

International Protection of Intellectual Property: An Empirical Investigation*

Edwin L.-C. Lai[†] and Isabel K. Yan[‡]

November 2004

Abstract

This paper is a pioneering study on the estimation of Grossman and Lai's (2004) game-theoretic model of international patent protection. The model yields clear predictions of the variation in degrees of patent protection across countries based on the market sizes and levels of innovative capability of the countries, and interdependence of the degrees of patent protection between countries. To correctly take into account the interaction between countries, we borrow an approach from the spatial econometrics literature. We find that the pattern of patent protection around the globe was broadly consistent with the predictions of the non-cooperative game model before the early 1990s. As conjectured, the non-cooperative game model was less applicable after the early 1990s, in view of the implementation of the TRIPS Agreement.

Keywords: intellectual property, patent, empirical study, spatial econometrics

JEL Classification Number: O34, F13

[†] Corresponding author. Department of Economics and Finance, City University of Hong Kong, Kowloon, Hong Kong. Phone: +(852)2788-7317; Fax: +(852)2788-8806; E-mail: edwin.lai@cityu.edu.hk.

[‡] Department of Economics and Finance, City University of Hong Kong, Kowloon, Hong Kong. Phone: +(852)2788-7315; Fax: +(852)2788-8806; E-mail: efsyan@cityu.edu.hk.

* The work in this paper has been supported by the Research Grants Council of Hong Kong, China (Project no. CityU 1145/99H) and the Research Center for International Economics at City University of Hong Kong. The authors would like to thank participants in the seminars in Hong Kong Univ. of Sci. and Tech., Chinese Univ. of Hong Kong, Univ. of Hong Kong, and 2004 IEFIS conference in San Diego, for helpful comments.

1 Introduction

Intellectual property protection (IPR) comes to the forefront of international trade negotiations nowadays. Since the early 1990s, the US and other EU countries began to exert ever higher pressure for other countries to adopt more stringent standards in protecting patents, trademarks, copyrights, trade secrets, geographical indications, new plant varieties and other biotechnological products in the real as well as cyber- world. These efforts culminated in the signing of the TRIPS (Trade-Related Intellectual Property) Agreement of the Uruguay Round of GATT in 1994. By January 1, 1996, all developed countries were required to adopt the universal minimum IPR standards; by January 2000, all developing and transition economies were required to adopt the same standards; and by January 2006, the poorest countries should also adopt these standards (2016 for pharmaceutical patent protection).

It is obvious that TRIPS has enormous income distribution implications among countries in the world. McCalman (2001) found that the US was by far the largest beneficiary of TRIPS (4.6 billion USD), followed by Germany (0.79 billion USD) and France (0.57 billion USD) as distant second and third beneficiaries. On the other hand, the greatest loser was Canada (1 billion USD), followed by Brazil (0.93 billion USD) and UK (0.54 billion USD). Obviously, many countries are reluctant to adopt the standards stipulated by TRIPS had there not been quid pro quo in other trade or non-trade issues.

Countries have long sought to coordinate on their intellectual property policies, but virtually all international agreements on IPR lacked any binding power, until the signing of the TRIPS Agreement in 1994. For example, the Paris Convention (1883) for the protection of industrial property and the Berne Convention (1886) for the protection of artistic and literary property (mainly copyrights), both managed by the World Intellectual Property Organization (WIPO), had been in place for a long time. Yet, the lack of a dispute settlement mechanism rendered the treaties rather toothless. On the other hand, using Section 301 of the Trade Act of 1974, the US had been able to pressure South Korea, Argentina, Brazil, Thailand, Taiwan and China to adopt stronger IPR

legislations in the 1990s, while the EU had been able to pressure Egypt and Turkey to do the same during the same period (Maskus 2000). Therefore, intuitive reasoning points to the hypothesis that countries by and large behaved non-cooperatively in setting their strengths of patent protection before the beginning of the 1990s; and then they became more and more cooperative over time in the 1990s as the US, and to a lesser extent, the EU, exerted more and more pressure on other countries to strengthen their patent protection. This culminated in the signing of the TRIPS agreement in 1994, which called for some universal minimum standards regarding intellectual property protection to be adopted by all members of the WTO.

This paper focuses on testing the empirical implications of the game-theoretic model of intellectual property protection developed by Grossman and Lai (2004). We focus on patent protection in this paper, and leave the research into other important aspects of intellectual property (IP) protection, such as trademarks and copyrights, to later studies. Specifically, we are interested in knowing whether the variation in the degrees of patent protection across countries, in the absence of binding international cooperation such as the TRIPS, can be explained by the Grossman-Lai model. Then, we ask whether TRIPS agreement has really led to cooperative increase in patent protection in all countries beyond the incentive for countries to protect patents non-cooperatively.

Grossman and Lai (2004) propose a theoretical framework for explaining the variation in degrees of patent protection across countries as the outcome of a non-cooperative Nash game with national governments as players choosing the degrees of patent protection. The Grossman-Lai model yields clear predictions of the variation in degrees of patent protection across countries based on the market sizes and levels of innovative capability of the countries as well as interdependence of the degrees of patent protection between countries. These hypotheses are tested using spatial econometric technique. We find that the pattern of patent protection around the globe is broadly consistent with the predictions of the non-cooperative game model in 1980, 1985 and 1990. We also find that the model is less applicable to the years 1995 and 2000, which is consistent with

our conjecture that countries in the world became more cooperative in providing patent protection after the TRIPS agreement was signed.

The significance of this study lies not only in its validation of the Grossman-Lai model, but also that it provides evidence that the degrees of patent protection adopted by countries are guided by economic considerations of marginal costs and marginal benefits as economic theory predicts. Furthermore, the study provides evidence that countries behaved strategically in determining their domestic strength of patent protection absent any binding international agreements.

In the literature, there are a number of empirical studies on the determinants of patent protection. Two examples are Maskus and Penubarti (1995) and Ginarte and Park (1997). A distinguishing feature of the present paper is that it is based on a fully specified theoretical model. The econometric estimations are tightly guided by the variables, the functional form and the game-theoretic structure of the theoretical model. Nevertheless, the findings of this paper are fairly consistent with the findings of the studies in the literature. Using linear regression, studies in the existing literature similarly find that proxy for market size (like GDP) and proxy for innovative capability (like the share of scientists and engineers in the labor force or the ratio of R&D to GDP) are significant in explaining the variation of the strengths of patent protection across countries. These studies, however, typically do not have a fully specified theory to guide their estimations.

In Section 2, we briefly describe the theory, which basically draws from Grossman and Lai (2004). We then partially linearize the main equation so as to allow us to apply spatial econometrics technique to test the hypotheses. In Section 3, we describe the estimation procedure using spatial econometrics approach, discuss the results and test the predictive power of the structural model. Section 4 augments the spatial econometrics model in Section 3 by taking into account the depth of trade relationship between country pairs in determining the interdependence of their strengths of patent protection. Section 5 concludes.

2 The Theory

The theory we want to test comes directly from Grossman and Lai (2004). For detail of the model, the reader should refer to the paper directly. Here is a summary of the assumptions and features of the model, with the assumptions being set in italics.

1. *Consumers decide how much of each good to purchase given their budget constraint.* Given *there is international trade*, they would benefit whenever there are more inventions from any country in the world.

2. The value of a firm's patent increases as the degree of patent protection in any of its markets increases. The value of the patent also increases with the size of each market. When the value of a patent increases, there is more incentive to invent. Therefore, there would be more inventions from all countries whenever the market size or the degree of protection of any country increases.

3. *Each government chooses the degree of patent protection to maximize the present discounted value of the sum of consumer surplus and the returns to capital (firm owners are owners of capital), given the degree of patent protection of other countries.* Given the behavior of the consumers and the firms mentioned above, each government can figure out its own best response function.

4. *Firms and government are forward-looking. When evaluating costs and benefits, they would take into account the present discounted value of them. Consumers simply maximize their utility subject to the budget constraint in each period.*

5. *When a government announces a change of the degree of patent protection at time 0, the new policy only applies to inventions that take place after time 0. Old inventions are subjected to whatever patent policy was in effect at the time of invention.*

6. *It is assumed that agents believe that the relative market size M_i/M_j and relative number of inventions I_i/I_j (for all countries $i \neq j$) remain more or less constant in the future.*

According to Grossman and Lai (2004), in the absence of international cooperation, countries play a Nash game in setting the degree of patent protection. The optimal degree

of patent protection for a country depends on the degrees of patent protection of all other countries in the world that trade with it. The best response function of a country is obtained by choosing the degree of protection such that the marginal cost of extending protection is equal to the marginal benefit of extending protection, given the degrees of protection of all other countries. Consider the choice of P_i , the degree of enforcement of patent protection at any moment in country i , by the government of that country. We can think of this degree as the fraction of country i 's market where patent protection is enforced. It is an index of the degree of patent protection. This country bears two costs from increasing the degree of enforcing patents slightly. First, it increases the fraction of the market that suffers a static deadweight loss of $C_c^i - C_m^i - \pi_i$ per consumer on each differentiated good *invented and sold in country i* , where C_c^i is the consumer surplus per consumer in the part of country i 's market where pricing is competitive because patents are not enforced; C_m^i is the consumer surplus per consumer in the part of country i 's market where the patent-holder can price with market power because patents are enforced; π_i is the profit per consumer earned by a typical patent-holding firm from country i in the part of the market where patents are enforced. Second, a slight increase in the degree of patent protection enlarges the fraction of the market where each of its consumers realizes surplus of only C_m^i instead of C_c^i on each good that was *invented in other countries but sold in country i* . Notice that the profits earned by foreign producers in country i are not an offset to this latter marginal cost, because they accrue to patent holders in foreign countries. On the other hand, the marginal benefit that comes to country i from strengthening patent protection reflects the increased incentive that foreign and domestic firms have to engage in R&D. If the welfare-maximizing degree of patent protection at time t , P_{it} , is positive and less than 1, then the marginal benefit per consumer of increasing P_{it} must match the marginal cost, which implies the following best response function

$$\bar{T} \left[I_{it}(C_c^i - C_m^i - \pi_i) + \left(\sum_{j \neq i} I_{jt} \right) (C_c^i - C_m^i) \right] = \left(\sum_{j \in \mathcal{N}} \frac{dI_{jt}}{dP_{it}} \right) \cdot [(1 - P_{it})\bar{T}C_c^i + P_{it}\bar{T}C_m^i],$$

where I_{it} is the number of inventions made by residents of country i , and \mathcal{N} is the set of all countries in the world that produce or trade patent-sensitive goods. Because our sample

contains all the major countries that produce and trade patent-sensitive goods, it should be a good proxy for \mathcal{N} .¹ Therefore, we shall treat \mathcal{N} as the set of our sample countries hereinafter. It is assumed that a firm in country i makes the same profit per consumer π_i regardless of where it sells its product. This would be true if manufacturing of a good is done in the country where it is invented using that country's labor and capital. \bar{T} is the present discounted value of one dollar over the life time, $\bar{\tau}$, of a product, or $\int_0^{\bar{\tau}} e^{-\rho t} dt$.

