Patent Examination and Disguised Protection

Fei YU and Yanrui WU

Economics (M251), Business School, University of Western Australia,
35 Stirling Highway, Crawley, WA 6009, Australia

Abstract: Could national patent offices use patent examination procedures as a form of protection of local industries? To explore this question, we propose a game theory model in which a foreign multinational corporation (MNC) and a domestic firm compete in the domestic market. In this model the domestic patent office can influence the profit curves of the two firms by controlling the pendency or grant probability of the MNC’s patents. Hence patent examination could be used as a tool to protect the domestic firm and help it to catch up and even leapfrog ahead technologically. We then conduct numerical simulations to identify potential features of such protection and establish hypotheses for empirical testing. With patent data from the world’s major patent offices, our empirical findings show that the Japanese, Korean and Chinese patent offices are linked with the practice of disguised protection. In contrast, there is no clear evidence of protection associated with the patent offices in the US, the UK and Germany.

Key words: Patent examination, disguised protection and technological leapfrogging

JEL Classification: D43, L13, O31, O34
1. Introduction

Overt discrimination against foreign inventors used to be stated clearly in patent laws. For example, according to the United States Patent Act in 1793 the patent rights were only granted to US citizens. The Act was amended in 1800 to allow the ownership of patents by foreigners who had been residents in the United States for two years. Though this restriction on foreigners was removed in the 1836 Act, British applicants were still charged an application fee of $500 each and other foreigners $300 each in comparison with only $30 each for US citizens.

Today, such overt discrimination has been removed from the patent laws in most nations. However, similar practices still exist in the form of lower grant probability and longer grant lag toward foreign applications (Kotabe, 1992; Wineberg, 1988; Linck and McGarry, 1993). A recent example is the patent war between Apple and Samsung ignited in April 2011. Among the seven utility patents that Apple accused Samsung for patent infringement, only one got successfully registered in the Korean patent office.¹ Samsung struck back with a suit against Apple for six patent infringements.² Although these six patents were all successfully registered in the US patent office, they were examined on average for 4.5 years. In contrast, on average Apple’s seven patents only waited for 3.8 years in the US patent office, eight months shorter than Samsung’s. In the electronic industry, eight months can be the life of a generation. For example, iPhone 3GS only dominated in the market for one year before iPhone 4 got released. Are these

¹ The US patent numbers for those seven patents are 7812828, 7669134, 6493002, 7469381, 7844915, 7853891 and 7863533.
² The US patent numbers for those six patents are 7668563, 6937700, 7835729, 6920602, 7050410 and 6882636.
variations in patent examination due to discriminatory policies against foreign patents or disguised protection of domestic firms?

The existing literature has focused on the design of patent regimes so as to provide the best mechanism for patent protection (Helpman, 1993; Gould and Gruben, 1996; Grossman and Lai, 2004). It is generally assumed that foreign patents receive national treatment. However, it is argued that even under a weak patent regime with no discrimination, imitations may occur when the imitation cost is low and thus imitation itself is profitable (Mansfield et al., 1981). Nonetheless, imitation will not occur if it requires substantial R&D investment or incurs large uncertainty. In this case discriminatory policies may be employed to protect domestic inventions or imitations.

The Agreement on Trade Related Aspects of Intellectual Property Rights (TRIPS) established in 1994 explicitly prohibits any discrimination. However, empirical analysis suggests that discrimination still persists (Webster et al., 2007). Although discriminatory policies in international trade have been widely documented, we have found up to now no studies in the field of patent protection. This paper, therefore, tries to unveil the mechanism underlying discriminatory policies in patent examination procedures. It also demonstrates how discriminatory practices have evolved since TRIPS were introduced.

---

3 Article 27.1 requires that “patents shall be available and patent rights enjoyable without discrimination as to the place of invention, the field of technology and whether products are imported or locally produced”.

4 For the trade literature, the readers may refer to Bhattacharjea, 1995; Brander, 1995; Dixit 1984; Kovac and Zigic, 2007; Lee, 2007; Trebilcock and Giri, 2004; and Zigic, 2000.
The reminder of this paper is organized as follows. We first propose a theoretical model in Section 2 and conduct numerical simulations to form hypotheses in Section 3. We then discuss the empirical models and data issues in Section 4. The regression results are presented in Section 5. Section 6 concludes.

2. The theoretical Model

Two firms are assumed to play a three-stage dynamic game. Firm 1 is a foreign multinational corporation (MNC) and Firm 2 is local. They compete with each other in Firm 2’s domestic market. In stage 1, Firm 1 introduces a new invention which will produce a higher quality \( s_1 \) than the existing products. To protect its invention, Firm 1 lodges a patent application for its invention in the host country. However, under the patent assessment system the application is automatically published in the public domain usually in 18 months after lodgment regardless whether the patent is eventually granted or rejected. During this period (in Japan it is six years on average), Firm 1 has no legal IPR to fight against Firm 2’s potential imitation until the patent is granted. Therefore in stage 2 Firm 2 decides whether to imitate Firm 1’s products and chooses its own quality \( s_2 \) if it enters the market. In the last stage, Firms 1 and 2 compete in prices \( p_1 \) and \( p_2 \).\(^5\)

Firm 2’s payoff function is its expected profit. There is risk to imitate during the examination period of Firm 1’s patent, because Firm 1 is possible to sue Firm 2 for patent

\(^5\) Motta (1993) argues firms differentiate more under Bertrand than under Cournot. His paper also shows the economy is better off when firms compete on prices. Enormous literature has contributed to the comparison of Bertrand and Cournot competitions in the context of product differentiation (Lambertini and Mantovani, 2010; Naimzada and Tramontana, 2012; Tremblay and Tremblay, 2011). Since it is not the theme of our paper, we assume the two firms are in Bertrand competition.
infringement during that period if the patent is granted. However the risk also bears with super profits. Suppose Firm 2 is the only firm that takes the risk to imitate Firm 1’s new product. Hence during the examination period \((T_\text{years})\), a dual oligopolistic market structure appears. However, once the patent is rejected or the patent is expired, more firms enter and thus the profit of producing that product is reduced to zero under perfect competition. We denote Firm 2’s oligopolistic profit as \(\Pi^O_2\). We assume that the production is costless and Firm 2 spends a fixed R&D expenditure \((C_2)\) to absorb Firm 1’s invention \((0 \leq s_2 \leq s_1)\) or even to develop a higher quality product \((s_2 \geq s_1 \geq 0)\). If the patent is rejected, then Firm 2 can keep its profit. If the patent is granted, Firm 1 can claim for patent infringement and hence get compensation \((F)\). If \(P\) is the probability of grant, then Firm 2’s expected profit can be written as

\[
E(\Pi_2) = [(1 - P)\Pi^O_2 + P(\Pi^O_2 - F)] - C_2
\]

Correspondingly, Firm 1’s profit also depends on whether its patent is granted or not. Due to the competition from Firm 2, Firm 1 only obtains the oligopolistic profit \((\Pi^O_1)\) during the examination period. If the patent is rejected, even the oligopolistic profit will disappear because other firms can imitate without risk. However, if the patent is granted,

---

\(\text{6}\) For example, under the US patent law “inventors can obtain reasonable royalties from others who make, use, sell, or import the invention during the period between the time the patent application is published and the patent is granted”. (http://www.uspto.gov/news/pr/2000/00-72.jsp)

\(\text{7}\) It is possible for more than one local firm to join in the competition during the patent examination period. The number of local firms entered is affected by the strictness of IPR protection and local firms’ capability in imitation and innovation. Thus the assumption of only one local firm’s entrance reflects a relatively strong IPR protection and weak capability of local firms. However, we will show even under this assumption the local firm can still leap-frog ahead if disguised protection exists.
Firm 1 can enjoy a monopolistic profit usually for 20 years ($\Pi_1^{M20}$). Hence, Firm 1’s expected profit can be written as

\[ E(\Pi_1) = [(1-P)\Pi_1^Q + P(\Pi_1^Q + \Pi_1^{M20} + F)] - C_1 \]

We derive a R&D cost function that is consistent with the knowledge production functions discussed in Griliches (1979), Färe (1974), Furmana et al. (2002) and Suliman (1997). Essentially, the cost function reflects the intuition that the marginal product of R&D investment is diminishing because the progress of new technologies is constrained by the existing knowledge stock. Thus Firm 1’s R&D cost function can be written as (see appendix A for the proof)

\[ C_1(s_i) = \frac{\alpha}{n} \frac{s_i}{1-s_i}, \quad (0 \leq s_i < 1) \]

where $n$ is the number of markets/countries that Firm 1 enters. Since Firm 1 is a multinational corporation, its R&D cost can be shared by its subsidiaries worldwide. $\alpha$ is a scaling parameter to capture industrial characteristics. A larger $\alpha$ indicates that a larger proportion of earnings is invested in R&D in an industry. The denominator ($1-s_i$) ensures the property of diminishing marginal product of R&D investment.

