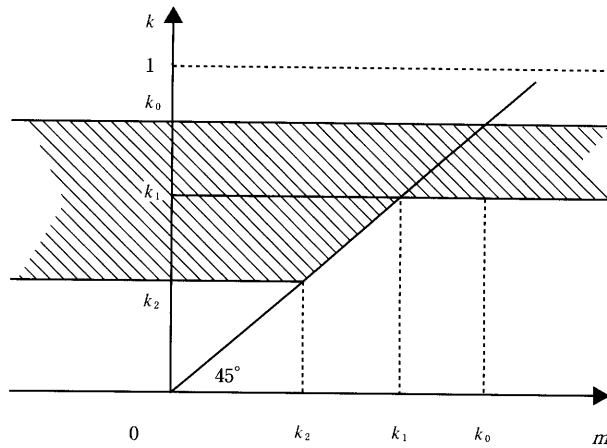

Fig. 6   the optimal patent breadth (zone)

Proposition 7: The optimal zone of patent breadth is:

$[k_1,k_0]$ if $m > k_1$; $[m,k_0]$ if $k_2 < m < k_1$; $[k_2,k_0]$ if $m < k_2$.

There are explicit policy implications for proposition 7. When innovation is cumulative, firms often play multi-stage patent race, where a firm must first complete previous innovations to do subsequent innovations. We show that social welfare is maximized when market outcome is (disclose, continue). Therefore, patent system should encourage firms to disclose their intermediate results on the one hand and keep enough competition on the other hand. To realize this aim, patent breadth should not too small or too big. Moreover, $k^* < k_0 = 1 - c/\mu V$ implies that, in an industry with small μ, i.e., where "development" is very difficult, social planner should set relatively small patent breadth to accommodate lagging firms to increase the total probability of accomplishing innovation 2.

2.4 Ex Ante Licensing and Pareto Improvement

We have just discussed ex post licensing. In practice, there is also ex ante licensing arrangement, in which licensing rate is determined before investment for innovation 2 has been made. Gallini and Scotchmer (1994) pointed out correctly that in the context of cumulative innovation, the pattern of licensing arrangement has very important effect on the final outcome. Moreover, ex ante licensing leads to higher social welfare in ordinary case. In this paper we support their insights from the point of view of public disclosure.

Proposition 8: When $k > k_0$, market outcome is (disclose, exit) with social welfare W^{DE} under ex post licensing. However, if the firms can make ex ante licensing choosing $k^* \in [k_1, k_0]$ is a Pareto improvement: $W^2 (2,2,k^*) \geq 0$ for firm 2 and $W^1 (2,2,k^*) > W^1 (2,0,k)$ for firm 1.

When $k > k_0$ firm 1 would rather choose $k^* \in [k_1, k_0]$ to accommodate firm 2 than exclude firm 2 with licensing rate k. However, under ex post licensing firm 1 can not commit a licensing rate $k < k_0$, because firm 2' investment, once made, will become sunk cost ex post and does not affect the result of litigation and then firm 1 will choose an licensing rate as large as possible to divide the revenue V from innovation 2. Therefore, firm 2 will exit with negative expected profits, leading to the inefficiency outcome (Disclose, exit). This inefficiency just represents the hold-up effect (Shapiro, 2001).

However, if ex ante licensing is possible, firm 1 can set $k^* < k_0 < k$ to accommodate firm 2 and social welfare is improved. Furthermore, this does not violate the antitrust law, by which raising licensing rate is ofren banned because of suspect of collusion between firms and being detrimental to social welfare. Therefore, ex ante licensing (if possible) can enhance social welfare without violating antitrust law in the case of cumulative innovation and should be encouraged to avoid the inefficiency due to excessive patent protection from the point of view of disclosure.

It seems that, by proposition 7, social planner can avoid the inefficiency outcome (disclose, exit) just by setting a patent breadth $k < k_0$. Noting that the degree of patent protection is almost the same across different industries with different combinations of c, V and μ, however, level of k_0 may be too high in one industry but too low in another industry.