The above best response function further implies that

$$\bar{T} \left[I_{it}(C_c^i - C_m^i - \pi_i) + \left(\sum_{j \neq i} I_{jt} \right) (C_c^i - C_m^i) \right] = \left(\sum_{j \in \mathcal{N}} \frac{I_{jt}}{v_{jt}} \gamma_j \pi_j \right) M_{it} \bar{T} \left[(1 - P_{it}) \bar{T} C_c^i + P_{it} \bar{T} C_m^i \right],$$

where M_{it} is the number of consumers of patent-sensitive goods in country i ; $v_{jt} = \bar{T} \pi_j \sum_{k \in \mathcal{N}} M_{kt} P_{kt}$ is the value of a global patent for a firm from country j ; and γ_j is the elasticity of I_{jt} with respect to v_{jt} . That is, $\gamma_j = \frac{dI_{jt}}{dv_{jt}} \frac{v_{jt}}{I_{jt}}$. It stands for the responsiveness of innovation in country j to changes in the value of a patent (in elasticity form) held by a typical firm from country j . In deriving the above best response function, we assume that agents expect I_{it}/I_{jt} and M_{it}/M_{jt} to stay constant over time for all $i \neq j$. Once these assumptions are made, it is easy to show that the current year M_{it} and I_{it} values are sufficient for calculating the present discounted value of the marginal cost and benefit. The intersection of all the best response functions gives the set of equilibrium degrees of patent protection of the non-cooperative game.

We can re-write the above best response function as

$$\left(\sum_{j \in \mathcal{N}} I_{jt} \right) - \theta_2 I_{it} = \left(\sum_{j \in \mathcal{N}} \frac{I_{jt}}{v_{jt}} \gamma_j \pi_j \right) \bar{T} M_{it} (\theta_1 - P_{it}) \quad \text{for } i = 1, 2, \dots, N, \quad (1)$$

where $N =$ total number of countries in \mathcal{N} . It is assumed that $\theta_1 = \theta_1^i \equiv \frac{C_c^i}{C_c^i - C_m^i}$, $\theta_2 = \theta_2^i \equiv \frac{\pi_i}{C_c^i - C_m^i}$ for all i . This assumption would be valid if the elasticities of demand in

¹As long as the countries in \mathcal{N} that are excluded from our sample all have very small I so that $\sum_j I_j$ (where $j \in \{\text{sample countries}\}$) $\approx \sum_{j \in \mathcal{N}} I_j$, and the excluded countries are so insignificant that changes in M , I , and P of these countries have little effect on the IP protection of the included countries, we can take the set of our sample countries as a good proxy for \mathcal{N} .

all markets are equal. The above equation is the core equation of the theory. It can be easily seen that a country with larger market (M_i) or more patented inventions (I_i) will protect patents more when we compare across countries at a certain point in time. We could estimate the above system of simultaneous equations if we are willing to assume that γ_j are equal across countries. However, the errors incurred in doing so can be very large given the fact that we have a very diverse sample of countries, which can have very different γ_j . Therefore, we transform the simultaneous equation system (1) into another system which does not require us to make any assumption about γ_j . This is shown below. The cost of doing this will be discussed in subsection 3.2.

Setting $i = k$ in (1), summing over k on both sides for $k \in \mathcal{N}$, and then re-arranging, we obtain

$$\left(\sum_{j \in \mathcal{N}} \frac{I_{jt}}{v_{jt}} \gamma_j \pi_j \right) \bar{T} = \frac{(\sum_{j \in \mathcal{N}} I_j) (N - \theta_2)}{\sum_{k \in \mathcal{N}} M_k (\theta_1 - P_k)}$$

Hereinafter we drop the time subscript for simplicity of exposition. Substituting this expression into (1) for country i , and re-arranging, we obtain

$$\frac{(\sum_{j \in \mathcal{N}} I_j) (N - \theta_2)}{\sum_{k \in \mathcal{N}} M_k (\theta_1 - P_k)} = \frac{(\sum_{j \in \mathcal{N}} I_j) - \theta_2 I_i}{M_i (\theta_1 - P_i)} \quad i = 1, 2, \dots, N.$$

We shall denote the operator $\sum_{j \in \mathcal{N}}$ simply by \sum_j . The above equation can be re-written as

$$P_i = \theta_1 - \frac{\sum_k M_k (\theta_1 - P_k)}{M_i} \left[\frac{(\sum_j I_j) - \theta_2 I_i}{(\sum_j I_j) (N - \theta_2)} \right]. \quad i = 1, 2, \dots, N$$

which can be further re-written as

$$P_i = \theta_1 - \frac{B_i}{M_i} \sum_k M_k (\theta_1 - P_k) \quad i = 1, 2, \dots, N \quad (2)$$

where

$$B_i = \left[\frac{\sum_j I_j - \theta_2 I_i}{(\sum_j I_j) (N - \theta_2)} \right] = \frac{1 - \theta_2 \left(\frac{I_i}{\sum_j I_j} \right)}{N - \theta_2} \quad \text{with } 0 < \theta_2 < 1.$$

Equation (2) defines the equilibrium patent protection of country i as a function of the market size and innovative capability variables as well as other countries' levels of patent protection.² It can be rewritten as

$$P_i = \theta_1 - \frac{B_i}{1 - B_i} \sum_{k \neq i} \frac{M_k}{M_i} (\theta_1 - P_k). \quad (3)$$

Using a first order Taylor's series expansion, we can linearize equation (3) as follows:

$$P_i = \beta_0 + \beta_1 \left(\frac{\sum_{k \neq i} M_k}{M_i} \right) + \beta_2 \left(\frac{I_i}{\sum_j I_j} \right) + \rho \sum_{k \neq i} w_{ik} P_k \quad (4)$$

The derivation of the above equation is given in Appendix A. As M_i refers to the number of consumers of patent-sensitive goods and is a hypothetical variable on which we do not have data, we proxy for it using the total consumption of the patent-sensitive goods Con_i . The justification for it is discussed in Section 3.1.

As a result, eqt. (4) can be rewritten as

$$P_i = \beta_0 + \beta_1 \left(\frac{\sum_{k \neq i} Con_k}{Con_i} \right) + \beta_2 \left(\frac{I_i}{\sum_j I_j} \right) + \rho \sum_{k \neq i} w_{ik} P_k \quad (5)$$

Equation (5) constitutes the basis of the estimations and predictions of this paper. The terms associated with $\frac{\sum_{k \neq i} Con_k}{Con_i}$ and $\frac{I_i}{\sum_j I_j}$ mainly capture the impacts of market size and innovative capability on the equilibrium patent protection. The term $\sum_{k \neq i} w_{ik} P_k$ captures the interaction between the strength of patent protection in country i (P_i) and the strengths of patent protection in other countries ($\{P_k\}_{k \neq i}$). The extent of interaction of patent protection is governed by the weight

$$w_{ik} = \left(\frac{B_i}{1 - B_i} \right) \frac{Con_k}{Con_i} \quad \text{with} \quad B_i = \left[1 - \theta_2 \left(\frac{I_i}{\sum_j I_j} \right) \right] / (N - \theta_2). \quad (6)$$

²Of the N equations in (2) for $i = 1, 2, \dots, N$, only $N - 1$ are independent. The solution of a system consisting of the best response function of an arbitrarily chosen benchmark country plus the $N - 1$ independent equations of the form (2) for the rest of the sample is equivalent to the solution of the N best response functions. More on this will be discussed in subsection 3.2.

Since the weights w_{ik} consist of B_i which depends on an unknown parameter θ_2 with $0 < \theta_2 < 1$, we use three different values of θ_2 in the estimation. The values include $\theta_2 = 0.2$ (a high demand elasticity scenario), $\theta_2 = 0.5$ (a medium demand elasticity scenario) and $\theta_2 = 0.9$ (a low demand elasticity scenario). As the strengths of patent protection of all countries are determined simultaneously in equilibrium, the variable associated with the levels of patent protection of other countries ($\sum_{k \neq i} w_{ik} P_k$) on the right hand side of equation (5) will be correlated with the error term, similar to the presence of an endogenous variable. Because of this, this equation is estimated using a spatial econometric technique which explicitly accounts for the simultaneity embedded in the term $\sum_{k \neq i} w_{ik} P_k$. The details of the spatial econometric technique are discussed in Section 3.2. According to the derivation, β_1 should be negative while both β_2 and ρ should be positive.

3 Testing the Structural Model

In this section, we present a test of Grossman and Lai's (2004) model. To account for the interaction between countries in setting the degrees of patent protection, we adopt the approach of spatial econometrics. The empirical analysis is based on the equilibrium condition derived from Grossman and Lai's model which is given in equation (5). This equation has three important implications which are stated in the following three propositions:

Proposition 1 *When one compares across countries, the degree of patent protection of a country (P_i) increases with the market size of its patent sensitive industries relative to the total market size of its trade partners ($\frac{Con_i}{\sum_{k \neq i} Con_k}$), holding the relative innovative capability constant.*

Proposition 2 *The degree of patent protection of a country (P_i) increases with the country's relative innovative capability ($\frac{I_i}{\sum_j I_j}$), holding its relative market size constant.*

Proposition 3 *The welfare-maximizing policy makers in any country take notice of the strengths of patent protection in other countries when determining the degree of domestic patent protection.*

These three propositions are tested using equation (5). In this equation, proposition 1 implies that the coefficient of $\frac{\sum_{k \neq i} Con_k}{Con_i}$ (β_1) should be negative, proposition 2 implies that the coefficient of $\frac{I_i}{\sum_j I_j}$ (β_2) should be positive while proposition 3 implies that the spatial interactive coefficient ρ should be nonzero (in fact positive in the present case).

After carrying out the hypothesis testing in subsection 3.2, we calculate the in-sample and out-of-sample predictions of the model and gauge its performance in the various years in subsection 3.3. Besides comparing the mean squared errors of the predictions, we also investigate if there is any evidence that more countries adopted stronger patent protection than the level predicted by the model after the implementation of TRIPS.

3.1 Data

Years of study: 1980, 1985, 1990, 1995, 2000

What are patent-sensitive sectors?

The list of patent-sensitive sectors is given in Table 1. There are many different ways to choose the patent-sensitive sectors (see, for example, Maskus and Penubarti 1995, Lee and Mansfield 1996, and Smith 1999). Our choice of sectors is mainly based on Maskus (2000), but we decide to include the entire chemicals sector instead of just the sector of polymerization products and perfumes as in Maskus (2000).

Steady state and Time lag

The original model of Grossman and Lai (2004) is a dynamic one. It assumes that a government maximizes the present discounted value of the net social welfare of the country by choosing an optimal degree of patent protection, given the optimal strengths of patent protection chosen by all other countries. The Grossman-Lai model assumes

that M_i and I_i are constant over time for all countries. As a result, there is a steady-state dynamic equilibrium with endogenous, time-invariant $\{P_i\}_{i \in \mathcal{N}}$, which is a function of $\{M_i\}_{i \in \mathcal{N}}$ and $\{I_i\}_{i \in \mathcal{N}}$. In reality, the data show that M_i , I_i and P_i are all changing over time, which obviously indicates that we cannot assume the world was in a steady state. Nonetheless, we can justify the application of the model to analyze real world data by assuming that agents believe that M_i/M_j and I_i/I_j ($i \neq j$) are stable in the future. Once this assumption is made, it is easy to show that $\{M_i\}_{i \in \mathcal{N}}$ and $\{I_i\}_{i \in \mathcal{N}}$ of the current year (or previous year if we allow for time-lag) are sufficient for calculating the best response functions that determine the equilibrium $\{P_i\}_{i \in \mathcal{N}}$ in that year. This relationship is shown in equation (1).

For each year, we use the average of the previous four years of M_i and I_i as the proxies for the explanatory variables. There are a couple of justifications for using that. First, because of legislative lag, the patent laws implemented in a certain year is more likely to be determined by market conditions in the last few years than by conditions in the same year. This precludes the use of contemporaneous measures of M and I . Second, using lagged variables on the right hand side avoids endogeneity problem. For example, it is plausible that stronger patent protection (P_i) increases the contemporaneous number of patents granted to domestic residents (one measure of I_i). Such effect of the dependent variable on the explanatory variable causes endogeneity problem if contemporaneous values of the explanatory variables are used.