Correspondingly, Firm 2’s R&D cost function can be written as (see appendix A for the algebraic proof)

---

\[ \text{footnote}6 \text{ Again this happens under a strong IPR protection regime. In a weaker regime, local firms may stay in the market during the 20-year protection term. However, we will show later that the strategic patent policy could work even under strong IPR protection regimes (also see footnote 6).} \]
\[
C_2(s_2 \mid s_1) = \frac{\alpha(s_2 - \gamma s_1)}{1 + \gamma s_1 - s_2}, \quad (\gamma s_1 \leq s_2 < 1 + \gamma s_1)
\]

where \( C_2 \) is not only a function of \( s_2 \) but also related to \( s_1 \). It implies Firm 2’s R&D activities benefit from Firm 1’s published knowledge \((s_1)\). \( \gamma \) reflects Firm 2’s capability of absorbing \( s_1 \) and hence is called the absorptive capacity.

Once a patent infringement is detected and a lawsuit is lodged in the court, the court shall decide the compensation for the damage \((F)\). Different countries follow different rules on determining the compensation. According to Coury (2003), the compensation in the US, the UK and France is calculated on the basis of the patentees’ lost profits caused by the infringers. However, the punishment in Italy, Canada, Japan and Germany is based on either patentees’ lost profits or infringers’ illegal profits. This system is also adopted in China and South Korea.\(^9\) Coury argues that in practice it is more common and easier to calculate the fine on the basis of infringers’ illegal profits. In addition, if the infringers do not have enough assets to cover the patentees’ losses, then the executed compensation has to be reduced. Therefore, in this paper we assume that the amount of compensation is a function of infringers’ illegal profit. Symbolically,

\[
F = \mu \beta \text{Max}\left(0, \Pi^0_2 - C_2\right)
\]

---

\(^9\) See Article 65 of the Chinese Patent Law and Article 128 of the Korean Patent Act
http://english.sipo.gov.cn/laws/lawsregulations/201101/t20110119_566244.html
where $\beta$ is the degree of punishment for patent infringement and $\mu$ is the degree of infringement. A higher degree of imitation may incur a higher probability of litigation.

We use the quality distance to measure the degree of infringement, that is, $\mu = 1 - |s_1 - s_2|$

Both Firm 2’s profit ($\Pi_2^O$) and Firm 1’s profits ($\Pi_1^O$ and $\Pi_1^{M20}$) depend on the length of examination ($T$) for Firm 1’s patent. Thus they are integrated values of annual profits, $\pi_2^O$, $\pi_1^O$ and $\pi_1^{M}$, respectively. If $\rho$ is the annual discount rate, we can write

$$\Pi_2^O = \frac{1}{\rho} \int_0^T e^{-\rho t} \pi_2^O dt = \frac{1-e^{-\rho T}}{\rho} \pi_2^O,$$

$$\Pi_1^O = \frac{1}{\rho} \int_0^T e^{-\rho t} \pi_1^O dt = \frac{1-e^{-\rho T}}{\rho} \pi_1^O$$ and

$$\Pi_1^{M20} = \frac{1}{\rho} \int_T^{T+20} e^{-\rho t} \pi_1^M dt = \frac{e^{-\rho T} (1-e^{-20\rho})}{\rho} \pi_1^M$$

In stage 3 of the game, Firms 1 and 2 compete with each other and receive oligopolistic profits $\pi_1^O = p_1 x_1(p_1, p_2)$ and $\pi_2^O = p_2 x_2(p_1, p_2)$, where $p$ and $q$ stand for price and quantity, respectively. The derivation of the demand functions $q_1(p_1, p_2)$ and $q_2(p_1, p_2)$ using the standard consumer utility functions for differentiated products is presented in appendix B (Li and Song, 2009; Motta, 1993). If Firm 2 produces a lower quality ($s_1 > s_2$), then the demand functions are

$$q_1 = 1 - \frac{p_1 - p_2}{s_1 - s_2}$$ and

$$q_2 = \frac{p_1 - p_2}{s_1 - s_2} - \frac{p_2}{s_2}$$
Therefore, the Bertrand duopoly solutions of profit maximization are (see appendix B for the algebraic proof)

\[
\begin{align*}
(11) \quad p_1^* &= \frac{2s_1^2 - 2s_1s_2}{4s_1 - s_2} \quad \text{and} \\
(12) \quad p_2^* &= \frac{s_1s_2 - s_2^2}{4s_1 - s_2}
\end{align*}
\]

The corresponding profits are

\[
\begin{align*}
(13) \quad \pi_1'^* &= \frac{4(s_1 - s_2)s_1^2}{(4s_1 - s_2)^2} \quad \text{and} \\
(14) \quad \pi_2'^* &= \frac{s_1s_2(s_1 - s_2)}{(4s_1 - s_2)^2}
\end{align*}
\]

If Firm 2 produces a higher quality \((s_2 > s_1)\), the demand functions are changed to (see appendix B):

\[
\begin{align*}
(15) \quad \bar{q}_2 &= 1 - \frac{p_2 - p_1}{s_2 - s_1}, \quad \text{and} \\
(16) \quad \bar{q}_1 &= \frac{p_2 - p_1}{s_2 - s_1} - \frac{p_1}{s_1}
\end{align*}
\]

Hence, the annual profits also change to

\[
\begin{align*}
(17) \quad \bar{\pi}_1'^* &= \frac{s_1s_2(s_2 - s_1)}{(4s_2 - s_1)^2} \quad \text{and} \\
(18) \quad \bar{\pi}_2'^* &= \frac{4(s_2 - s_1)s_2^2}{(4s_2 - s_1)^2}
\end{align*}
\]
Appendix B shows that if firm 1’s patent is granted and hence protected for 20 years, the demand for Firm 1’s product is (see appendix B)

\[ q_i = 1 - \frac{P_i}{s_i} \]  

In this case, Firm 1 receives a monopolistic profit

\[ \pi^M_1 = \frac{s_1}{4} \]  

By integrating the above results into the expected profit functions (1) and (4), we obtain Equations (20) and (21) in the case of \( s_2 < s_1 \):

\[ E(\Pi_2) = [(1-P)\Pi_2^0 + P(\Pi_2^0 - F)] - C_2 = \Pi_2^0 - P \times F - C_2 = \frac{1 - e^{-\rho T}}{\rho} s_1 s_2 (s_1 - s_2) (4s_1 - s_2) \]

\[ -P \times (1 - s_1 + s_2) \beta \max \left\{ 0, \frac{1 - e^{-\rho T} s_1 s_2 (s_1 - s_2)}{\rho (4s_1 - s_2)^2} - \frac{\alpha(s_2 - \gamma s_1)}{1 + \gamma s_1 - s_2} \right\} \]

\[ \pi^M_1 = \frac{s_1}{4} \]  

In the case of \( s_2 > s_1 \), the expected profits are

\[ E(\Pi_1) = [(1-P)\Pi_1^0 + P(\Pi_1^0 + \Pi_1^{M20} + F)] - C_1 = \Pi_1^0 + P(\Pi_1^{M20} + F) - C_1 \]

\[ = \frac{1 - e^{-\rho T}}{\rho} 4(s_1 - s_2) s_1^2 + \frac{P e^{-\rho T} (1 - e^{-20\rho T}) s_1}{4} \]

\[ -P \times (1 - s_1 + s_2) \beta \max \left\{ 0, \frac{1 - e^{-\rho T} s_1 s_2 (s_1 - s_2)}{\rho (4s_1 - s_2)^2} - \frac{\alpha(s_2 - \gamma s_1)}{1 + \gamma s_1 - s_2} \right\} - \frac{1}{n} \frac{\alpha s_1}{1 - s_1} \]

and

\[ E(\Pi_2) = \Pi_2^0 - P \times F - C_2 = \frac{1 - e^{-\rho T} 4(s_2 - s_1) s_2^2}{\rho (4s_2 - s_1)^2} \]

\[ -P \times (1 - s_2 + s_1) \beta \max \left\{ 0, \frac{1 - e^{-\rho T} 4(s_2 - s_1) s_2^2}{\rho (4s_2 - s_1)^2} - \frac{\alpha(s_2 - \gamma s_1)}{1 + \gamma s_1 - s_2} \right\} \alpha(s_2 - \gamma s_1) \]

\[ \frac{1}{1 + \gamma s_1 - s_2} \]

and
In the case of \( s_2=s_1 \), Equations (20) and (22) become identical and so do Equations (21) and (23). In this situation the game shrinks to the identical-quality Bertrand competition.