2.5 The Determination and Implementation of Patent Breadth

Although we explain patent breadth in cumulative innovation as the licensing rate between sequential innovators from the point of view of litigation and settlement, litigation costs do

not appear in social welfare function at all. Is this pathological? More importantly, our discussion about optimal patent breadth is meaningless unless court or government can control patent breadth.

In practice, the determination and implementation of patent breadth is a very complicated process. For simplicity, we consider it from the following aspects: First, patent breadth increases with the claim scope in patent application. However, a larger claim scope also increases the risk of being refused by the Patent Office during patent application and inspection. In case of refusal, the applicant has to curtail its claim scope to pass the examination. Worse still, by so-called prosecution history estoppel, applicants can not re-claim in the future the parts she has given up during application. Therefore, the claim scope in application letter is virtually controlled by the patent office. Secnd, patent breadth depends on the use of the doctrine of the equivalents. Due to incompleteness of contract, claim scope of patent can not characterize clearly every contingency in the future and thus patent breadth is ambiguous. However, as far as the essentially same function is concerned, if the accused products realize the essentially same results in the essentially same way, then there is infringement under the doctrine of equivalents. Obviously, how the courts define "the essentially same" matters in case of patent litigation. If the court's discretion is pro-plaintiff, paten breadth is enlarbed[8].

Now we simplify the determination and implementation of patent breadth into the following two steps. First, the claim scope in patent application (controlled by patent office) together with the use of the doctrine of equivalents (controlled by the courts) determines the probability of patentee winning in patent litigation. Second, given this probability, the two firms determine to settle or go to trial. Firm i's litigation cost is T_1 and its subjective estimation of the probability of firm 1 winning the litigation is α_i $(i = 1,2)$[9]. Alternatively, they can also settle with licensing rate k (determined by the patent breadth), saving the litigation costs of both parties. For firm 1 to settle, licensing rate must satisfy $\alpha_1 V - T_1 \le$

---

[8] See chandler (2000) for the detail of the relationship between Prosecution History Estoppel, the doctrine of equivalents and patent scope.

[9] For the allocation of litigation cost, the two main rules are American rule of litigation cost allocation and British rule. In the former case, each party affords its own litigation cost regardless of the final outcome. In the latter case, the loser will afford the whole costs of both parties. Here we consider the case of American rule. However, Aoki and Hu (1998) show that changing from American rule to British rule will not affect the final results greatly.

kV, i.e., $k \geq \theta_1 = \alpha_1 - T_1/V$. For firm 2 to settle, licensing rate must satisfy $(1-\alpha_2)V - T_2 \leq (1-k)V$, i.e., $k \leq \theta_2 = \alpha_2 + T_2/V$. Litigation can be avoided if $\theta_2 \geq \theta_1$, which, obviously, is a very loose condition. This explains why most of patent disputes end up in settlement rather than litigation. In the following, we give a sufficient condition for settlement.

Proposition 9: If only firm 1 and firm 2 have the same expected results of litigation, i.e., $\alpha_1 = \alpha_2$, then $\theta_2 \geq \theta_1$ and the final outcome is that no litigation appears. The determination and implementation of patent breadth is costless.

In general case, the empirical study of Lanjouw and Shankerman (2001) shows that the main factors leading to litigation include: (i) parties involved having different subjective estimation of the court judgement. (ii) the difference of each party's litigation cost; (iii) the stake involved. (iv) the benefits of settlement. Their findings coincide well with our explanations: a necessary condition for litigation is $\alpha_1 > \alpha_2$. Furthermore, if T_j/V $(j = 1,2)$, i.e., litigation cost relative to the stake is smaller, litigation will occur more frequently because now the condition $\theta_2 \geq \theta_1$ is more difficult to satisfy.

3. Conclusion

Empirical studies show that many innovations are held as secrets (Cohen et al., 2000; Arundel, 2001), which contradicts to the public knowledge assumption in patent literature. Innovations, once kept as private knowledge, can not be used as inputs by subsequent innovators. We show that one important function of patent system is transform private knowledge into public knowledge by awarding finite monopoly power to innovators.