Data for the degree of patent protection

In the model, P_i is literally the percentage of market of country i that receives patent protection during the lifetime of the product. One can also interpret it as the probability that patent rights are enforced at any given moment throughout the life of the product, or the percentage of time patent is protected during the lifetime of the product in country i . Basically, this is the degree of enforcement of patent protection in a country.

An example of how this degree of protection can be captured in the real world is given by Ginarte and Park (1997) in their construction of the patent right index for a large

number of countries for 1960-1995 at five-year intervals. The Ginarte-Park (GP) index includes five aspects: (i) coverage of the patent laws in the country, (ii) membership in international agreements, (iii) the risks of having patent rights forfeited in the country, (iv) enforcement as stipulated by the law, and (v) duration of protection.

We use a linear monotonic function of the GP Index to proxy for P . The choice of the GP index to proxy for the degree of patent protection can be subject to a few criticisms, but we believe this index is the best we can get to proxy for patent rights, and we can defend the use of it in our analysis. First, although the GP index can stand for the degree of patent protection in a country, it is an ordinal measure: a doubling of the index does not mean a doubling of the probability that patent rights are enforced. Therefore, the GP index should only be a monotonic transformation of the true P in our model. To address this concern, we assume the true P is an affine transformation of the GP index. We can interpret this form as a linearized version of the Taylor series approximation of the true non-linear transformation function between the GP index and the true P . Specifically, we assume that the true probability of enforcement of patent protection, P_i , to be related to the GP index, GP_i , by the linear approximation $P_i = a + b \cdot GP_i$, where a is non-negative and b is a positive constant. It turns out that it does not matter whether we can observe the parameters a and b .

The second possible criticism of the use of the GP Index to stand for the degree of patent rights protection is that P should include not just the laws in the book, but also execution of the laws, which the GP index does not capture. Although execution of the laws is not captured by the index, Ginarte and Park argued from the evidence they gathered that “...the main complaints overall are not about the execution of patent laws, but of statutory and institutional differences which the indexes already reflect.” Park followed up on their 1997 study by updating the index for the year 2000. He also found that the index is strongly correlated (having a correlation coefficient of 0.8) with the Global Competitiveness Report’s intellectual property rights index, which is based on surveys of opinions of firms and executives about how patent protections are implemented

in different countries (Park and Wagh, 2002).

For the purpose of this study, the data for 1980, 1985, 1990, 1995 and 2000 are used.

Data for the market size

As M_i refers to the number of consumers of patent-sensitive goods and is a hypothetical variable on which we do not have data, we proxy for it using the total consumption of the patent-sensitive goods Con_i . The UNESCO has data that allows us to calculate Con_i . In the UNESCO dataset, it is defined that

$$\text{consumption} = \text{domestic production} + \text{imports} - \text{exports} \quad \text{in each country } i$$

or

$$Con_i = Prod_i + Im_i - Ex_i ,$$

For each country i , we sum up the dollar values of the consumption variable in this dataset for all the patent-sensitive sectors to arrive at the value for Con_i .

In using this variable as a proxy for M_i , we need to face the issue that this proxy for M_i is not the number of consumers or buyers but the expenditure by buyers. If prices are different across countries, then countries with lower price would show a higher expenditure even if the number of buyers are the same in all countries, as long as demand is elastic. We could actually estimate the price elasticity of demand if we had data on the relative prices of patent-sensitive goods in all countries. The problem is that there are no data on these price indexes. A priori, it is not clear whether, for example, a developing country (LDC) would have a higher relative price for patent-sensitive goods than a developed country (DC), since two counteracting effects are present. On the one hand, the prices of the non-traded goods should be lower in a developing country, since wages are much lower there. On the other hand, tariffs and trade barriers in developing countries are usually higher, which tend to raise the prices of tradable goods above those in the developed world. Without more information, we can only conclude that prices in LDCs are equally likely to be higher or lower than in DCs. Consequently, it could well be true that prices of patent-sensitive goods do not differ very much across countries on the average. As

mentioned above, we use the average of the previous four years' value of Con_i , to proxy for M_i .

Since the data essential for computing Con_i , viz. $Prod_i$, Im_i and Ex_i , are available only from a relatively small set of countries, we use instruments to proxy for this variable to augment the set of countries. To find such instruments, we regress the consumption of patent-sensitive goods on real GDP and real GDP per capita. We find that the R^2 is higher than 0.95 using pooled data for the years 1980, 1985 and 1990. So, we use the fitted expression as the proxy for the consumption of patent-sensitive goods in the country. A detail description of the steps is provided in Appendix C.

Data for I

In Grossman and Lai (2004), I is the number of useful commercializable patents obtained by the residents of a country in a given period. This reflects the number of patent-sensitive innovations and hence the innovative capability of the country. It is assumed in the model that all these innovations are equally valuable, and that a useful invention would be patented in all countries of the world.

We want to find a proxy for the number of patent-sensitive innovations that is comparable across countries. One option is to use an *output measure* — for example the number of patents granted to domestic residents. The shortcoming with this measure is that different countries have different criteria for granting patents. Some countries grant mini-patents or utility models while others do not. Another problem with using the number of patents granted to domestic residents is the potential simultaneity/endogeneity problem, i.e. the degree of patent protection also affects the number of patents granted to domestic residents. Nonetheless, to address this simultaneity problem, we use lagged measures of the number of patents granted to domestic residents (the average of the previous four years).

We use the number of patents granted to domestic residents obtained from World Intellectual Property Organization (WIPO) to proxy for the input measure of I . There are a few issues to be addressed when using this proxy. First, most patents are never put to

commercial use, and therefore should not be counted. Our response is that as long as the number of useful patents is approximately a constant fraction of total patents, our proxy would be fine. Second, there is home bias in application for patents. For example, US citizens tend to apply only for US patents but not patents from other countries. Therefore, the assumption that every invention would obtain patents from all countries of the world may not be true. Our response is that, given that all countries exhibits home-bias, as long as the proportion of inventions that only apply for home patents is constant across countries, our proxy is proportional to the number of inventions that obtain patents in all of the (major) markets in the world . Third, not all patents generate the same commercial value. Some patents are much more valuable than others. Thus, the assumption that all consumers spend the same amount on each patented good is not valid. Our response is that we are looking at the average patent. We can treat a particularly high-value patent as equivalent to several or many average patents, while a particularly low-value patent as equivalent to a fraction of an average patent. If we further assume that the values of patents have about the same mean and variance across all countries, then the equivalent number of equally-valuable commercializable patents in a country would be proportional to the number of commercializable patents in that country. In summary, if the assumptions mentioned in the first, second and third points above are not too far from the fact, then the equivalent number of equal-value commercializable inventions that are patented in all the major markets would be proportional to the number of patents granted to domestic residents. Like in the case of M , proportionality of the proxy to the real variable is all we need, as can be seen from the inspection of (5).

The other option to proxy for the number of patent-sensitive innovations is to use some *input measures* — such as the share of scientist and engineers in the labor force ($SciEng_{it}/LF_{it}$). The advantage of using such measure to proxy for the number of innovations is that they can avoid the endogeneity problem altogether. A higher $SciEng_{it}/LF_{it}$ signifies that a higher fraction of labor force are engaged in creative innovation as opposed to minor non-patentable innovations or routine manufacturing activities. In fact, in Gross-

man and Lai (2004), the ratio of the number of inventions between any two countries is equal to the ratio of the innovative capability (i.e. supply of human capital) between the two countries as long as the innovation production function is CES. Since the CES function is quite general, the theory is compatible with the use of Scientist and Engineers as a proxy for I . The data source is the UNESCO Statistical Yearbook, supplemented by the World Competitiveness Yearbook of the IMD.

In the estimation, we present the results for both the output and input measures of innovations.

3.2 Estimating the Structural Model Using Spatial Econometrics

One special contribution of the Grossman and Lai (2004) model is that it provides a specific functional form to describe how countries interacted in regard to patent protection and how the degrees of patent protection are related to the market sizes and levels of innovative capability of different countries. At the Nash equilibrium, the degrees of patent protection (P_i) across countries are determined by equation (5). This is the core equation that forms the basis of the spatial estimation. Given equation (5), the equilibrium patent protection of country i is characterized by:

$$P_i = \mathbf{X}_i \boldsymbol{\beta} + \rho \sum_{k \neq i} w_{ik} P_k + \varepsilon_i \quad (7)$$

where $\mathbf{X}_i = [1, \frac{I_i}{\sum_j I_j}, \frac{\sum_{j \neq i} Con_j}{Con_i}]$; $\boldsymbol{\beta} = [\beta_0, \beta_1, \beta_2]'$; $\varepsilon_i \sim N(0, \sigma^2)$; ρ is the “spatial” dependence parameter which measures the extent of strategic interdependence across countries and w_{ik} is the spatial weight assigned to the k^{th} trading partner of country i . Based on the structural model, w_{ik} is given by equation (6), with $0 < \theta_2 < 1$. As mentioned in Section 2, we do not attempt to estimate θ_2 as a parameter in the model. Instead, we tried three different values of θ_2 (0.2, 0.5 and 0.9) to conduct sensitivity analysis with respect to demand elasticity. The impact of changes in the market size and innovative capability on the equilibrium patent protection is investigated using the spatial equation

(7). The structural model in Section 2 suggests that the coefficient of $\frac{I_i}{\sum_j I_j}$ should be positive while the coefficient of $\frac{\sum_{j \neq i} Con_j}{Con_i}$ should be negative. In addition, the spatial parameter ρ should be positive.

Spatial estimation was first introduced by Paelinck and Klaassen in the early 1970s and further developed by Anselin in the 1980s. The main assumption of this model is that the endogenous variables have reached a state of equilibrium so that the structural parameters work uniformly across space. In terms of functional form, the spatial model is similar to the simultaneous equation model except that the spatial model can be considered as a special case of a simultaneous equation model with the number of observations being equal to the number of equations. Spatial Econometrics has been applied to economic studies in a variety of fields, including urban economics (e.g. Baranes and Tropeano (2003)), health economics (e.g. Mobley (2003)), public finance (e.g. Case, Rosen and Hines (1993), Hoyt (1993), Shroder (1996), Brueckner (1998), Figlio (1999) and Saavedra (2000)) and international economics (e.g. Kalnins (2003)).

This spatial system specifies how the change in the patent protection in any country affects the equilibrium patent protection of other countries. In other words, the spatial model setting implies that there is global interaction in the determination of patent protection. The welfare maximizing policy makers in any country take notice of the patent protection in other countries when determining the degree of domestic patent protection. The spatial model is also used to estimate how the equilibrium strengths of patent protection are impacted by shifts in the reaction functions due to the variations in the market size and innovative capability variables.

Given the setting of the spatial model, equation (7) can be written in matrix form as:

$$\mathbf{P}^N = \mathbf{X}^N \boldsymbol{\beta} + \rho \mathbf{W}^N \mathbf{P}^N + \boldsymbol{\varepsilon}^N \quad (8)$$

where \mathbf{P}^N is a $N \times 1$ vector of P_i ; \mathbf{X}^N is a $N \times 3$ matrix with row i being \mathbf{X}_i ; \mathbf{W}^N is a $N \times N$ matrix with the diagonal elements being 0 and the off-diagonal elements in row i and column k being $w_{ik} = \left(\frac{B_i}{1-B_i}\right) \frac{Con_k}{Con_i}$. The structure of the error term is specified as $\boldsymbol{\varepsilon}^N \sim N(\mathbf{0}, \Omega^N)$ with $\Omega^N = \sigma^2 \mathbf{1}_N$, where $\mathbf{1}_N$ is a $N \times N$ identity matrix.