It will not occur according to proposition 1.

**Proposition 1**: The case in which \( s_2 \) equals \( s_1 \) is not an equilibrium solution.

**Proof**: In the case of \( s_2=s_1 \), Equations (20) and (22) give the same profit function for Firm 2:

\[
E(\Pi_2) = 0 - P \times \beta \max \left( 0, \frac{\alpha(s_2 - s_1)}{1 + \gamma s_1 - s_2} - \frac{\alpha(s_2 - s_1)}{1 + \gamma s_1 - s_2} \right) \leq 0
\]

Thus if Firm 2 chooses the same quality as \( s_1 \), it will incur a loss (that is negative profit). The choice of not entering the market would be a better strategy for Firm 2 than entering the market and choosing \( s_2=s_1 \). Therefore, the case of \( s_2=s_1 \) will not occur as an equilibrium.

In stage 2, Firm 2 sets its quality to maximize its profit. Thus Firm 2’s maximization problem is

\[
\text{Max} \left[ E(\Pi_2) \right], \text{ subject to } 0 \leq \gamma s_1 \leq s_2 < s_1 < 1 \text{ or }
\]

\[
\text{Max} \left[ E(\bar{\Pi}_2) \right], \text{ subject to } 0 \leq s_1 < s_2 < 1 + \gamma s_1
\]
Solutions to these maximization problems depend on both $s_2$ and $s_1$. In stage 1 Firm 1 can calculate Firm 2’s best response curve and decides its best quality level ($s_1^*$), then Firm 2 can compare $\text{Max} \left[ E(\Pi_2) \right]_{s_1 = s_1^*}$ and $\text{Max} \left[ E(\Pi_2) \right]_{s_1 = s_1^*}$ and decide whether or not to produce a better quality (leap-frogging). Proposition 2 ensures the existence of an equilibrium.

**Proposition 2**: A pure Nash equilibrium exists in this game.

**Proof**: The strategic delegation problem is a finite game of perfect information, i.e. Firms 1 and 2 know all parameters and the structure of the game (and each firm knows what the other firm knows). Then by Zermelo’s Theorem a Nash equilibrium exists (See Schwalbe and Walker (2001) for Zermelo’s Theorem).

3. **Simulation results**

There are seven parameters in the maximization problem, namely, $\rho$, $\alpha$, $\beta$, $\gamma$, $n$, $T$ and $P$, and hence it is impossible to solve the problem in a closed form. Moreover, the compensation function, Equation (5), is non-differentiable, which makes it impossible to solve the maximization problem algebraically. Therefore, we try to solve it with numerical simulations.

To start, the seven parameters are assigned with initial values. We follow the suggestion of Weitzman (2001) to set the annual discount rate ($\rho$) to be 0.04. $\beta$ reflects the strictness of the patent law. If $\beta > 1$, not only the loss due to infringement is covered, but
also the infringer is punished with a fine. However, punitive damages are not generally available for patent infringement among the G7 countries. In fact, Coury (2003) points out that the United States stands alone for awarding punitive damages. In some countries such as Italy and Japan, such awards have typically been low. Therefore we set $\beta=1$ as the initial value. $\gamma$ reflects the capability of absorbing the existing knowledge stock and falls into interval $[0, 1]$. Thus we first use 0.5 as the initial value for $\gamma$. $n$ is the number of patent offices worldwide with which a patent is lodged. Firm 1 lodges at least two patent applications for its invention, one in its home country and the other in Firm 2’s home country. Thus we set the initial value of $n$ to be 2. $\alpha$ is the ratio of R&D expenditures over profits in a sector and it varies across industries. We set its initial value to be 0.1 (see more detail in appendix A). Finally, for the grant probability ($P$) and grant lag ($T$) we use the US patent office practice as the baseline and set the initial values to be close to their means, 0.5 and 2 years, respectively.

Figure 1 shows the simulation results of the baseline setting in which the parameters \{ $\rho$, $\alpha$, $\beta$, $\gamma$, $n$, $T$, $P$ \} are fixed at \{0.04, 0.1, 1, 0.5, 2, 2, 0.5\}\textsuperscript{10}. Figure 1(a) illustrates Firm 2’s best response quality ($s_2$) as $s_1$ varies from zero to one. $s_2^*(s_1)$ is Firm 2’s best response curve when Firm 2’s quality is inferior to Firm 1’s quality. $\tilde{s}_2^*(s_1)$ is Firm 2’s best response curve when leap-frogging occurs. Figure 1(b) gives out profits curves for $\Pi_1(s_1|s_2^*)$, $\tilde{\Pi}_1(s_1|\tilde{s}_2^*)$, $\Pi_2(s_2^*|s_1)$ and $\tilde{\Pi}_1(\tilde{s}_2^*|s_1)$. In the case of $s_2<s_1$, Firm 1 will achieve its maximum profit 1.43 by setting $s_1$ to 0.83. The follower Firm 2 will in turn set $s_2$ to 0.18

\textsuperscript{10} Matlab is used for the calculation. The program script is available upon request.
and obtain a profit of 0.03. In the case of \( s_2 > s_1 \), Firm 1 receives a smaller maximum profit 1.11 \((\tilde{s}_1^* = 0.82)\). Firm 2 will receive negative profit (-0.01) and thus choose not to enter the market. Therefore, there is a unique equilibrium under the condition \( s_1 = 0.83 > s_2 \). In this case, Firm 1 takes the first-mover advantage and doesn’t allow Firm 2 to pursue leap-frogging.

**Figure 1** simulation results for \{ρ, α, β, γ, n, T, P\}={0.04, 0.1, 1, 0.5, 2, 2, 0.5}.

However, the patent authority can help Firm 2 to leap-frog by prolonging Firm 1’s patent examination pendency. Figure 2 shows that when the grant lag is extended to six years, then leap-frogging will occur. Even though Firm 1 may obtain its maximum profit 1.74 when \( s_1 \) is set to 0.85, Firm 2 will not bear with that. Instead, Firm 2 will always choose a higher quality than Firm 1, which will entitle Firm 2 a higher profit. Therefore, by taking Firm 2’s later choice into consideration, Firm 1 will realize in the beginning that the profit 1.74 is not feasible. Hence it will accept the second best result (1.13) by setting \( s_1 \) to 0.80. Firm 2 then enters the market, chooses leap-frogging \((s_2 = 1.17 > s_1)\) and obtains profit 0.26. Therefore, the equilibrium brings about leap-frogging.
Figure 2 simulation results for \( \{ \rho, \alpha, \beta, \gamma, n, T, P \} = \{0.04, 0.1, 1, 0.5, 2, 6, 0.5\} \).

This example shows that a patent office can promote technological leap-frogging by prolonging foreign patent examination pendency. Later we will illustrate that both grant lag and grant probability can be used for stimulating leap-frogging. If in a patent office all foreign patents are discriminated in this way, we say this patent office practice institutional discrimination. However, to evade international anti-discrimination protocols (such as TRIPS), a country can turn from institutional discrimination to more disguised strategic discrimination, which doesn’t apply to all foreign patents but is targeted at key patents or patent fields. In other words, this strategy will spare less important patents and “snipe” at high ranked ones. The aim of strategic discrimination, along with other industrial policies (e.g. R&D subsidies, see Spencer and Brander, 1983), is to promote technological catching-up or leap-frogging in these key patent fields. In the real world, do patent offices really strategically discriminate against foreign patent? To answer this question, we need to bring our model into empirical tests with real data. Therefore, in the following discussions hypotheses of strategic discrimination are developed for real-data examination.
In our model, whether leap-frogging can occur or not is determined by the seven parameters $\rho, \alpha, \beta, \gamma, n, T$ and $P$. However, patent offices can regulate only on two of them, the grant lag ($T$) and grant probability ($P$). Thus we will focus on what the threshold values of $T$ and $P$ are needed for leap-frogging when other parameters vary.