We construct a two-firm cumulative innovation race model to discuss the effect of patent breadth on social welfare and technological progress. First, we explain patent breadth as licensing rate k between sequential innovator for the point of view of litigation and settlement. Our analysis shows that whether or not the first firm discloses the basic innovation depends on both patent breadth and the magnitude of strategic first mover advantage by holding it secret. If patent breadth is very small, the strategic effect of holding secret dominates because the first firm can only cave small licensing revenue if the other firm first accomplishes the second innovation, or conversely, the firm need not pay too much of licensing fee if the other firm gets the patent of the basic innovation. On the contrary, if patent breadth is very big the first firm would rather disclose the basic innovation to get (or avoid) a large amount of licensing revenues (or fees). Our first basic conclusion is that big

breadth is helpful to disclosure. However, this does not imply the better the bigger patent breadth is. Another important result we find is that patent breadth has a quasi inverted-U-shaped effect on social welfare and the technological progress and thus optimal patent breadth is finite. In contrast, the optimal patent breadth in existing literature is often a corner solution. Gilbert and Shapiro (1990) argues that it be as small as possible. Gallini (1992) reverses Gilbert and Shapiro's conclusion and shows that patent breadth be as large as possible to avoid wastefully costly imitations. In Klemperer (1990) it can be as large or small as possible depending on the concrete situations. Denicolo (1996) concludes above literature, showing that different conclusions result from different definitions of patent breadth and market structure. Undoubtedly, they give us many insights about patent system. However, as mentioned in introduction, all of these papers omit the issue of disclosure to focus on the tradeoff between patent length and breadth. On the contrary, to focus on the issue of public disclosure, we assume that innovation hazard rate λ and μ is irrelevant to flow cost c. Therefore, our analysis is complementary to the above literature.

The paper considers the case of file-to-file system. However, with only a tiny modification of the Bellman equations, our analysis can be easily expanded to the case of first-to-invent system, under which, by intuition, firm's incentive to disclose will be smaller because the corresponding risk of paying licensing fee to the other firm (if it patents the basic innovation) does not exist any more. This coincides with the results of Scotchmer and Green (1990) who compares the incentives of disclosure under the two systems. Furthermore, we can consider the case with interim profits of basic invention. Again, by modification of the Bellman equations we can know that firm 1's incentive to disclose will increase.

References

1) Aoki, R. and J. Hu, 1999, "Licensing vs. Litigation: Effect of the Legal System on Incentives to Innovate", *Journal of Economics and Management Strategy*, 8, 133-160.

2) Aoki, R. and Y. Spiegel, 1999, "Public Disclosure of Patent Applications, R&D, and Welfare", Working paper, The Berglas School of Economics, Tel Aviv University.

3) Arrow, K., 1962, "Economic Welfare and the Allocation of Resources for Invention", is R. Nelson eds. *The Rate and Direction of Inventive Activities: Economic and Social Factors*. Princeton University Press.

4) Arundel, A., 2001. "The Relative Effectiveness of Patents and Secrecy for Appropriation",

Research Policy, 30, 611-624.

5) Chandler, W., 2000, "Prosecution History Estoppel, the Doctrine of Equivalents, and the Scope of Patents", *Harvard Journal of Law & Technology*, 13, No. 3, 466-519.

6) Choi, J., 1991, "Dynamic R&D Competition under 'Hazard Rate' Uncertainty", *RAND Journal of Economics*, 22, 596-610.

7) Cohen, W. M., Nelson, R. R., Walsh, J. P., 2000, "Protecting Their Intellectual Assets: Appropriability Conditions and Why US Manufacturing Firms Patent (or Not), NBER Working Paper No. 7552.

8) Dasgupta, P., and P. David, 1994, "Towards A New Economics of Science", *Research Policy*, 23, 487-521.