However, note that equation (8) is originated from equation (2), the sum of which over all countries always gives an identity. Thus the system actually only consists of $N - 1$ linearly independent equations. This has an important econometric implication on the error terms as the sum of the error terms always equals to zero in the identity. This suggests that the variance-covariance matrix Ω^N is singular. To handle this singularity problem, we need to drop an equation from (8). We choose to drop the equation of the country whose patent protection index is closest to the mean, and this country turns out to be Philippines. We shall call it the benchmark country. We do not try to explain the level of patent protection of the benchmark country but will take it as given. We then estimate the remaining $N - 1$ linearly independent equations specified by equation (8) to obtain the estimates of β and ρ , taking into account the data of the benchmark country.³

By simple matrix manipulation, the $N - 1$ independent equations from (8) can be rewritten in the following reduced form:

$$\mathbf{A}_{-b}\mathbf{P}_{-b} = \mathbf{X}_{-b}\beta + \rho\mathbf{W}_bP_b + \boldsymbol{\varepsilon}_{-b} \quad (9)$$

where $\mathbf{A}_{-b} = (\mathbf{1} - \rho\mathbf{W}_{-b})$ is a $(N - 1) \times (N - 1)$ matrix with the diagonal elements being 1 and the off-diagonal elements in row i and column k being $a_{ik, k \neq i} = -\rho w_{ik}$, $\forall i, k \neq$ the benchmark country. \mathbf{P}_{-b} is a $(N - 1) \times 1$ vector of the patent protection indices of all the countries except the benchmark country. \mathbf{W}_{-b} equals the original weighting matrix \mathbf{W}^N but with the row and column associated with the benchmark country excluded. \mathbf{X}_{-b} is equal to the original matrix \mathbf{X}^N but with the row associated with the benchmark country excluded. \mathbf{W}_b is a $(N - 1) \times 1$ vector that corresponds to the column of the original \mathbf{W}^N matrix that is associated with the benchmark country, excluding the row of the benchmark country. P_b is the patent protection index of the benchmark country. $\boldsymbol{\varepsilon}_{-b}$ is a $(N - 1) \times 1$ vector of error terms and Ω_{-b} is the corresponding variance-covariance matrix. To simplify the notations, we will thereafter define $\mathbf{A} \equiv \mathbf{A}_{-b}$, $\mathbf{W} \equiv \mathbf{W}_{-b}$, $\mathbf{P} \equiv \mathbf{P}_{-b}$, $\mathbf{X} \equiv \mathbf{X}_{-b}$, $\boldsymbol{\varepsilon} \equiv \boldsymbol{\varepsilon}_{-b}$

³We could have estimated the best response function (1) directly and be able to make use of all N observations in our sample. The cost of doing this is that we have to assume that γ_j is uniform across countries. We believe the error incurred in doing so is even larger than adopting the current approach.

and $\Omega \equiv \Omega_{-b}$.

The parameters $[\boldsymbol{\beta}, \rho, \Omega]$ can be estimated by maximizing the following log-likelihood function for the joint distribution of \mathbf{P} :

$$\begin{aligned} \underset{\boldsymbol{\beta}, \rho, \Omega}{Max} \ln L &= -((N - 1)/2) \ln(\pi) - (1/2) \ln |\Omega| + \ln |\mathbf{A}| \\ &\quad - \frac{1}{2} (\mathbf{A}\mathbf{P} - \mathbf{X}\boldsymbol{\beta} - \rho \mathbf{W}_b P_b)' \Omega^{-1} (\mathbf{A}\mathbf{P} - \mathbf{X}\boldsymbol{\beta} - \rho \mathbf{W}_b P_b) \end{aligned} \quad (10)$$

The maximization of the log-likelihood can be reduced to a simpler expression in the form of a concentrated likelihood function. The detailed steps in maximizing the concentrated likelihood are available in Anselin (1986) and summarized in Appendix B.⁴

3.2.1 Estimation Results Using Output Measure of Innovative Capability

The estimation results for the year 1980, 1985, 1990, 1995 and 2000 with $\theta_2 = 0.2, 0.5$ and 0.9 are provided in Table 2, Table 3 and Table 4 respectively. The innovative capability measure (I_i) in these estimation are based on the output measure, that is, the number of patents granted to domestic residents. The estimation show that for the year 1980, 1985 and 1990, the coefficients of $\frac{\phi_i}{\sum_j \phi_j}$ (the output measure of the innovative capability $\frac{I_i}{\sum_j I_j}$) and the spatial parameter ρ are both significantly positive, which are consistent with the implications of the structural model. The coefficients of $\frac{\sum_{j \neq i} Con_j}{Con_i}$ are significantly negative at the 5% significance level for all the years prior to and including 1995. These findings are consistent with our hypothesis that the degrees of patent protection across countries are determined by a non-cooperative game Nash Equilibrium before the TRIPS Agreement was signed in 1994. For the year 2000, the estimation results show that the spatial parameter ρ becomes insignificant even at the 10 percent significance level. This indicates that the non-cooperative game model is less applicable after the TRIPS Agreement was signed. These conclusions are robust to the variation of θ_2 .

Similar conclusions are reached when we pool the data of 1980, 1985 and 1990 together in one group and the data of 1995 and 2000 together in another group in the estimation

⁴Refer to Chapter 12 of Anselin (1986).

of the model. The estimation results for the case of $\theta_2 = 0.5$ are reported in Table 5.⁵ For the year 1980, 1985 and 1990, the coefficients are all significant and have signs that are consistent with the structural model. For the year 1995 and 2000, the coefficient of $\frac{\sum_{j \neq i} Con_j}{Con_i}$ is insignificant at even the 10 percent significance level, suggesting that the implication of the non-cooperative game model on the market size does not hold in these years. The spatial parameter even has the opposite sign to the predicted one.

In the pooled estimation, a fixed effect model is used. In addition, we use Zellner's SURE setting in the error term to allow for non-zero correlation between the error terms of a country in different time periods. The correlation structure across time is specified by the matrix Σ . For a three period pooled estimation $\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix}$. The structure of the error term is thus given by $\varepsilon \sim N(\mathbf{0}, \Omega)$ with $\Omega = \Sigma \otimes \mathbf{1}_{N-1}$.⁶

3.2.2 Estimation Results Using Input Measure of Innovative Capability

To check if the results above are robust to alternative measure of innovative capability (I_i), we have also estimated the model using an input measure of innovative capability, that is, the share of scientists and engineers in the labor force ($SciEngLF_i$). The estimation results for the pooled sample of 1980, 1985 and 1990 as well as the pooled sample of 1995 and 2000 are reported in Table 6. All the results are based on case of $\theta_2 = 0.5$. We find that the results are similar to those based on the output measure of innovative capability. For the year 1980, 1985 and 1990, the coefficients are all significant at the 5 percent significance level and have signs that are consistent with the structural model. For the year 1995 and 2000, both the coefficient of $\frac{\sum_{j \neq i} Con_j}{Con_i}$ and the spatial parameter are insignificant at even the 10 percent significance level, again suggesting that the non-cooperative game model is less applicable in these years.

Next, we calculate the predicted values of P_i based on our estimated coefficients. By

⁵The results are similar for all the cases of $\theta_2 = 0.2, 0.5$ and 0.9 . The estimation results for the case of $\theta_2 = 0.5$ are provided in this paper. The estimation results for the cases of $\theta_2 = 0.2$ and $\theta_2 = 0.9$ are available upon request.

⁶In the pooled estimation, the data is first grouped by country and then by time period.

doing this, we can gauge the predictive power of the model. Moreover, we can find out whether more countries behaved more cooperatively than predicted by the model after the implementation of the TRIPS, as we conjectured.

3.3 Testing the Predictive Power of the Structural Model

3.3.1 In-Sample Prediction

To obtain predictions based on the structural model, we make use of the reduced form of the spatial representation of equation (9). Equation (9) implies

$$\mathbf{P} = \mathbf{A}^{-1}\mathbf{X}\boldsymbol{\beta} + \mathbf{A}^{-1}\rho\mathbf{W}_bP_b + \mathbf{A}^{-1}\boldsymbol{\varepsilon} \quad (11)$$

Equation (11) is the equation used in both the in-sample and out-of-sample prediction in this paper.

The in-sample prediction of the patent protection is based on the estimate of the structural model for the year 1980, 1985 and 1990. The mean squared error ($MSE_{in,80,85,90}$) equals to 0.0273 (see Table 7). To test whether the $MSE_{in,80,85,90}$ is too high (i.e. whether the structural model can explain the variation of the data), we use bootstrap method to infer the statistical significance of the MSE. First of all, we calculate the estimated $MSE_{in,80,85,90}$ as described below (to simplify the notation, we define $\{T_1\} = \{1980, 1985, 1990\}$):

$$\begin{aligned} \widehat{MSE}_{in,\{T_1\}} &= \frac{1}{n}\widehat{\alpha}'_{in,\{T_1\}}\widehat{\alpha}_{in,\{T_1\}} \\ \text{where } \widehat{\alpha}_{in,\{T_1\}} &= P_{\{T_1\}} - \widehat{P}_{\{T_1\}}^{in} \text{ with } \widehat{P}_{\{T_1\}}^{in} = f(I_{\{T_1\}}, Con_{\{T_1\}}, P_{b,\{T_1\}}|\widehat{\theta}_{\{T_1\}}) \\ &\text{and } n \text{ is the total number of observations.} \end{aligned}$$

Then we re-sample the data as described by the bootstrap method explained in Appendix D. We define the bootstrap MSE as

$$\begin{aligned} \widehat{MSE}_{in,\{T_1\}}^{bs} &= \frac{1}{n}(\widehat{\alpha}_{in,\{T_1\}}^{bs} - \widehat{\alpha}_{in,\{T_1\}})'(\widehat{\alpha}_{in,\{T_1\}}^{bs} - \widehat{\alpha}_{in,\{T_1\}}) \\ &\text{where } \widehat{\alpha}_{in,\{T_1\}}^{bs} \text{ is the bootstrap version of } \widehat{\alpha}_{in,\{T_1\}}. \end{aligned}$$

The $\widehat{MSE}_{in,\{T_1\}}$ is then compared with the bootstrapped distribution of $\widehat{MSE}_{in,\{T_1\}}^{bs}$ to test for the statistical significance of the MSE. The p-value is defined as

$$\hat{p}_{in} = \frac{1}{B} \sum_{bs=1}^B I\{\widehat{MSE}_{in,\{T_1\}}^{bs} > \widehat{MSE}_{in,\{T_1\}}\} \quad (12)$$

where $I\{\widehat{MSE}_{in,\{T_1\}}^{bs} > \widehat{MSE}_{in,\{T_1\}}\}$ is an indicator function for the event $\{\widehat{MSE}_{in,\{T_1\}}^{bs} > \widehat{MSE}_{in,\{T_1\}}\}$. A small p-value obtained from equation (12) would imply that the $\widehat{MSE}_{in,\{T_1\}}$ is “too high” and the structural model would fail to explain the variation in the patent protection data. The in-sample MSE and the bootstrap result are reported in Table 7. The p value for the $\widehat{MSE}_{in,\{T_1\}}$ is 0.1010 and hence the null hypothesis that the MSE is not too high cannot be rejected at the 10% significance level. This provides an evidence that the structural model is able to explain the variation in the data.