Figure 3(a) depicts the simulation result when $\{\rho, \alpha, \beta, \gamma, n\} = \{0.04, 0.1, 1, 0.5, 2\}$. Leap-frogging will occur in the shadowed domain for $P$ and $T$. If $P$ and $T$ do not reach the threshold boundaries, the domestic firm (Firm 2) will not be motivated enough to pursue for leap-frogging. In fact, the boundary curve is the only Nash equilibrium solution, because excessive discrimination will draw the other countries’ retaliation as it has happened in international trade. Thus the ideal strategic behavior is to protect domestic firms by just choosing the threshold boundary level of discrimination. Any combination of $T$ and $P$ along that boundary curve will be the Nash equilibrium solution.

Thus this boundary curve can be named as the iso-equilibrium curve.

**Figure 3** simulation results when $\{\rho, \alpha, \beta, n\} = \{0.04, 0.1, 0.5, 2\}$.

Figure 3(b) illustrates how the iso-equilibrium curve shifts as the degree of punishment for patent infringement ($\beta$) varies from 0 to 1. A significant divergence appears after $P$.
exceeds 0.5. In the real world, $\beta$ is independently determined by the court rather than the patent office. Thus it is taken granted by patent offices before any discriminative polices are orchestrated. Figure 3(b) suggests that in a stricter anti-infringement law system, severer discriminative tools are needed for leap-frogging. Thus we obtain our first hypothesis.

**H1:** In a regime where the punishment for patent infringement ($\beta$) is severe,

a lower $P$ or longer $T$ is needed for the realization of leap-frogging.

$\alpha$ and $\gamma$ mainly reflect industrial characteristics. $\alpha$ measures the proportion of R&D investment. $\gamma$ reflects Firm 2’s absorptive capability of $s_1$, which indicates a country’s comparative advantages in international specialization. Figure 4(a) shows how the iso-equilibrium curves shift while $\alpha$ changes with other variables \{\rho, \beta, \gamma, n\} fixed to \{0.04, 1, 0.5, 2\}. Figure 4(b) demonstrates the situation when $\gamma$ varies with \{\rho, \alpha, \beta, n\} set to \{0.04, 0.1, 1, 2\}. These two figures lead to our second and third hypotheses.

![Figure 4](attachment:image.png)

**Figure 4** simulation results when $\alpha$ and $\gamma$ changes, respectively.
**H2:** In a sector with high R&D intensity (bigger $\alpha$), a lower $P$ or longer $T$ is needed to achieve leap-frogging.

**H3:** In a sector where domestic firms’ absorptive capability is weak (smaller $\gamma$), a lower $P$ or longer $T$ is needed to promote leap-frogging.

Each patent corresponds to a specific $n$, which measures the importance or profitability of an invention. When no discrimination exists, more important patent applications will receive quicker process (Dranove and Meltzer, 1994). However, if strategic discrimination is employed, the opposite is true. Figure 5(a) is obtained given $\{\rho, \alpha, \beta, \gamma\}$ set to $\{0.04, 0.1, 1, 0.5\}$ and it suggests

**H4:** The more countries a foreign invention gets registered (bigger $n$), the lower $P$ or longer $T$ is needed to stimulate domestic firms’ leap-frogging.

$\rho$ is the annual discount rate and it varies across firms. Figure 5(b) depicts how the iso-equilibrium curve shifts as $\rho$ varies given the other variables $\{\alpha, \beta, \gamma, n\}$ fixed to $\{0.1, 1, 0.5, 2\}$. The simulation results suggest that when firms face a larger discount rate, a lower $P$ or longer $T$ is needed to promote leap-frogging. However, this hypothesis is difficult to test because in business $\rho$ is determined by each firm’s cost of capital. There is rarely an instrumental variable that captures the variation of a company’s discount rate. Nonetheless, this hypothesis may exist in theory but is rarely practical, because even patent examiners hardly obtain such information. Therefore, we will not test this hypothesis in our econometric models.
Finally, all the iso-equilibrium curves in Figures 3-5 are monotonous increasing. This indicates a clear substitute effect between $P$ and $T$. If Firm 1’s patent has higher grant probability ($P$), then a longer grant lag ($T$) is needed to stimulate Firm 2 for leap-frogging. Thus we get

**H5**: A combined tool of grant probability ($P$) and grant lag ($T$) makes it easier to promote leap-frogging.

### 4. Econometric model and data

In this section, we want to test whether the grant probability ($P$) and grant lag ($T$) are affected by variables $\{\alpha, \beta, \gamma, n\}$ in the real world. The reason is that even through the strategic patent policy demonstrated in the theoretical model can foster domestic firms to leap-frog, such policy might not be employed in national patent offices. Patent examiners usually have less freedom of employing strategic actions than trade policy makers. For example, in the WTO framework national treatment applies to both traded products and patents (Article 3 of TRIPS). However, national treatment for trade only applies once a

---

**Figure 5** simulation results when $n$ and $\rho$ changes, respectively.
product has entered the market. Therefore, charging customs duty on an import is not a violation of national treatment even if locally-produced products are not charged an equivalent tax. In patent field such policy tools are not available. Moreover, even if patent offices intend to employ strategic patent policies, they may have other goals such as maximize a weighted welfare function of producers and costumers rather than only focusing domestic producers’ leap-frogging. The welfare between producers and costumers might inconsistent. For example, Figures 1 and 2 demonstrates leap-frogging can be incurred once the patent grant lag is extended from two years to six years. However, this is achieved at the cost of a lowered quality for consumers from 1.43 to 1.17.

Nonetheless, recent literature on political economics has discussed how special interest groups can affect the weight of governments’ welfare function. For example, Grossman and Helpman (1994) develop a model in which special-interest groups make political contributions in order to influence an incumbent government's choice of trade policy. Goldberg and Maggi (1999) test the Grossman-Helpman model and find that the weight of welfare in the US government's objective function is many times larger than the weight of contribution, which supports the model. Branstetter and Freenstra (2002) find that in China the weight applied to consumer welfare is between one-seventh and one-quarter of the weight applied to the output of state-owned enterprises. If the Grossman-Helpman hypothesis is justified in certain governments, then patent offices probably also put a higher weight on firms’ welfare than consumers’ welfare since they are more interactive with patent applicants from firms rather than consumers. Therefore, there is
possibility that in the real world certain patent offices will assist local firms’ leapfrogging even at the cost of domestic consumers’ welfare loss.

To put the theoretical hypotheses into test, first we need measure the variables proposed in the model. The grant probability ($P$) is not observable in the real data. However, we can observe whether a patent is granted or not. Denote $g_i$ as a dichotomic variable for patent $i$. $g_i = 1$ stands for the success of grant and $g_i = 0$ the failure of grant. Thus the grant probability can be written as $P_i(g_i = 1 | X)$ which is conditional on vector $X$. $X$ is the vector of $\{A, R, N\}$ that measures variables $\{\alpha, \gamma, n\}$. Because variable $g_i$ only takes two discrete values, we can specify $P_i(g_i = 1 | X)$ in a Logit model (Equation 27). The logistic function ensures the right-hand side value always falls into interval $[0, 1]$.

\[
(27) \quad P_i(g_i = 1 | X) = \frac{1}{1 + \exp[-(X'\beta + \epsilon_i)]}
\]

Similarly, the grant lag ($T$) can be featured into a count model. By definition, the lag can be calculated as the time period between a patent’s first publication date (not granted yet) and the last publication date when it is granted. Since $T$ is a count number and always positive, it can be modeled as follows

\[
(28) \quad T_i = \exp(X'\tilde{\beta} + \tilde{\epsilon}_i)
\]

where the exponential form ensures the right-hand side being positive.
In Equations 27 and 28, $\mathbf{B}$ and $\tilde{\mathbf{B}}$ are the vectors of coefficients while $\varepsilon_i$ and $\tilde{\varepsilon}_i$ stand for error terms. More specifically, $\mathbf{X}'\mathbf{B}$ can be expanded as ($\mathbf{X}'\tilde{\mathbf{B}}$ is similar):

$$
(29) \quad \mathbf{X}'\mathbf{B} = b_0 + b_1v + b_4A + b_6R + b_8N + b_{10}A\times v + b_{12}R\times v + b_{14}N\times v
$$

where $A\times v$, $R\times v$ and $N\times v$ are the interaction terms. Dummy variable $v$ distinguishes the origin of a patent. $v=0$ indicates it is a domestic application whereas $v=1$ indicates it is filed by a foreign applicant. $\alpha$ is the proportion of R&D investment and thus it is a sector-specific variable. We use industry indicator $A$ to measure $\alpha$. Each patent application in a patent office is assigned with one or more International Patent Classification (IPC) codes. However, the IPC codes are technological categories rather than industrial categories (e.g. NACE).\(^{11}\) Hence, following Schmoch et al. (2003) we converted each IPC code to a NACE code.\(^{12}\) Furthermore, the NACE codes fall into four groups of industrial levels, namely, high-tech, medium-high-tech, medium-low-tech and low-tech industries.\(^{13}\) Therefore, we construct the industry indicator ($A$) by assigning integers 1-4 to the four industrial categories, respectively. If a patent is assigned with multiple IPC codes, we use the average value of the integers.