9) Dasgupta, P. and J. Stiglitz, 1980b, "Uncertainty, Industrial Structure and the Speed of R&D", *Bell Journal of Economics*, 11, 1-8.

10) Denicolò, V., 1996, "Patent Races and Optimal Patent Breadth and Length", *Journal of Industrial Economics*, 44, 249-266.

11) Denicolò, V., 2000, "Two Stage Races and Patent Policy", *RAND Journal of Economics*, 31, 488-501.

12) Denicolò, V., and P. Zanchettin, 2002, "How Should Forward Patent Protection Be Provided?", *International Journal of Industrial Organization*, 20, 801-827.

13) Fudenberg, D., R. Gilbert, J. Stiglitz and J. Tirole, 1983, "Preemption, Leapfrogging and Competition in patent Paces", *European Economic Review*, 22, 3-31.

14) Gallini, 1992, "Patent Policy and Costly Imitation", *RAND Journal of Economics*, 23, 52-63.

15) Gallini, N., 1984, "Deterrence by Market Sharing: A Strategic Incentive for Licensing", *American Economic Review*, 74, 931-941.

16) Gallini, N. and S. Scotchmer, 2001, "Intellectual Property: When is it the Best Incentive Mechanism?", *Innovation Policy and the Economy*, Vol. 2, A. Jaffe, J. Lerner and S. Stern, MIT Press, 51-78.

17) Gilbert, R., and D. Newberry, 1982, "Preemptive Patenting and the Persistence of Monopoly", *American Economic Review*, 72, 514-26.

18) Gilbert, R. and C. Shapiro, 1990, "Optimal Patent Length and Breadth", *RAND Journal of Economics*, 21, 106-112.

19) Green, J. and S. Scotchmer, 1995, "On the Division of Profit in Sequential Innovation", *RAND Journal of Economics*, 26, 20-33.

20) Grossman, G. and C. Shapiro, 1987, "Dynamic R&D Competition", *Economic Journal*, 97, 372-387.

21) Horowitz, A. and E. Lai, 1996, "Patent Length and Rate of Innovation", *International Economic Review*, 37, 785-601.

22) Klemperer, P., 1990, "How Broad Should the Scope of Patent Protection be?" *RAND Journal of*

Economics, Vol., 21, 113-130.

23) Lanjouw, J. and M. Shankerman, 2001, "Characteristics of Patent Litigation: A Window on Competition", *RAND Journal of Economics*, 32, 129-151.

24) Lee, T. and L. Wilde, 1980, "Market Structure and Innovation: A Reformulation", *Quarterly Journal of Economics* 94, 429-436.

25) Loury, G., 1979, "Market Structure and Innovation", *Quarterly Journal of Economics* 93, 395-410.

26) Machlup, F. and E. Penrose, 1950, "The Patent Controversy in the Nineteenth Century", *Journal of Economic History*, 10, 1-29.

27) Merges, R. and R. Nelson, 1990, "On the Complex Economics of Patent Scope", *Columbia Law Review*, 90, 839-916.

28) O'Donoghue, T., 1998, "A Patentability Requirement for Sequential Innovation", *RAND Journal of Economics*, 29, 654-67.

29) O'Donoghue, T. S. Scotchmer and J. Thisse, 1998, "Patent Breadth, Patent Life, and the Pace of Technological Progress", *Journal of Economics and Management Strategy*, 71-32.

30) Reinganum, J., 1989, "The Timing of Innovation: Research, Development, and Diffusion", in R. Schmalensee and R. D. Willig eds. *Handbook of Industrial Organization*, Vol. 1, Amsterdam: The Netherlands: Elsevier.

31) Scotchmer, S. and J. Green, 1990, "Novelty and Disclosure in Patent Law", *RAND Journal of Economics*, 21, 131-146.

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

33) Shapiro, C., 2001, "Antitrust Limit to Patent Settlements", Working Paper, UC Berkeley, http://www.law.berkeley.edu/institutes/bclt/pubs/wp/501.pdf.