3.3.2 Out-of-Sample Prediction

To perform the out-of-sample prediction, we first of all have to establish that the coefficients are stable over the period used for the estimation and the period used for the prediction. An F-test on the equivalence of the time-specific fixed effect coefficients over the period of 1980-1990 ($H_0 : \beta_{0,80} = \beta_{0,85} = \beta_{0,90}$) is not rejected. So we can assume that the structural coefficients are stable over the period 1980-1990 and we can proceed to perform out-of-sample prediction.

In this paper, we use the rolling window approach to examine the one-period ahead prediction performance. We first estimate the structural model using the data of 1980 and 1985 and perform out-of-sample prediction for 1990. The estimated MSE ($\widehat{MSE}_{out,90}$) equals 0.0229 (see Table 7). It is quite close to the MSE when we did an in-sample prediction for the whole period of 1980-1990. This is consistent with our conjecture that the countries are all playing a non-cooperative game before the TRIPS Agreement was signed in 1994. We then try to predict the levels of patent protection for different countries in 1995 *assuming* the countries are still playing a non-cooperative game. Since the TRIPS

Agreement may introduce a major change from non-cooperative to more cooperative behavior, we expect the MSE of the out-of-sample prediction for 1995 using the estimates from 1985 and 1990 to be significantly higher. The prediction results reported in Table 7 show that this is indeed the case. The estimated MSE for 1995 ($\widehat{MSE}_{out,95}$) is equal to 0.0301, which is more than thirty percent higher than that for the out-of-sample prediction for 1990. To formally test whether the out-of-sample prediction for 1990 is significantly better than the prediction for 1995, we use the test statistic of White (2000) which is defined as:

$$H_n = \widehat{\delta}\sqrt{n} \text{ where } \widehat{\delta} = \widehat{MSE}_{out,95} - \widehat{MSE}_{out,90} \text{ and } n = N - 1.$$

This can be evaluated using the bootstrap method again. We re-sample the data for many times to get a set of bootstrapped test statistics for $bs = 1, 2, \dots, B$:

$$H_{bs,n} = \widetilde{\delta}_{bs}\sqrt{n}$$

where $\widetilde{\delta}_{bs}$ is the re-centered value of the bootstrap version of $\widehat{\delta}$:

$$\widetilde{\delta}_{bs} = \widehat{\delta}_{bs} - h(\widehat{\delta})$$

Since Hansen (2003) proves that using $h(\widehat{\delta}) = \widehat{\delta}$ as originally suggested by White will cause finite sample power problems, he suggests using the following truncated function:

$$h(\widehat{\delta}) = \begin{cases} \widehat{\delta} & \text{if } \widehat{\delta} \geq -A \\ 0 & \text{otherwise} \end{cases}$$

where $A = \frac{1}{4}n^{-1/4}\widehat{\sigma}(n^{1/2}\widehat{\delta})$ with $\widehat{\sigma}^2(n^{1/2}\widehat{\delta}) = \frac{1}{B} \sum_{bs=1}^B (\sqrt{n}\widehat{\delta}_{bs} - \sqrt{n}\widehat{\delta})^2$.

To test if the out-of-sample prediction for 1990 is significantly better than the prediction for 1995 (i.e. whether the null hypothesis $H_0 : MSE_{out,90} \geq MSE_{out,95}$ can be rejected), we compare the test statistic H_n with the set of bootstrapped test statistic $H_{bs,n}$ to obtain the p-value for testing the null hypothesis. The p value is calculated as:

$$\widehat{p}_{out} = \frac{1}{B} \sum_{bs=1}^B I\{H_{bs,n} > H_n\}$$

A p-value close to 0 would mean the rejection of the null hypothesis. As reported in Table 7, $\hat{\delta} = \widehat{MSE}_{out,95} - \widehat{MSE}_{out,90} = 0.0072$ and its p value is 0.0750, so the null hypothesis is rejected at 10% level of significance and we conclude that the out-of-sample prediction for 1990 significantly out-perform the out-of-sample prediction for 1995.

In addition, we expect the actual patent protection after the TRIPS Agreement was signed will be in general higher than the predicted patent protection of the non-cooperative game model. To study this, we calculate the percentage of observations which have a higher patent protection than what is predicted by the non-cooperative game model ($P_{i,.} > \hat{P}_{i,.}^{out}$). The results in Table 9 show that it is 55.73% for the out-of-sample prediction of 1990 but jumps significantly to 68.51% (78.68%) for the out-of-sample prediction of 1995 (2000). This confirms our conjecture that more countries behaved more cooperatively after the TRIPS agreement was signed.

4 Extension: Effects of Trade on the interdependence of IP protection

One key element of the Grossman and Lai model is the interdependence of IP protection among countries. This intensity of dependence of country i 's degree of IP protection on the degree of IP protection of country k is captured by the term w_{ik} in equation (5), which is derived based on free trade among all countries. Suppose we take into account the existence of trade barriers between countries, how should we modify the expression for w_{ik} in (6)? A simple way to take into account trade barriers between countries is to adjust w_{ik} according to the propensity of country k to import goods from i , as well as the propensity of country i to import from k . This is because the importance of country k as a market for innovative goods all over the world has to be adjusted by how easy foreign goods are imported into the country, which is positively related to country k 's propensity to import. Moreover, w_{ik} should also be adjusted to capture the fact that B_i

in equation (6) is positively related to country i 's propensity to import.⁷ Thus, w_{ik} should be positively related to both i 's propensity to import from k and k 's propensity to import from i .

To calculate the propensity to import, we do not use any tariff or trade-barriers data, as tariff by itself is considered as an insufficient measure of overall trade protection, while availability of reliable measures on non-tariff barriers is very limited. Moreover, general measures of non-tariff barriers may not be appropriate for the circumstances. Therefore, we develop a measure of propensity to import based on the actual data on trade flows and production. To derive the augmented model, we first of all define a few variables. Let η_{ji} be the fraction of goods invented in country j that are sold in i , η_{ii} be the fraction of goods invented in country i that are sold in i (it is not necessarily equal to one as some goods that are invented end up do not have markets), and Im_{ij} be the value of trade flow from country i to country j in patent-sensitive goods. If there is no overlap of inventions between any two countries, then

$$\text{Im}_{ik} = I_i \eta_{ik} M_k. \quad (13)$$

Moreover,

$$\begin{aligned} \text{Con}_i &= \left(\sum_{j \in \mathcal{N}} I_j \eta_{ji} \right) M_i \\ \text{Im}_i &= \sum_{j \neq i} \text{Im}_{ji} = \left(\sum_{j \neq i} I_j \eta_{ji} \right) M_i \end{aligned} \quad (14)$$

From equation (13) we can write

$$\eta_{ik} = \frac{\text{Im}_{ik}}{M_k I_i} \quad \text{for } k = 1, 2, \dots, i-1, i+1, \dots, N \quad (15)$$

From equations (14), we have

$$\text{Con}_k - \text{Im}_k = M_k I_k \eta_{kk} \quad \Rightarrow \quad \eta_{kk} = \frac{(\text{Con}_k - \text{Im}_k)}{M_k \cdot I_k}. \quad (16)$$

⁷When country i 's propensity to import is higher, the deadweight loss caused by an increase in IP protection is larger, which leads to a higher B_i , which in turn leads to a stronger effect of other countries' IP protection on country i 's IP protection.

If we define k 's propensity to import from i (Λ_{ik}) as the ratio $\frac{\eta_{ik}}{\eta_{kk}}$, then

$$\Lambda_{ik} \equiv \frac{\eta_{ik}}{\eta_{kk}} = \frac{\text{Im}_{ik} I_k}{(\text{Con}_k - \text{Im}_k) I_i} \quad \text{for } k \neq i \quad (17)$$

Eq. (17) indicates that k 's propensity to import from i can be measured using the trade flow from i to k (Im_{ik}), the total inventions of economy i and k (I_i and I_k) as well as the total consumption and imports of economy k (Con_k and Im_k) on which we have reliable data. The variable Λ_{ii} is by definition equal to one.

$\{\Lambda_{ik}\}_{k \neq i}$ and $\{\Lambda_{ki}\}_{k \neq i}$ are then used to construct three alternative measures that capture three aspects of the extent of trade relationship between country i and k . The three measures include:

- (i) a measure of the importance of country k as an importer of goods from country i $\left(\frac{\Lambda_{ik}}{\sum_j \Lambda_{ij}/N} \right)$.
- (ii) a measure of the importance of country k as an exporter of goods to country i $\left(\frac{\Lambda_{ki}}{\sum_j \Lambda_{ji}/N} \right)$.
- (iii) an average of measures (i) and (ii) $\left[\left(\frac{\Lambda_{ik}}{\sum_j \Lambda_{ij}/N} + \frac{\Lambda_{ki}}{\sum_j \Lambda_{ji}/N} \right) / 2 \right]$.

In the extended spatial model, the weight $\{w_{ik}\}_{k \neq i}$ that governs the interaction of patent protection between country i and k are augmented by these three alternative measures. The spatial model is thus modified from (5) to

$$P_i = \beta'_0 + \beta'_1 \left(\frac{\sum_{k \neq i} \text{Con}_k}{\text{Con}_i} \right) + \beta'_2 \left(\frac{I_i}{\sum_j I_j} \right) + \rho' \sum_{k \neq i} w'_{ik} P_k$$

where w'_{ik} has three specifications:

- (i) $w'_{ik} = \left(\frac{B_i}{1-B_i} \right) \frac{\text{Con}_k}{\text{Con}_i} \left(\frac{\Lambda_{ik}}{\sum_j \Lambda_{ij}/N} \right)$ with $B_i = \left[1 - \theta_2 \frac{I_i}{\sum_j I_j} \right] / [N - \theta_2]$
- (ii) $w'_{ik} = \left(\frac{B_i}{1-B_i} \right) \frac{\text{Con}_k}{\text{Con}_i} \left(\frac{\Lambda_{ki}}{\sum_j \Lambda_{ji}/N} \right)$ with $B_i = \left[1 - \theta_2 \frac{I_i}{\sum_j I_j} \right] / [N - \theta_2]$
- (iii) $w'_{ik} = \left(\frac{B_i}{1-B_i} \right) \frac{\text{Con}_k}{\text{Con}_i} \left[\left(\frac{\Lambda_{ik}}{\sum_j \Lambda_{ij}/N} + \frac{\Lambda_{ki}}{\sum_j \Lambda_{ji}/N} \right) / 2 \right]$ with $B_i = \left[1 - \theta_2 \frac{I_i}{\sum_j I_j} \right] / [N - \theta_2]$

The weights w'_{ik} in the interaction term in the augmented model now depends on the trade propensity variable. The theoretical model predicts that β'_1 should be negative while both β'_2 and ρ' should be positive.

The estimation results using the output measure of innovative capability are given in Table (10) to Table (12), corresponding to the three specifications of the augmented model

(model (i) to (iii)). The result for model (iii) using the input measure of innovative capability is given in Table 13.⁸ As in the free trade model, the results indicate that market size and innovative capability have significant positive impacts on the patent protection over the periods prior to the TRIPS Agreement. In addition, there is significant interaction in the determination of patent protection policy across different countries. These conclusions hold for all three specifications of the augmented model. In the period after the implementation of the TRIPS Agreement (1995 onwards), the coefficient associated with the interaction term becomes insignificant, which confirms that the non-cooperative model becomes less applicable in these periods.

Since alternative specifications of w_{ik} cannot falsify the theoretical model, we conclude that the model is robust to different measures of the degree of interdependence of IP protection between countries.