$\gamma$ reflects Firm 2’s ability to absorb Firm 1’s technology. We define $R_i$ for patent $i$ as the ratio of the total number of native patent applications to the total number of the native

\(^{11}\) NACE stands for Nomenclature Generale des Activites Economiques dans l’Union Europeenne and is used in European Union countries.
\(^{12}\) Currently two versions of NACE are available, NACE Rev 1.1 and NACE Rev 2. We use NACE Rev 1.1 in this paper.
\(^{13}\) The four industrial levels are defined by the statistical office of the European Union (Eurostat). See the Eurostat web site (accessed on Nov. 17, 2010): http://epp.eurostat.ec.europa.eu/cache/ITY_SDDS/Annexes/htec_esms_an2.pdf
and foreign applications in patent \( i \)'s IPC field.\(^{14}\) Patents in different technological fields have different values for \( R_i \). A higher value of \( R_i \) indicates that the domestic firms have stronger absorption ability in their technological fields.\(^{15}\) \( R_i \) is time-variant because new applications are lodged every day. Therefore, even patents in the same technological fields can correspond to different values of \( R_i \), because they could be granted in different time.

\( n \) indicates the number of markets/countries entered by a multi-national corporation. We measure it by the number of patents (\( N \)) in a patent family (Lanjouw et al. 1998, Burke and Reitzig 2007). A patent family contains the same invention which is patented in more than one country.\(^{16}\) A patent applicant only lodges applications in which countries the invention can earn a profit, because each patent costs application fee and maintenance fee. Thus a larger \( N \) implies a larger number of markets to share the R&D cost of the invention.

The patent data of this study are drawn from the European Patent Organization Worldwide Patent Statistical Database (also known as PATSTAT). This database covers about 70 million patent records from over 80 countries. However, we only extract the patent applications registered in the patent offices of the US, the UK, Germany, Japan, South Korea and China. Furthermore, we only use the patents which are applied in more than one country (\( N \geq 2 \)). One reason is that the hypotheses stated in the preceding

\(^{14}\) Each IPC code subgroup corresponds to a patent field. The IPC eighth edition (in 2006) consists of 61,397 subgroups.

\(^{15}\) If a patent is categorized into multiple IPC code subgroups, then only the main (first) code is used.

\(^{16}\) A patent family may contain more than one patent granted in a patent office. In this case, we only keep one patent for each patent office so that \( N \) is not overestimated (United States Patent and Trademark Office web site http://www.uspto.gov/main/glossary/#patentfamily (accessed on Nov. 17, 2010).
section are based on the analysis of firms’ behavior. However, a large number of the patents are actually filled by individual inventors who are not covered by the theoretical model. Individual inventors have stricter financial constraints than firms and thus they usually only lodge applications in the home patent office. Therefore, we can eliminate most of the individual patent applications by focusing the applications that are lodged in more than one country. The other reason for this constraint is due to the measure of domestic firms’ absorption ability ($R$). $R$ is the ratio of the number of native patent applications and foreign ones in each patent field. However, a higher $R$ may not represent higher absorption ability if the patent field is crowded with less important native patents. Thus by raising the threshold for important patents, $R$ will more accurately reflect domestic firms’ absorption ability. For the same reason, only invention patents are used in our sample and utility models as well as design patents are excluded.

In the empirical analysis, we will use different samples for the two sets of regressions. Both issued and rejected patents are used in the Logit model (Equation 27). However, only issued patents are used in the count model (Equation 28). The reason is that the grant lag can be prolonged by both patent examiners and patent applicants. As Webster et al. (2008) point out, applicants prefer a short examination process when the expected grant probability is high, but favor a slow examination process when the expected grant probability is low. Hence, the grant lag of rejected patents can be deliberately delayed by applicants. On the contrary, the grant lag of approved patents is mainly affected by the patent examiners. Since we want to test the behavior of patent offices rather than
applicants, only approved patent data are used to estimate Equation 22 so that the effect of applicants can be eliminated.

Figure 6 demonstrates the distribution of patents from 1974 to 2009. Those patents are divided by their origin (home or foreign countries) and assessment results (issued or rejected). Figure 6 shows that the PATSTAT dataset is not complete in certain years when the rejected patents are not reported. For example, the rejected foreign patents in the US patent office are only recorded from 2002 onwards (Figure 6a). Since the Logit model captures the grant ratio, both the granted and rejected patent data are required. As a result, only data in selected years are used for our regressions.

Figure 6 Grant-rejection distribution of patents (\( N \geq 2 \))
Similarly, the grant lags also vary across countries and over time (see Figure 7). There are two distinct regimes of patent publication. Under the first regime, patent applications are not published until they are granted. Germany and Korea adopt this regime which is clearly shown by the data in Figure 7c and 7e. Our count model only is applied to data recorded under the second regime.

Figure 7 Distribution of granted patents on grant lags ( $N \geq 2$ )
Summary information about our database is listed in Table 1. It shows that the data clearance helps trim off huge amounts of unusable data. For instance, there are over 2.5 million applications for utility patents in the US patent office during 2002-2009. However, only 62% (1.6 million) of them were lodged in other patents offices ($N \geq 2$).

In the Japanese patent office, only 18% of the original data enter our regression. The British and Japanese data are split at the year 1994 when the TRIPS was established. The separation of the sample on that year will allow us to capture policy changes in the post TRIPS era.

### Table 1 Data summary (in thousand)

<table>
<thead>
<tr>
<th></th>
<th>US</th>
<th>Britain Pre-TRIPS</th>
<th>Britain Post-TRIPS</th>
<th>Germany Pre-TRIPS</th>
<th>Japan Pre-TRIPS</th>
<th>Japan Post-TRIPS</th>
<th>Korea</th>
<th>China</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native-Rejected</td>
<td>327</td>
<td>26</td>
<td>19</td>
<td>239</td>
<td>N/A</td>
<td>183</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>Native-Granted</td>
<td>216</td>
<td>31</td>
<td>27</td>
<td>78</td>
<td>N/A</td>
<td>151</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Foreign-Rejected</td>
<td>542</td>
<td>1</td>
<td>1</td>
<td>642</td>
<td>N/A</td>
<td>600</td>
<td>184</td>
<td>417</td>
</tr>
<tr>
<td>Foreign-Granted</td>
<td>520</td>
<td>77</td>
<td>59</td>
<td>514</td>
<td>N/A</td>
<td>242</td>
<td>20</td>
<td>309</td>
</tr>
<tr>
<td>Total</td>
<td>1,605</td>
<td>135</td>
<td>106</td>
<td>1,473</td>
<td>N/A</td>
<td>1,176</td>
<td>227</td>
<td>747</td>
</tr>
<tr>
<td>Original data</td>
<td>2,589</td>
<td>199</td>
<td>204</td>
<td>2,194</td>
<td>398</td>
<td>6,529</td>
<td>1,043</td>
<td>1,538</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>US</th>
<th>Britain Pre-TRIPS</th>
<th>Britain Post-TRIPS</th>
<th>Germany Pre-TRIPS</th>
<th>Japan Pre-TRIPS</th>
<th>Japan Post-TRIPS</th>
<th>Korea</th>
<th>China</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native-Granted</td>
<td>160</td>
<td>30</td>
<td>15</td>
<td>N/A</td>
<td>2</td>
<td>129</td>
<td>N/A</td>
<td>6</td>
</tr>
<tr>
<td>Foreign-Granted</td>
<td>408</td>
<td>76</td>
<td>34</td>
<td>N/A</td>
<td>107</td>
<td>225</td>
<td>N/A</td>
<td>309</td>
</tr>
<tr>
<td>Total</td>
<td>568</td>
<td>106</td>
<td>49</td>
<td>N/A</td>
<td>108</td>
<td>354</td>
<td>N/A</td>
<td>315</td>
</tr>
<tr>
<td>Original data</td>
<td>1,279</td>
<td>146</td>
<td>91</td>
<td>972</td>
<td>837</td>
<td>2,185</td>
<td>198</td>
<td>551</td>
</tr>
</tbody>
</table>