5 Conclusion

This paper validates the Grossman-Lai model in confirming that the cross sectional differences in patent protection in the pre-TRIPS period can be largely explained by three factors: the market sizes of the patent-sensitive industries, the levels of innovative capability and the interdependence of patent protection between countries. The structural model is written as a spatial model which allows for interdependence of patent protection between countries. Hypothesis testing on the coefficients of the spatial model during the year 1980, 1985 and 1990 shows that the patent protection of a country increases with the market size of its patent sensitive industry, innovative capability as well as the degree of patent protection of its trading partners, exactly as implied by the structural model. The same result is obtained whether input measure or output measure is used to proxy for innovative capability. With the signing of the TRIPS Agreement since 1994, the non-

⁸The estimation results for the case $\theta_2 = 0.5$ are reported. The main conclusions remain unchanged when θ_2 is changed to 0.2 or 0.9. For the estimation using input measure, we only include the result for model (iii) due to the limitation of space. Similar patterns are found for the other two models.

cooperative game model is expected to be less applicable in the year 1995 and 2000. The results indeed show that the some of the structural coefficients are no longer significant or even reverse in sign in these two years.

To evaluate the predictive power of the structural model, we then employ the structural estimates of the model to perform in-sample predictions of the patent protection for all the years before the TRIPS Agreement was signed. The p-value test on the significance of the mean squared errors (MSE) shows that the in-sample MSE of the structural model is not “too high”, meaning that the structural model performs well in predicting the variations of patent protection across countries. A rolling window approach is also used to test the one-period ahead out-of-sample predictive ability of the structural model. The result shows that the out-of-sample prediction of 1990 using the estimates of 1980 and 1985 significantly outperform the out-of-sample prediction of 1995 using the estimates of 1985 and 1990. This implies that the non-cooperative game model is more applicable before the signing of the TRIPS Agreement. This is consistent with the conjecture that the countries were playing non-cooperative game before the signing of the TRIPS Agreement but were behaving more cooperatively after the TRIPS Agreement was signed. Moreover, based on the one-period ahead out-of-sample prediction, we find that the percentage of countries whose actual patent protection exceeded the predicted patent protection increased from 55.73% in 1990 to 68.51% in 1995 and to 78.68% in 2000. This confirms our conjecture that more countries were practicing stronger patent protection in 1995 and 2000 than the degrees predicted by the non-cooperative game model.

An alternative measure of the degree of interdependence of patent protection that takes into account trade propensity between country pairs was constructed and used for estimation. It was found that the conclusions of the model are robust to the alternative measures of the degree of interdependence of patent protection between countries.

Appendix A: Derivation of the Spatial Model

We hereby derive the structural equation used in the estimation by linearizing the nonlinear structural equation (3). In this appendix, we show the detail steps in the linearization. Equation (3) gives

$$P_i = \theta_1 - \frac{B_i}{1 - B_i} \sum_{k \neq i} \frac{M_k}{M_i} (\theta_1 - P_k).$$

which is equivalent to

$$P_i = \theta_1 \left[1 - \frac{B_i}{1 - B_i} \sum_{k \neq i} \frac{M_k}{M_i} \right] + \frac{B_i}{1 - B_i} \sum_{k \neq i} \frac{M_k}{M_i} P_k \quad (18)$$

with

$$B_i = \left[\frac{\sum_j I_j - \theta_2 I_i}{(\sum_j I_j)(N - \theta_2)} \right] = \frac{1 - \theta_2 \left(\frac{I_i}{\sum_j I_j} \right)}{N - \theta_2} \text{ with } 0 < \theta_2 < 1.$$

Equation (18) involves two terms. The first term $\theta_1 \left[1 - \frac{B_i}{1 - B_i} \sum_{k \neq i} \frac{M_k}{M_i} \right]$ mainly captures the impact of market size and innovative capability on the equilibrium patent protection. This term is independent of the patent protection in other countries ($\{P_k\}_{k \neq i}$). To simplify the notation, we denote this term by $P_{0i} \left\{ \frac{\sum_{k \neq i} M_k}{M_i}, \frac{I_i}{\sum_j I_j} \right\}$. The second term is $\frac{B_i}{1 - B_i} \sum_{k \neq i} \frac{M_k}{M_i} P_k$ which measures the extent of interaction between the patent protection in country i (P_i) and the patent protection in other countries ($\{P_k\}_{k \neq i}$). The extent of interdependence of patent protection is governed by the weight $w_{ik} = \frac{B_i}{1 - B_i} \frac{M_k}{M_i}$, which can be easily shown to be directly related to country k 's relative market size $\frac{M_k}{M_i}$ and innovative capability (I_k).⁹ Using these notations, equation (18) can be written as:

$$P_i = P_{0i} \left\{ \frac{\sum_{k \neq i} M_k}{M_i}, \frac{I_i}{\sum_j I_j} \right\} + \sum_{k \neq i} w_{ik} P_k$$

⁹The positive relationship between w_{ik} and I_k can be easily shown by calculating $\frac{\partial \left(\frac{B_i}{1 - B_i} \right)}{\partial \left(\frac{I_i}{\sum_j I_j} \right)} \frac{\partial \left(\frac{I_i}{\sum_j I_j} \right)}{\partial I_k}$ which is strictly positive.

We linearize the term $P_{0i} \left\{ \frac{\sum_{k \neq i} M_k}{M_i}, \frac{I_i}{\sum_j I_j} \right\}$ around the mean values of the variates $\left\{ m_i \equiv \frac{\sum_{k \neq i} M_k}{M_i} \text{ and } l_i \equiv \frac{I_i}{\sum_j I_j} \right\}$ by using a first order Taylor's series expansion in order to investigate how this term reacts to changes in market size and innovative capability. We denote the mean values of m_i and l_i by \bar{m} and \bar{l} respectively. We keep the second term associated with the interaction of patent protection across countries in its original form without linearizing it because the structural model gives specific weights w_{ik} to the patent protection of different partners k . These weights are employed in the estimation of the structural model. Given the weights, equation (18) indicates that the patent protection of country i (P_i) is linear in the weighted patent protection of its partner countries ($\{P_k\}_{k \neq i}$). The linearized version of equation (3) is thus

$$P_i = P_{0i}\{\bar{m}, \bar{l}\} + \frac{\partial P_{0i}}{\partial m_i} \Big|_{\bar{m}, \bar{l}} (m_i - \bar{m}) + \frac{\partial P_{0i}}{\partial l_i} \Big|_{\bar{m}, \bar{l}} (l_i - \bar{l}) + \sum_{k \neq i} w_{ik} P_k + \varepsilon. \quad (19)$$

where

$$\begin{aligned} \frac{\partial P_{0i}}{\partial m_i} &= -\theta_1 \left(\frac{B_i}{1 - B_i} \right) \\ \frac{\partial P_{0i}}{\partial l_i} &= \frac{\partial \left(\frac{B_i}{1 - B_i} \right)}{\partial B_i} \left(\frac{\theta_2}{N - \theta_2} \right) \theta_1 m_i \\ &= \frac{1}{(1 - B_i)^2} \left(\frac{\theta_2}{N - \theta_2} \right) \theta_1 m_i \end{aligned}$$

Evaluating the first order partial derivatives at the points \bar{m} and \bar{l} , we get

$$\begin{aligned} \frac{\partial P_{0i}}{\partial m_i} \Big|_{\bar{m}, \bar{l}} &= -\theta_1 \left(\frac{B_i\{\bar{l}\}}{1 - B_i\{\bar{l}\}} \right) \equiv \beta_1 < 0 \\ \frac{\partial P_{0i}}{\partial l_i} \Big|_{\bar{m}, \bar{l}} &= \frac{\partial \left(\frac{B_i}{1 - B_i} \right)}{\partial B_i} \Big|_{\bar{l}} \left(\frac{\theta_2}{N - \theta_2} \right) \theta_1 \bar{m} \\ &= \frac{1}{(1 - B_i\{\bar{l}\})^2} \left(\frac{\theta_2}{N - \theta_2} \right) \theta_1 \bar{m} \equiv \beta_2 > 0 \end{aligned}$$

Hence the linearized version of equation (3) can be written as

$$P_i = \beta_0 + \beta_1 \frac{\sum_{k \neq i} M_k}{M_i} + \beta_2 \frac{I_i}{\sum I_j} + \rho \sum_{k \neq i} w_{ik} P_k$$

This is specified as equation (4) in Section 2. The structural model indicates that $\beta_1 < 0$, $\beta_2 > 0$ and $\rho > 0$.

Appendix B: Procedures for Estimating the Spatial Model

The model is $\mathbf{AP} = \mathbf{X}\boldsymbol{\beta} + \rho \mathbf{W}_b P_b + \boldsymbol{\varepsilon}$ and the full maximum likelihood is

$$\begin{aligned} \ln L = & -((N-1)/2) \ln(\pi) - ((N-1)/2) \ln \sigma^2 + \ln |\mathbf{A}| \\ & - \frac{1}{2\sigma^2} (\mathbf{AP} - \mathbf{X}\boldsymbol{\beta} - \rho \mathbf{W}_b P_b)' (\mathbf{AP} - \mathbf{X}\boldsymbol{\beta} - \rho \mathbf{W}_b P_b) \end{aligned} \quad (20)$$

where $\mathbf{A} = (\mathbf{1} - \rho \mathbf{W})$ and $\varepsilon_i \sim N(0, \sigma^2)$. The maximization procedure for the parameters $\theta \equiv [\rho, \boldsymbol{\beta}, \sigma^2]'$ can be simplified to a maximization problem in ρ only by using the following steps stated in Anselin (1988) Chapter 12:

1. Let $\mathbf{X}^* \equiv \begin{bmatrix} \mathbf{X} & \mathbf{W}_b P_b \end{bmatrix}$ where \mathbf{X} is a $(N-1) \times k$ matrix with k being the number of explanatory variables and $\mathbf{W}_b P_b$ is a $(N-1) \times 1$ vector. The estimator of $\boldsymbol{\beta}$ is given by the first k elements of the following vector:

$$(\mathbf{X}^{*'} \mathbf{X}^*)^{-1} (\mathbf{X}^{*'} \mathbf{AP}) = (\mathbf{X}^{*'} \mathbf{X}^*)^{-1} \mathbf{X}^{*'} \mathbf{P} - \rho (\mathbf{X}^{*'} \mathbf{X}^*)^{-1} \mathbf{X}^{*'} \mathbf{W} \mathbf{P} = \widehat{\boldsymbol{\beta}}_o^* - \rho \widehat{\boldsymbol{\beta}}_L^* \quad (21)$$

where the OLS estimate of $\widehat{\boldsymbol{\beta}}_o^*$ and $\widehat{\boldsymbol{\beta}}_L^*$ are obtained from a regression of \mathbf{P} on \mathbf{X}^* and $\mathbf{W} \mathbf{P}$ on \mathbf{X}^* respectively. Clearly, the ML estimate $\widehat{\boldsymbol{\beta}}$ (the first k elements of the vector in eqt.(21) above) is a function of the auxiliary linear regression coefficients $\widehat{\boldsymbol{\beta}}_o^*$ and $\widehat{\boldsymbol{\beta}}_L^*$ (the first k elements of $\widehat{\boldsymbol{\beta}}_o^*$ and $\widehat{\boldsymbol{\beta}}_L^*$ respectively) as well as of ρ . You can think of $\widehat{\boldsymbol{\beta}}_o^*$ and $\widehat{\boldsymbol{\beta}}_L^*$ as the coefficients of regressing \mathbf{P} and $\mathbf{W} \mathbf{P}$ on an orthogonalized \mathbf{X} (orthogonalized with respect to $\mathbf{W}_b P_b$). Note that whereas the estimate of ρ cannot be expressed analytically, neither $\widehat{\boldsymbol{\beta}}_o^*$ nor $\widehat{\boldsymbol{\beta}}_L^*$ are a function of any other parameters. Therefore the estimate for $\boldsymbol{\beta}$ can be found directly, once a value for ρ has been determined.

2. The two coefficient vectors $\widehat{\beta}_o^*$ and $\widehat{\beta}_L^*$ lead to two sets of residuals, e_o and e_L , which depends on \mathbf{X}^* , \mathbf{P} and \mathbf{WP} only:

$$\begin{aligned} e_o &= \mathbf{P} - \mathbf{X}^* \widehat{\beta}_o^* \\ e_L &= \mathbf{WP} - \mathbf{X}^* \widehat{\beta}_L^* \end{aligned}$$

3. Further application of the first order conditions and taking into account the auxiliary residuals yields the estimate for the error variance σ^2 as:

$$\sigma^2 = \left(\frac{1}{N-1}\right)(e_o - \rho e_L)'(e_o - \rho e_L)$$

Again, this estimate can be readily obtained once a value for ρ has been determined.

4. Substitution of the estimates of β and σ^2 into the likelihood results in a concentrated likelihood of the following form:

$$\ln L_C = C - \frac{N-1}{2} \ln \left[\left(\frac{1}{N-1}\right)(e_o - \rho e_L)'(e_o - \rho e_L) \right] + \ln |\mathbf{1}_{N-1} - \rho \mathbf{W}|$$

where C is the usual constant. This expression is a nonlinear function in one parameter only, ρ , and can be easily maximized by means of numerical techniques.

The ML estimates are asymptotically efficient. This means they achieve the Cramer-Rao lower variance bound, given by the inverse of the information matrix: $[I(\theta)]^{-1} = -E[\partial^2 L / \partial \theta \partial \theta']^{-1}$. The Hessian matrix $[\partial^2 L / \partial \theta \partial \theta']$ can be obtained from the full likelihood function

$$\begin{aligned} \ln L &= -\frac{N-1}{2} \ln(\pi) - \frac{N-1}{2} \ln \sigma^2 + \ln |\mathbf{A}| \\ &\quad - \frac{1}{2\sigma^2} (\mathbf{AP} - \mathbf{X}\beta - \rho \mathbf{W}_b P_b)' (\mathbf{AP} - \mathbf{X}\beta - \rho \mathbf{W}_b P_b) \end{aligned}$$

using standard numerical function of GAUSS.

Appendix C: Description of Data

This section describe in more detail how we come up with the data for market size and innovative capability (output measure is patent count; input measure is scientist and

engineers as a percentage of labor force) that we actually used in the regressions. The market size variable is measured in billion USD. The patent count is measured in 10,000. The patent protection index $P_{it} \in [0, 1]$ is taken as the Ginarte and Park index (which originally ranges between 0 and 5) divided by 5.

The actual data on the market size of the patent sensitive industries is obtained from the Industrial Demand-Supply Database of UNESCO. Since the data is only available for about 35 countries each year, we extend the set of countries in our sample to about 62 by proxying the market size using data on the real GDP and real GDP per capita. The procedure in constructing the market size variable and patent count variable is described below:

Firstly, we regress the actual market size data on the real GDP and the real GDP per capita for the countries where the actual market size data is available (about 35 countries per year)¹⁰. The estimated coefficients are then applied to obtain the fitted market size for all the countries where the real GDP and real GDP per capita data is available (about 62 countries per year).

Next we smooth out the short-term up-and-down fluctuations of the long-term underlying market size and innovative capability measure by calculating the four-year moving average of the fitted market size variable and the innovative capability variable respectively.

In addition, since both the market size variable and the innovative capability variable have persistent time trend and hence may not be stationary, we detrend the data by projecting the moving average of the fitted market size variable and the innovative capability variable on a set of time dummies. The de-trended values are used in all the estimations.

The sector-specific bilateral import data used in the augmented model is also taken from the Industrial Demand-Supply Database of UNESCO.

¹⁰The R^2 in the regressions are close to 0.9.

Appendix D: Bootstrap of the MSE

The algorithm of the bootstrap scheme is described below. Let $\hat{\theta} = g(\mathbf{X})$ be the estimator of θ based on the data set $\mathbf{X} = \{X_1, X_2, \dots, X_n\}$. The pseudo data series $\mathbf{X}^{bs} = \{X_1^{bs}, X_2^{bs}, \dots, X_n^{bs}\}$ is generated by drawing with replacement from the actual sample of size n . When the pseudo data series is generated, we compute $\hat{\theta}^{bs} = g(\mathbf{X}^{bs})$. The conditional sampling distribution of $\widehat{MSE}(\mathbf{X}^{bs}, \hat{\theta}^{bs})$ given \mathbf{X}^{bs} is the bootstrap approximation of the true sampling distribution of the true $MSE(\mathbf{X}, \hat{\theta})$. By simulating a large number B of the pseudo data series, the true distribution of MSE can be approximated by the empirical distribution of the B number of $\widehat{MSE}(\mathbf{X}^{bs}, \hat{\theta}^{bs})$. The bootstrap results in this paper are based on $B = 200$.

References

- [1] Acemoglu, Daron and Linn, Joshua (2003), "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry", NBER Working Paper.
- [2] Amemiya, T., (1971), "The estimation of the variances in a variance-components model", *International Economic Review*, 12, 1-13.
- [3] Anselin, Luc (1988), *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers.
- [4] Aneuryn-Evans, Gwyn; Deaton, Angus, "Testing Linear versus Logarithmic Regression Models", *The Review of Economic Studies*, Volume 47, Issue 1, Econometrics Issue, Jan 1980, 275-291.
- [5] Baltagi, Badi H. (2001), *Econometric Analysis of Panel Data*, 2nd Ed. Wiley.
- [6] Baranes, Edmond and Tropeano, Jean-Philippe (2003), "Why are Technological Spillovers Spatially Bounded? A Market Orientated Approach", *Regional Science and Urban Economics* 33, p.445-466.
- [7] Brueckner, Jan K. (1998), "Testing For Strategic Interaction Among Local Governments: The Case of Growth Controls", *Journal of Urban Economics* 44 (3), p.438-467.
- [8] Case, A.C.; Rosen, H.S. and Hines, J.R. (1993), "Budget Spillovers and Fiscal Policy Interdependence: Evidence from the States", *Journal of Public Economics* 52, p.285-307.
- [9] Figlio, D.N.; Kolpin V.W. and Reid, W. (1999), "Do States Play Welfare Games?", *Journal of Urban Economics*, 46 (3), p.437-454.
- [10] Ginarte, Juan C. and Park, Walter G. (1997), "Determinants of Patent Rights: A Cross-National Study," *Research Policy* 26, 283-301.

- [11] Grossman, Gene M. and Lai, Edwin L.-C. (2004), "International Protection of Intellectual Property," manuscript, City University of Hong Kong and Princeton University, May 2004; forthcoming, *American Economic Review*.
- [12] Hansen, Peter Reinhard (2003), "A test for superior predictive ability", Working Paper, Department of Economics, Brown University.
- [13] Hoyt, W.H. (1993), "Tax Competition, Nash Equilibria and Residential Mobility", *Journal of Urban Economics* 34, p.358-379.
- [14] Kalnins, Arturs (2003), "Hamburger Prices and Spatial Econometrics", *Journal of Economics and Management Strategy* 12 (4), p.591-616.
- [15] Lee, Jeong-Yeon and Mansfield, Edwin (1996), "Intellectual Property Protection and U.S. Foreign Direct Investment," *The Review of Economics and Statistics* 78(2), 181-186.
- [16] Maskus, Keith E. (2000), *Intellectual Property Rights in the Global Economy*, Washington D.C.: The Institute for International Economics.
- [17] Maskus, Keith E. and Eby-Konan, Denise (1994), "Trade-Related Intellectual Property Rights: Issues and Explanatory Results," In *Analytical and Negotiating Issues in the Global Trading System*, ed. by Alan V. Deardorff and Robert M. Stern. Ann Arbor, MI: University of Michigan Press.
- [18] Maskus, Keith E. and Mohan Penubarti (1995), "How Trade-related are Intellectual Property Rights?" *Journal of International Economics* 39, 227-248.
- [19] McCalman, Phillip (2001), "Reaping What you Sow: An Empirical Analysis of International Patent Harmonization," *Journal of International Economics* 55, 161-185.
- [20] Mobley, Lee R. (2003), "Estimating Hospital Market Pricing: An Equilibrium Approach using Spatial Econometrics", *Regional Science and Urban Economics* 33, p.489-516.

- [21] Paelinck, Jean and Klaassen, L. (1979), *Spatial Econometrics*, Farnborough: Saxon House.
- [22] Park, Walter G. and Smita Wagh (2002), "Index of Patent Rights," in *Economic Freedom of the World: 2002 Annual Report*, Chapter 2.
- [23] Saavedra, L.A. (2000), "A Model of Welfare Competition with Evidence from AFDC", *Journal of Urban Economics*, 47 (2), March 2000.
- [24] Scotchmer, Suzanne (2004), "The Political Economy of Intellectual Property Treaties," *Journal of Law, Economics and Organizations*, 20 (2): pp.415-437.
- [25] Shroder, M. (1996), "Games the States Don't Play: Welfare Benefits and the Theory of Fiscal Federalism", *Review of Economics and Statistics* 77, p.183-191.
- [26] Smith, Pamela J. (1999), "Are Weak Patent Rights a Barrier to U.S. Exports?" *Journal of International Economics* 48, 151-177.
- [27] White, Halbert (2000), "A Reality Check for Data-snooping", *Econometrica* 68, 1097-1126.

Patent-sensitive sectors	
1. Chemicals	5. Electromedical machines
2. Special industry machines	6. Electric microcircuits
3. Metalworking machines	7. Measuring, control instruments
4. Data processing equipment	

Table 1: List of Patent-Sensitive Sectors

Dependent Variable: Patent Protection Index P_i ($\theta_2 = 0.2$)					
Variables	1980	1985	1990	1995	2000
β_0	0.8942 (12.132)**	0.9291 (11.820)**	0.9552 (11.015)**	1.1156 (13.9848)**	0.8862 (12.539)**
$\frac{\phi_i}{\sum_j \phi_j}$	0.0049 (1.8936)*	0.0038 (1.6884)*	0.0030 (1.6704)*	0.0024 (1.1961)	0.0382 (4.3913)**
$\frac{\sum_{j \neq i} Con_j}{Con_i}$	-0.0789 (-5.0377)**	-0.0746 (-4.8989)**	-0.0773 (-4.6752)**	-0.0877 (-3.8282)**	0.0197 (1.0132)
WPL	0.0037 (3.7352)**	0.0031 (3.4353)**	0.0033 (3.5402)**	0.0026 (2.4588)**	-0.0023 (-1.4893)
\bar{R}^2	0.4395	0.4211	0.4462	0.4320	0.3495
observations	61	61	61	61	61
<p>Note: The numbers in parentheses are the t-statistics. “**” means the t statistic is significant at the 5% level and “*” means the t statistic is significant at the 10% level. The variable ϕ is the number of patents granted to domestic residents and Con is the market size of the patent sensitive industries.</p>					

Table 2: Structural Estimates ($\theta_2=0.2$) based on output measure of innovation

Dependent Variable: Patent Protection Index P_i ($\theta_2 = 0.5$)					
Variables	1980	1985	1990	1995	2000
β_0	0.8968 (12.127)**	0.9318 (11.809)**	0.9584 (10.971)**	1.1208 (16.5807)**	0.8892 (12.639)**
$\frac{\phi_i}{\sum_j \phi_j}$	0.0051 (1.9436)*	0.0036 (1.8005)*	0.0029 (1.6344)*	0.0023 (1.1363)	0.0383 (4.4054)**
$\frac{\sum_{j \neq i} Con_j}{Con_i}$	-0.0797 (-5.0644)**	-0.0753 (-4.9205)**	-0.0781 (-4.6783)**	-0.0889 (-4.2947)**	0.0190 (0.9792)
WP_L	0.0038 (3.7746)**	0.0032 (3.4714)**	0.0033 (3.5543)**	0.0026 (2.6971)**	-0.0022 (-1.4355)
$\overline{R^2}$	0.4381	0.4192	0.4457	0.4300	0.3503
observations	61	61	61	61	61
Note: The numbers in parentheses are the t-statistics. “**” means the t statistic is significant at the 5% level and “*” means the t statistic is significant at the 10% level.					