5. **Regression results and discussions**

The mean and standard deviation values of the key variables are presented in Table 2. In the regression, because $R$ and $N$ do not take zero values, it would be more meaningful to center them at their means. Therefore variables $A$, $R$ and $N$ are centered to 1, 0.1 and 5 (near their means), respectively. As a result, the estimated coefficients of the foreign

---

17 The data summary doesn’t match in the two models. For example, there are 160 thousand patents granted to native applicants in the US used in the count model, but the number is 216 thousand in the Logit model. This discrepancy occurs because we delete the patents with zero-day grant lag. In the early publication system, this may happen under the request of patent applicants. Thus these patents do not reflect the discriminatory policies discussed in this paper and should not enter the regression.
dummy \((v)\) depict how foreign applications are treated differently in comparison with home applications in the reference group \(A=1\) (high-tech industries), \(R=0.1\) (low native patent concentration) and \(N=5\) (important inventions). The theoretical model implies that strategic discrimination will be more apparent in high-tech industries (\(H2\)) with lower native patent concentration (\(H3\)) and more important inventions (\(H4\)). Thus, the estimated coefficients of \(v\) reflect the degree of discrimination.

Table 2 The mean and standard deviation (in parenthesis) of key variables

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Patents family ((N))</td>
<td>4.05 (3.35)</td>
<td>5.08 (4.48)</td>
<td>4.21 (3.43)</td>
<td>6.33 (4.01)</td>
<td>N/A (0.73)</td>
<td>5.60 (0.79)</td>
<td>6.10 (0.82)</td>
<td>6.17 (0.74)</td>
<td>N/A (0.72)</td>
<td>5.60 (0.70)</td>
<td>6.10 (0.72)</td>
<td>N/A (0.72)</td>
<td>6.10 (0.72)</td>
</tr>
<tr>
<td>Industrial level ((A))</td>
<td>1.52 (0.73)</td>
<td>1.98 (0.79)</td>
<td>1.88 (0.82)</td>
<td>1.79 (0.74)</td>
<td>N/A (0.72)</td>
<td>1.63 (0.74)</td>
<td>1.55 (0.74)</td>
<td>1.60 (0.74)</td>
<td>N/A (0.72)</td>
<td>1.55 (0.72)</td>
<td>1.60 (0.72)</td>
<td>N/A (0.72)</td>
<td>1.60 (0.72)</td>
</tr>
<tr>
<td>Native ratio ((R))</td>
<td>0.32 (0.21)</td>
<td>0.16 (0.15)</td>
<td>0.23 (0.18)</td>
<td>0.24 (0.15)</td>
<td>N/A (0.15)</td>
<td>0.15 (0.15)</td>
<td>0.06 (0.15)</td>
<td>0.01 (0.15)</td>
<td>N/A (0.13)</td>
<td>0.06 (0.13)</td>
<td>0.01 (0.13)</td>
<td>N/A (0.13)</td>
<td>0.01 (0.13)</td>
</tr>
<tr>
<td>Grant ratio</td>
<td>0.46 (0.50)</td>
<td>0.80 (0.40)</td>
<td>0.81 (0.39)</td>
<td>0.40 (0.49)</td>
<td>N/A (0.47)</td>
<td>0.33 (0.47)</td>
<td>0.10 (0.30)</td>
<td>0.42 (0.49)</td>
<td>N/A (0.47)</td>
<td>0.10 (0.30)</td>
<td>0.42 (0.49)</td>
<td>N/A (0.47)</td>
<td>0.10 (0.30)</td>
</tr>
</tbody>
</table>

Tables 3 and 4 list the regression results for the Logit and count models, respectively. The maximum-likelihood estimation (MLE) method is employed in both cases. For the Logit model, we have assumed that the response variable followed the logistic distribution. For the count model, two types of distribution can be applied, namely, the Poisson distribution and the negative binomial distribution (Hausman et al., 1984). The Poisson distribution has an implicit restriction, that is, the variance of the sample is equal to the mean. However, real count data are commonly observed with inconsistence.
between the variance and mean. Therefore, researchers routinely employ a more general specification, usually the negative binomial distribution which allows for over-dispersion (the variance being larger than the mean). The dispersion value in Table 4 measures the degree of dispersion. If the dispersion value equals zero, the model is reduced to the Poisson model. However, we observed all the dispersion values in Models 8-13 are greater than zero, which confirms that the dependent variable is over-dispersed. This means that the negative binomial distribution fits our count models better than the Poisson distribution.

Table 3 Regression results of the Logit model

<table>
<thead>
<tr>
<th>Model</th>
<th>US Pre-TRIPS</th>
<th>Britain Post-TRIPS</th>
<th>Germany Post-TRIPS</th>
<th>Japan Post-TRIPS</th>
<th>Korea</th>
<th>China</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.22 (0.01)***</td>
<td>-0.04 (0.01)***</td>
<td>0.03 (0.01)***</td>
<td>-1.40 (0.01)***</td>
<td>0.51 (0.01)***</td>
<td>-2.00 (0.03)***</td>
</tr>
<tr>
<td>Foreign dummy (V)</td>
<td>-0.12 (0.01)***</td>
<td>3.97 (0.04)***</td>
<td>4.06 (0.05)***</td>
<td>1.17 (0.01)***</td>
<td>-1.51 (0.01)***</td>
<td>-0.56 (0.03)***</td>
</tr>
<tr>
<td>Patent family (N)</td>
<td>0.05 (0.00)***</td>
<td>-0.03 (0.00)***</td>
<td>0.01 (0.00)***</td>
<td>-0.11 (0.00)***</td>
<td>0.20 (0.00)***</td>
<td>-0.06 (0.01)***</td>
</tr>
<tr>
<td>Industrial level (A)</td>
<td>0.15 (0.00)***</td>
<td>-0.07 (0.01)***</td>
<td>-0.09 (0.01)***</td>
<td>0.21 (0.00)***</td>
<td>0.04 (0.02)***</td>
<td>0.19 (0.01)***</td>
</tr>
<tr>
<td>Native ratio (R)</td>
<td>-1.99 (0.02)***</td>
<td>2.29 (0.06)***</td>
<td>2.37 (0.06)***</td>
<td>0.19 (0.03)***</td>
<td>-2.22 (0.02)***</td>
<td>-0.58 (0.09)***</td>
</tr>
<tr>
<td>V×N</td>
<td>0.00 (0.00)***</td>
<td>-0.01 (0.01)***</td>
<td>-0.06 (0.01)***</td>
<td>0.12 (0.00)***</td>
<td>-0.16 (0.01)***</td>
<td>0.09 (0.01)***</td>
</tr>
<tr>
<td>V×A</td>
<td>-0.12 (0.00)***</td>
<td>0.21 (0.04)***</td>
<td>0.12 (0.04)***</td>
<td>-0.18 (0.01)***</td>
<td>0.03 (0.03)***</td>
<td>-0.33 (0.02)***</td>
</tr>
<tr>
<td>V×R</td>
<td>1.22 (0.02)***</td>
<td>-3.51 (0.18)***</td>
<td>-3.08 (0.19)***</td>
<td>-0.43 (0.03)***</td>
<td>-2.49 (0.03)***</td>
<td>-4.61 (0.17)***</td>
</tr>
</tbody>
</table>

*Note: Standard error in parentheses. *p<0.1, **p<0.05, ***p<0.01.*
The estimated coefficients in Tables 3 and 4 are not the direct marginal effects because the two original models are non-linear (see Equations 27 and 28). The estimated coefficients stand for the marginal effects on the log odds ratio in the Logit regressions and on the log grant lag in the negative binomial regressions. However, the log odds ratio is a monotonously increasing function of the grant probability and the log grant lag is monotonously increasing against the grant lag. Hence, the non-linearity doesn’t affect the interpretation of the results.

18 The odds ratio is defined as the ratio of “the probability that a patent application is granted” to “the probability that the application is rejected”.

Table 4 Regression results of the count model (Negative Binomial regressions)

<table>
<thead>
<tr>
<th>Model</th>
<th>US Pre-TRIPS</th>
<th>US Post-TRIPS</th>
<th>Britain Pre-TRIPS</th>
<th>Britain Post-TRIPS</th>
<th>Japan Pre-TRIPS</th>
<th>Japan Post-TRIPS</th>
<th>Japan</th>
<th>China</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>6.75 (0.00)***</td>
<td>5.27 (0.00)***</td>
<td>5.14 (0.01)***</td>
<td>7.45 (0.02)***</td>
<td>7.58 (0.00)***</td>
<td>6.87 (0.01)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foreign dummy (V)</td>
<td>-0.11 (0.00)***</td>
<td>0.02 (0.00)***</td>
<td>-0.14 (0.01)***</td>
<td>0.18 (0.02)***</td>
<td>0.06 (0.00)***</td>
<td>0.20 (0.01)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patent family (N)</td>
<td>0.00 (0.00)***</td>
<td>0.00 (0.00)***</td>
<td>-0.02 (0.01)***</td>
<td>0.00 (0.02)***</td>
<td>-0.01 (0.00)***</td>
<td>0.02 (0.01)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industrial level (A)</td>
<td>-0.13 (0.00)***</td>
<td>0.01 (0.00)***</td>
<td>-0.01 (0.01)***</td>
<td>-0.06 (0.02)***</td>
<td>0.00 (0.00)***</td>
<td>-0.03 (0.00)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Native ratio (R)</td>
<td>-0.09 (0.00)***</td>
<td>-0.11 (0.00)***</td>
<td>-0.12 (0.01)***</td>
<td>-1.25 (0.02)***</td>
<td>-0.81 (0.03)***</td>
<td>-0.03 (0.04)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V×N</td>
<td>0.01 (0.00)***</td>
<td>0.01 (0.00)***</td>
<td>0.00 (0.00)***</td>
<td>0.00 (0.00)***</td>
<td>0.01 (0.01)***</td>
<td>0.01 (0.00)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V×A</td>
<td>0.01 (0.00)***</td>
<td>0.02 (0.00)***</td>
<td>0.03 (0.00)***</td>
<td>0.03 (0.00)***</td>
<td>0.00 (0.00)***</td>
<td>0.00 (0.00)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V×R</td>
<td>0.05 (0.00)***</td>
<td>0.15 (0.00)***</td>
<td>0.09 (0.00)***</td>
<td>-0.88 (0.00)***</td>
<td>-0.58 (0.00)***</td>
<td>-0.24 (0.00)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dispersion</td>
<td>0.39 (0.01)***</td>
<td>0.12 (0.00)***</td>
<td>0.26 (0.01)***</td>
<td>0.16 (0.02)***</td>
<td>0.32 (0.01)***</td>
<td>0.17 (0.04)***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Standard error in parentheses. *p<0.1, **p<0.05, ***p<0.01.
If the estimated coefficients of $v$ are negative in the Logit models (Models 1-7), it means that the grant probability on average is lower for foreign patents than that for native ones. If the estimated coefficients of $v$ are positive in the count models (Models 8-13), it implies the grant lag is on average longer for foreign patents than that for native ones. Therefore, the sign of the estimated coefficients of $v$ is the critical signal for discriminatory policies. For example, the estimated coefficients of $v$ are positive in Models 3-4 and negative in Model 10, which indicates no evidence of discriminatory policies employed in Britain and Germany during the post-TRIPS period.

In Model 1 the negative coefficients of $v$ seem to suggest foreign innovators receive lower grant probability in the US patent office. However, the corresponding coefficient of $v$ in Model 8 is negative, which implies that foreign patents also receive shorter pendency. Recall that a combined tool of grant probability ($P$) and grant lag ($T$) is required for promoting leap-frogging (H5) and hence the evidence is not sufficient to conclude the US practices discrimination. The similar case applies to pre-TRIPS Britain. In contrast, there is strong evidence of the existence of discrimination in Japan, Korea and China. Foreign patents in these countries’ patent offices tend to receive lower grant probability (negative coefficients of $v$ in Models 5-7) as well as longer grant lags (positive coefficients of $v$ in Models 11-13). Nonetheless, other indicators have to be examined in order to judge whether such discrimination is employed strategically or institutionally. Strategic discrimination means that the grant probability for foreign patents will be negatively related to $N$ (H4) and positively related to $A$ and $R$ (H2 and H3) and the grant lag for foreign patents will be positively related to $N$ and negatively related
to A and R. In other words, hypotheses 2-4 imply that the estimated coefficient of \( v \times N \) shall take a negative value in the Logit model and a positive one in the count model, while those of \( v \times A \) and \( v \times R \) shall be positive in the Logit model and negative in the count model.

The regression results show that **H2, H3 and H4** are verified in Models 5 and 12 which represent the post-TRIPS Japan. Only the negative coefficient of \( v \times R \) in Model 5 is unexpected. Due to data limitation, only the Logit model is applicable for the Korean samples (Model 6). The estimated coefficients of \( v \times N \), \( v \times A \) and \( v \times R \) in Model 6 don’t support **H2-H4**. Thus the discriminatory policy in Korea is institutional rather than strategic. China and pre-TRIPS Japan stand between these two cases. As we stated before, institutional discrimination is more serious and overt protection for native firms, while strategic discrimination is rather selective and disguised. Given the size of the Japanese economy, if discriminatory policies are to be implemented, strategic discrimination is a better choice since it is more likely to escape international monitor and retaliation. The same argument applies to today’s China. However, as a relatively smaller economy, Korea has some room to employ the all-out institutional discrimination.

The results also show that TRIPS does affect the behavior of patent offices. Since the passing of TRIPS in 1994, foreign applications have been treated better in the Japanese and British patent offices (these two offices are the only ones we can observe given the sample). For example, foreign patents in the UK tend to have higher grant probability (compare coefficients of \( v \) in Models 2 and 3) and shorter pendency (Model 9 vs. Model
Furthermore, the estimated coefficient of $v \times A$ becomes insignificant in Model 12, which indicates that in post-TRIPS Japan foreign patents in the low-tech industries are no longer subjected to severe discrimination.

Finally, $H1$ suggests that in a regime where the punishment of patent infringement ($\beta$) is stricter, more severe discrimination is needed to promote leap-frogging. The OECD index GTRIC-e is roughly negatively correlated with $\beta$ (OECD, 2008, P107).\textsuperscript{19} The GTRIC-e score for China is 0.9748, much higher than those for Korea (0.6085) and Japan (0.0419). Therefore, if both Japan and China employ strategic discrimination, the discrimination will be more severe in Japan than in China according to $H1$. Models 5, 7, 12 and 13 provide some insight into this question.\textsuperscript{20} The magnitude of the estimated coefficients of $v \times N$, $v \times A$ and $v \times R$ indicates the degree of discrimination. Therefore according to $H1$, the estimated coefficient of $v \times N$ in Model 5 should be smaller than that in Model 7 and the coefficient of $v \times N$ in Model 12 should be larger than that in Model 13. These are confirmed by the regression results. The estimated coefficients of other interaction terms (except $v \times R$ in Models 5 and 7) also support $H1$.

6. Conclusions

This paper answers two questions, namely, 1) whether discrimination against foreigners exists in the patent examination procedure, and 2) whether such discriminatory policy is

\textsuperscript{19} GTRIC-e stands for General Trade-Related Index of Counterfeiting for economies

\textsuperscript{20} The Japanese and Chinese data are comparable because the data cover the same time period (1994-2009) and variables $A$, $R$ and $N$ are centered to the same values.
introduced with a strategic method. In parallel with strategic trade barriers, discriminatory policies in the invention field can be used to foster technological catching-up or even leap-frogging. A combined tool of lower grant probability and longer grant lags can be utilized to discriminate against foreign patents. We built a game theory model to reveal how this tool can be employed strategically. The traditional discrimination is institutional in the sense that all foreign patents are discriminated. However this overt discrimination can be very easily detected. Instead of abolishing the discriminatory policies, a country could switch to a type of more disguised, strategic discrimination which targets at certain key patent fields and spares others. Our simulation results help to identify the characteristics of these key patents, that is, they are important inventions, in high-tech industries and costly for imitation.