Table 3: Structural Estimates ($\theta_2=0.5$) based on output measure of innovation

Dependent Variable: Patent Protection Index P_i ($\theta_2 = 0.9$)					
Variables	1980	1985	1990	1995	2000
β_0	0.9005 (12.119)**	0.9356 (11.793)**	0.9629 (10.909)**	1.1279 (21.0480)**	0.8934 (12.776)**
$\frac{\phi_i}{\sum_j \phi_j}$	0.0053 (2.0112)**	0.0034 (1.6975)*	0.0029 (1.6912)*	0.0022 (1.0684)	0.0383 (4.4243)**
$\frac{\sum_{j \neq i} Con_j}{Con_i}$	-0.0808 (-5.0992)**	-0.0763 (-4.9479)**	-0.0791 (-4.6801)**	-0.0906 (-4.9172)**	0.0180 (0.9315)
WP_L	0.0039 (3.8263)**	0.0032 (3.5184)**	0.0034 (3.5701)**	0.0026 (2.9900)**	-0.0020 (-1.3591)
$\overline{R^2}$	0.4360	0.4165	0.4451	0.4272	0.3513
observations	61	61	61	61	61
Note: The numbers in parentheses are the t-statistics. “**” means the t statistic is significant at the 5% level and “*” means the t statistic is significant at the 10% level.					

Table 4: Structural Estimates ($\theta_2=0.9$) based on output measure of innovation

Dependent Variable: Patent Protection Index P_i ($\theta_2 = 0.5$)		
Variables	1980, 1985, 1990	1995, 2000
$\beta_{0,80}$	0.6272 (9.3762)**	
$\beta_{0,85}$	0.6326 (9.4085)**	
$\beta_{0,90}$	0.6381 (9.4482)**	
$\beta_{0,95}$		0.8389 (11.6119)**
$\beta_{0,00}$		0.9049 (12.7908)**
$\frac{\phi_i}{\sum_j \phi_j}$	0.0048 (2.0298)**	0.0122 (2.2892)**
$\frac{\sum_{j \neq i} Con_j}{Con_i}$	-0.0798 (-4.4720)**	-0.0197 (-1.1571)
WP_L	0.0013 (2.1580)**	-0.0022 (-1.7243)*
\bar{R}^2	0.3851	0.3243
observations	183	122

Note: The numbers in parentheses are the t-statistics. “**” means the t statistic is significant at the 5% level and “*” means the t statistic is significant at the 10% level. The variable ϕ is the number of patents granted to domestic residents and Con is the market size of the patent sensitive industries.

Table 5: Structural Estimates using the Pooled Sample of 1980-1990 and 1995-2000, based on the Output Measure of Innovative Capability

Dependent Variable: Patent Protection Index P_i ($\theta_2 = 0.5$)		
Variables	1980, 1985, 1990	1995, 2000
$\beta_{0,80}$	0.7738 (13.3095)**	
$\beta_{0,85}$	0.7777 (13.2649)**	
$\beta_{0,90}$	0.7859 (13.2658)**	
$\beta_{0,95}$		0.8414 (10.5211)**
$\beta_{0,00}$		0.9170 (11.6611)**
$\frac{SciEngLF_i}{\sum_j SciEngLF_j}$	0.0020 (2.2504)**	0.0181 (2.5585)**
$\frac{\sum_{j \neq i} Con_j}{Con_i}$	-0.0530 (-5.0813)**	-0.0215 (-1.1145)
WP_L	0.0023 (3.1504)**	-0.0018 (-0.8700)
\bar{R}^2	0.4010	0.3224
observations	183	134
Note: The numbers in parentheses are the t-statistics. “**” means the t statistic is significant at the 5% level and “*” means the t statistic is significant at the 10% level. The variable <i>SciEngLF</i> is the share of scientists and engineers in the labor force and <i>Con</i> is the market size of the patent sensitive industries.		

Table 6: Structural Estimates using the Pooled Sample of 1980-1990 and 1995-2000, based on the Input Measure of Innovative Capability

In-Sample and Out-Of-Sample Predictions				
	$MSE_{in,80,85,90}$	$MSE_{out,90}$	$MSE_{out,95}$	$MSE_{out,95} - MSE_{out,90}$
MSE	0.0273	0.0229	0.0301	0.0072
p-value	(0.1010)			(0.0750)*
Note: The numbers in parentheses are the p-values.				

Table 7: MSE of the In-Sample and Out-Of-Sample Predictions

Percentage of $P_{i,80} > \widehat{P}_{i,80}^{in}$	Percentage of $P_{i,85} > \widehat{P}_{i,85}^{in}$	Percentage of $P_{i,90} > \widehat{P}_{i,90}^{in}$
55.73%	55.73%	54.09%

Table 8: Percentage of Cases with Actual Patent Protection exceeding the In-Sample Predicted Patent Protection

Percentage of $P_{i,90} > \widehat{P}_{i,90}^{out}$	Percentage of $P_{i,95} > \widehat{P}_{i,95}^{out}$	Percentage of $P_{i,00} > \widehat{P}_{i,00}^{out}$
55.73%	68.51%	78.68%

Table 9: Percentage of Cases with Actual Patent Protection exceeding the Out-of-sample Predicted Patent Protection

Dependent Variable: Patent Protection Index P_i ($\theta_2 = 0.5$)		
Variables	1980, 1985, 1990	1995, 2000
$\beta_{0,80}$	0.6858 (11.9702)**	
$\beta_{0,85}$	0.6916 (11.9983)**	
$\beta_{0,90}$	0.6973 (12.0242)**	
$\beta_{0,95}$		0.8568 (12.4707)**
$\beta_{0,00}$		0.9214 (13.8381)**
$\frac{\phi_i}{\sum_j \phi_j}$	0.0040 (1.8698)*	0.0085 (1.7412)*
$\frac{\sum_{j \neq i} Con_j}{Con_i}$	-0.0267 (-2.8079)**	-0.0300 (-2.0940)**
WP_L	0.0016 (1.6806)*	-0.0026 (-1.6157)
\overline{R}^2	0.3943	0.3039
observations	183	122
<p>Note: The numbers in parentheses are the t-statistics. “**” means the t statistic is significant at the 5% level and “*” means the t statistic is significant at the 10% level. The variable ϕ is the number of patents granted to domestic residents and Con is the market size of the patent sensitive industries.</p>		

Table 10: Spatial Model based on Output Measure of Innovative Capability and w-ik Augmented with Measure on the Importance of k as an importer from i, theta2=0.5, pooled sample

Dependent Variable: Patent Protection Index P_i ($\theta_2 = 0.5$)		
Variables	1980, 1985, 1990	1995, 2000
$\beta_{0,80}$	0.7336 (13.2930)**	
$\beta_{0,85}$	0.7401 (13.3283)**	
$\beta_{0,90}$	0.7464 (13.3679)**	
$\beta_{0,95}$		0.8932 (14.2585)**
$\beta_{0,00}$		0.9606 (15.9098)**
$\frac{\phi_i}{\sum_j \phi_j}$	0.0080 (1.8432)*	0.0112 (2.0832)**
$\frac{\sum_{j \neq i} Con_j}{Con_i}$	-0.0370 (-4.1309)**	-0.0356 (-2.6675)**
WP_L	0.0008 (1.8315)*	-0.0012 (-0.7847)
\overline{R}^2	0.3984	0.3294
observations	183	122
<p>Note: The numbers in parentheses are the t-statistics. “**” means the t statistic is significant at the 5% level and “*” means the t statistic is significant at the 10% level. The variable ϕ is the number of patents granted to domestic residents and Con is the market size of the patent sensitive industries.</p>		

Table 11: Spatial Model based on Output Measure of Innovative Capability and w-ik Augmented with Measure on the Importance of k as an exporter to i, theta2=0.5, pooled sample

Dependent Variable: Patent Protection Index P_i ($\theta_2 = 0.5$)		
Variables	1980, 1985, 1990	1995, 2000
$\beta_{0,80}$	0.8091 (13.4548)**	
$\beta_{0,85}$	0.8166 (13.5199)**	
$\beta_{0,90}$	0.8241 (13.5568)**	
$\beta_{0,95}$		0.8630 (12.8550)**
$\beta_{0,00}$		0.9284 (14.3052)**
$\frac{\phi_i}{\sum_j \phi_j}$	0.0060 (1.7700)*	0.0093 (1.8441)*
$\frac{\sum_{j \neq i} Con_j}{Con_i}$	-0.0534 (-5.2701)**	-0.0306 (-2.1656)**
WP_L	0.0026 (4.1794)**	-0.0029 (-1.5698)
$\overline{R^2}$	0.3929	0.3109
observations	183	122
<p>Note: The numbers in parentheses are the t-statistics. “**” means the t statistic is significant at the 5% level and “*” means the t statistic is significant at the 10% level. The variable ϕ is the number of patents granted to domestic residents and Con is the market size of the patent sensitive industries.</p>		

Table 12: Spatial Model based on the Output Measure of Innovative Capability and w-ik Augmented with Measure on the Average of the Importance of k as exporter to and importer from i, theta2=0.5, pooled sample

Dependent Variable: Patent Protection Index P_i ($\theta_2 = 0.5$)		
Variables	1980, 1985, 1990	1995, 2000
$\beta_{0,80}$	0.7572 (12.7026)**	
$\beta_{0,85}$	0.7611 (12.6646)**	
$\beta_{0,90}$	0.7691 (12.6866)**	
$\beta_{0,95}$		0.8393 (11.0235)**
$\beta_{0,00}$		0.9141 (12.2909)**
$\frac{SciEngLF_i}{\sum_j SciEngLF_j}$	0.0050 (1.6521)*	0.0164 (2.3727)**
$\frac{\sum_{j \neq i} Con_j}{Con_i}$	-0.0490 (-4.5267)**	-0.0241 (-1.4740)
WP_L	0.0019 (2.1550)**	-0.0024 (-1.6417)
\overline{R}^2	0.3986	0.2985
observations	183	134
<p>Note: The numbers in parentheses are the t-statistics. “***” means the t statistic is significant at the 5% level and “**” means the t statistic is significant at the 10% level. The variable <i>SciEngLF</i> is the number of patents granted to domestic residents and <i>Con</i> is the market size of the patent sensitive industries.</p>		

Table 13: Spatial Model based on the Input Measure of Innovative Capability and w-ik Augmented with Measure on the Average of the Importance of k as exporter to and importer from i, theta2=0.5, pooled sample