Our empirical findings suggest no clear evidence of discrimination in the US, British and German patent offices. However, Japanese, Korean and Chinese patent offices are very likely practicing discriminatory policies. TRIPS has pushed the Japanese patent office away from practicing institutional discrimination and to switch to strategic discrimination. Korean patent office still maintains its institutional discrimination after TRIPS because Korean economy is much smaller and thus is less monitored. Chinese patent office stands between the two, but may move toward the Japanese model as the Chinese economy is heading to the world’s No. 1.
Appendix A: Derivation of R&D cost functions

We derive the R&D cost function from the knowledge production function. New knowledge is created by new R&D investment based on the established knowledge (Griliches, 1979; Furmana et al., 2002). Thus the knowledge production function can be written as

\[ s_t = K(s_{t-1}, r_t) \]

where \( s_t \) denotes new technology in period \( t \), created by R&D investment \( (r_t) \) in the period \( t \) on the basis of the existing technology \( (s_{t-1}) \). Without loss of generality, we assign an additive form for A1 and get

\[ s_t = \gamma s_{t-1} + \tilde{K}(r_t) \]

where \( \gamma \in [0,1] \) reflects the absorptive capability of the existing knowledge stock. Function \( \tilde{K}(r_t) \) measures how much new knowledge is induced by R&D investment.

The marginal product of R&D investment is not unlimited but constrained by the existing knowledge stock. For example, even infinite R&D investment cannot introduce iPhone 4 in year 1950. Thus the marginal product of R&D investment must converge to zero as the investment increases to infinity.\(^{21}\) This means that given a certain stock of knowledge \( (s_{t-1}) \) only a bounded technology improvement \( (\tilde{s}) \) is feasible.

\[ \lim_{r_t \to \infty} s_t = \gamma s_{t-1} + \lim_{r_t \to \infty} \tilde{K}(r_t) = \gamma s_{t-1} + \tilde{s}, \text{ and} \]

---

\(^{21}\) The zero marginal product of a specific input in a production function has long been acknowledged. For instance, Suliman (1997) and Färe (1974) have studied about zero marginal product of labor in production functions.
(A4) \( \lim_{t \to 0} s_t = \gamma s_{t-1} + \lim_{t \to 0} K(r_t) = \gamma s_{t-1} \)

A proper cost function should be in harmony with the knowledge production function. However, the commonly employed quadratic form of cost functions doesn’t satisfy this condition. A typical quadratic cost function can be represented as \( r = x^2 \). This cost function implies that as \( t \to \infty \), \( s \to \infty \) which violates condition A3.

The simplest form of a cost function that satisfies conditions A3 and A4 is

\[
(A5) \quad r_t = C(s_t) = \frac{s_t - \gamma s_{t-1}}{\gamma s_{t-1} + \bar{s} - s_t}
\]

This is the baseline cost function that we assume for Firms 1 and 2. Firm 1 is a multinational corporation. The cost for R&D is divided by the number of markets/countries \( (n) \) that it enters. Thus the R&D cost function of Firm 1 can be written as

\[
(A6) \quad C_1(s_{1,t}) = \frac{1}{n} \left( \frac{s_{1,t} - \gamma s_{1,t-1}}{\gamma s_{1,t-1} + \bar{s} - s_{1,t}} \right)
\]

where \( s_{1,t-1} \) denotes the stock of knowledge that Firm 1 faces when it conducts innovation on quality improvement. Without loss of generality, we normalize the initial knowledge stock \( (s_{1,t-1}) \) to zero and standardize the upper bound of R&D productivity \( (\bar{s}) \) to one. This gives out Firm 1’s cost function.

---

22 Other forms of the cost function also exist. In fact, for any positive real numbers \( n \) and \( a_n \),

\[
C(s_t) = \sum_n a_n \left( \frac{s_t - \gamma s_{t-1}}{\gamma s_{t-1} + \bar{s} - s_t} \right)^n
\]

are all eligible choices. However, since the selection of cost functions doesn’t affect the main conclusions of this paper, we will adopt the simplest function A5.
Firm 2 only operates in the domestic market and hence its cost cannot be shared in other markets. However, its R&D activities can benefit from the knowledge of Firm 1’s published patent \( s_i \). Thus Firm 2’s cost function can be written as

\[
C_2(s_2 \mid s_i) = \frac{s_2 - \gamma s_i}{\gamma s_i + 1 - s_2}, \quad (\gamma s_i \leq s_2 < 1 + \gamma s_i)
\]

Finally, we add a scaling parameter \( \alpha \) to capture the industrial characteristics. The share of R&D investment in firms’ revenue changes substantially across sectors. Thus we use a larger \( \alpha \) to indicate a larger proportion of earnings is invested in R&D investment in an industry:

\[
C_i(s_i) = \frac{1}{n} \frac{\alpha s_i}{1 - s_i}, \quad (0 \leq s_i < 1)
\]

\[
C_2(s_2 \mid s_i) = \frac{\alpha(s_2 - \gamma s_i)}{1 + \gamma s_i - s_2}, \quad (\gamma s_i \leq s_2 < 1 + \gamma s_i)
\]

In the numeric simulation exercises (Section 3) we have to choose an initial value for \( \alpha \). Higona and Antolinb (2012) estimated that the share of R&D investment in revenue was 0.4019 for foreign multinationals in the UK. Chan et al. (2001) also showed that R&D expenditures typically amounted to 30%-80% of firms’ earnings in the US. Hence, we assume the share of R&D investment in revenue is 0.4 and obtain

\[
\frac{1}{n} \frac{\alpha s_i}{1 - s_i} = 0.4 \times \pi_i
\]

During the period of patent protection, Firm 1 can receive profit \( \pi'_i = \frac{s_i}{4} \) (Equation 20).

Thus we get
The initial value of \( n \) and \( s_i \) are taken as 2 and 0.5, respectively (Section 3). Thus a proper initial value for \( \alpha \) is 0.1.

**Appendix B: Derivation of demand functions**

We assume that products are vertically differentiated. Each consumer is characterized by a parameter \( \theta \) and has the following utility function (Motta, 1993):

\[
U_\theta = \begin{cases} 
\theta s - p, & \text{if the consumer buys good with quality } s \text{ and price } p, \\
0, & \text{if the consumer does not buy.}
\end{cases}
\]

\( \theta \) indicates a consumer’s appreciation of quality. We assume that \( \theta \) is uniformly distributed over the interval \([0, 1]\) (Li and Song, 2009).

Based on the qualities and prices offered by Firms 1 and 2, the consumers choose buying the product from Firm 1 or from Firm 2 or not buying at all. Consumers’ choices determine the two firms’ demand. In the case of \( s_1 > s_2 \), the marginal consumer who is indifferent about buying either quality \( s_1 \) or quality \( s_2 \) is determined by \( \theta s_1 - p_1 = \theta s_2 - p_2 \).

Hence \( \theta_{12} = \frac{p_1 - p_2}{s_1 - s_2} \). As a result, the demand function for the good of quality \( s_1 \) can be derived as

\[
q_1 = \int_{s_2}^{s_1} d\theta = 1 - \frac{p_1 - p_2}{s_1 - s_2}
\]
Similarly, the marginal consumer who is indifferent between buying the good of quality $s_2$ and not buying at all is determined by $\theta s_2 - p_2 = 0$, Hence, $\theta_{s_2} = \frac{p_2}{s_2}$. Thus the demand function for the good of quality $s_2$ is

(B3) \[ q_2 = \int_{\theta_{s_2}}^{\Phi_{s_2}} d\theta = \frac{p_1 - p_2}{s_1 - s_2} - \frac{p_2}{s_2}. \]

Based on the same method, in the case where Firm 2 produces a higher quality ($s_2 > s_1$), the demands are

(B4) \[ q_2 = 1 - \frac{p_2 - p_1}{s_2 - s_1} \]

and

(B5) \[ q_1 = \frac{p_2 - p_1}{s_2 - s_1} - \frac{p_1}{s_1}. \]

Firm 1 will get the monopolistic profit $\pi_1^M$ once its patent is granted. In this case, the marginal consumer who is indifferent between buying the good of quality $s_1$ and not buying at all is determined by

(B6) \[ \theta s_1 - p_1 = 0 \]

Hence, $\theta_{s_1} = \frac{p_1}{s_1}$. Now the demand function is

(B7) \[ q_1 = \int_{\theta_0}^{\Phi_1} d\theta = 1 - \frac{p_1}{s_1}. \]

References

Economics 47: 287-312.


Review 89 (5), 1135–1155.


Kovac, E. and K. Zigic (2007). "International competition in vertically differentiated markets with innovation and imitation: Trade policy versus free trade." The Center for Economic Research and Graduate Education - Economic Institute


