INTERNATIONAL PROTECTION OF INTELLECTUAL PROPERTY: AN EMPIRICAL INVESTIGATION*

Edwin L.-C. Lai[†], **Samuel Wong**[a] **and Isabel K. Yan**[‡]

December 2008

Abstract

In this paper, we examine the validity of Grossman and Lai's (2004) game-theoretic model of international patent protection which yields clear predictions of the variation in the optimal degrees of public patent protection across countries. Grossman and Lai show that the welfare-maximizing levels of IPR protection can differ across countries according to their market sizes and innovative capability, so the optimal degree of IPR enforcement is not necessarily full enforcement. Their model provides a specific functional form that takes into account the interdependence in the determination of the welfare maximizing levels of patent protection across countries. This paper uses data in 1980-2000 to (i) check the validity of the non-cooperative game model of IPR protection before and after the TRIPS agreement was enforced; (ii) perform one-step-ahead prediction of the level of IPR protection and contrast the prediction performance of the structural model with various alternative models in the literature; and (iii) assess the impact of the implementation of the TRIPS agreement on the IPR protection. A salient feature of this paper is that this paper directly examines the empirical implications of a structural model.

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 implementation of the TRIPS agreement in 1995, but becomes less so after 1995. A horse race on the predictive power of this tightly specified nonlinear structural model against the linearized version demonstrates that the linearized model, though being more general and is not subject to any structural restrictions on the parameters, does not consistently outperform the structural model which is subject to strict restrictions on both the parameters and functional form. However, a comparison of the structural model against alternative empirical models with additional regressors like institutional factors and trade openness reveal that these factors possess surprising abilities to explain large variations in cross-sectional IPR protection.

Keywords: intellectual property, patent, empirical study, structural model
JEL Classification Number: O34, F13

---

† Corresponding author. Federal Reserve Bank of Dallas, Research Department, P.O. Box 655906, Dallas, TX 75265-5906, USA. Phone: (214) 922-6941; Fax: +(214) 922-5194; E-mail: Edwin.L.Lai@gmail.com or Edwin.Lai@dal.frb.org

[a] Department of Statistics, the Chinese University of Hong Kong, Shatin, Hong Kong. Phone: +(852)2609-7942; Fax: +(852)2603-5188; Email: samwong@sta.cuhk.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: efyan@cityu.edu.hk.

∗ The authors would like to thank helpful comments from Gene Grossman, Tom Prusa, Isaac Ehrlich, Keith Maskus, Eric Bond, Giovanni Maggi, Stephen Redding and participants in the seminars in Princeton, Columbia, Rutgers, Georgia Tech, University at Buffalo, Hitotsubashi University and Singapore Management University. The work in this paper has been supported by the Research Grants Council of Hong Kong, China (Projects No. CityU 1145/99H and CityU 1476/05H) and the Research Center for International Economics at City University of Hong Kong. Lai would like to thank the International Economics Section of Princeton University for its support while he was a Visiting Fellow there.

INTERNATIONAL PROTECTION OF INTELLECTUAL PROPERTY: AN EMPIRICAL INVESTIGATION

1 Introduction

Intellectual property rights (IPR) protection 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 Aspects of Intellectual Property Rights) 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. The corresponding deadline for all developing and transition economies was January 1, 2000, and that for the poorest countries is January 1, 2006 (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 would have been 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.

The motivation of this paper is two-fold. First, given that structural models unavoidably have to be constructed based on some simplifying assumptions in order to facilitate analytical derivation, this paper aims at gauging the adequacy of the game-theoretic model of intellectual property protection developed by Grossman and Lai (2004) in explaining the variations of IPR protection around the globe. If the model is an appropriate representation of the data generation process of the variables being predicted, the structural model should be able to generate predictions that match the stylized facts of the data. In this paper, we focus on patent protection, and leave the research into other aspects of intellectual property (IP) protection, such as trademarks and copyrights, to later studies.[1] 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. Second, we are interested in the utility of the model for operational prediction and policy design. In particular, 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 shows that the socially optimal degree of IPR protection is not necessarily full enforcement, and it yields clear predictions of the variation in the optimal strengths of patent protection across countries based on the market sizes and levels of innovative capability of the countries. One salient feature of Grossman and Lai's structural model is that it takes into account interdependence in the determination of the strengths of patent protection between countries. We find that the

[1] The patent-sensitive industry is a huge one that warrants studying in its own rights. In 2000, the total output of patent-sensitive goods in the world was 3,535 billion US dollars.

pattern of patent protection around the globe is broadly consistent with the predictions of the non-cooperative game model in 1980, 1985 and 1990. For the sake of comparison, we find that the model is less applicable to the years 1995 and 2000. This is expected since global patent protection became more cooperative after the TRIPS agreement was signed.

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). However, despite the considerable theoretical effort in the literature, it is somewhat surprising that little empirical work has focused on assessing the empirical performance of the structural models. From a policy-making point of view, this is certainly a gap that needs to be filled because structural models are much more appropriate when used for policy analysis.

In Section 2, we briefly describe the theory, which basically draws from Grossman and Lai (2004). We then derive an estimable equation from the structural model. In Section 3, we describe the GMM-IV estimation procedure and discuss the results. Section 4 performs model validation by testing the out-of-sample predictive power of the structural model before and after the implementation of the TRIPS agreement and against other alternative models. Section 5 concludes.

2 The Theory

The theory we want to test comes directly from Grossman and Lai (2004). For details of the model, the reader could refer to the paper directly. Below 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.[2]

[2]Lerner (2002) has presented evidence that tightening of patent laws does not increase the amount of patenting activities by domestic firms, but does increase the amount of patenting activities by foreigners in the domestic market.

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.*

According to Grossman and Lai's (2004) theory, countries play a Nash game in setting the degree of patent protection in the absence of international cooperation. 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, or as the probability that the inventor of an innovation can enforce her patent in court and prevent imitation. It is an index of the degree of patent protection.[3] In Grossman and Lai's model, 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*

[3] A similar interpretation is used in Eicher and García-Peñalosa (2008).

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

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

where I_{it} is the number of inventions made by residents of country i, which captures its innovative capability, 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 $\mathcal{N}$.[4] 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. $\overline{T}$ is the present discounted value of one dollar over the life time, $\overline{\tau}$, of a product, or $\int_0^{\overline{\tau}} e^{-\rho t}dt$.

The above best response function further implies that

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

where M_{it} is the number of consumers of patent-sensitive goods in country i, which captures its market size; $v_{jt} = \overline{T}\pi_j\sum_{k\in\mathcal{N}} M_{kt}P_{kt}$ 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. The intersection of all the best response functions gives the set of equilibrium degrees of patent protection of the non-cooperative game.

Given $v_{jt} = \overline{T}\pi_j\sum_{k\in\mathcal{N}} M_{kt}P_{kt}$,

[4] As 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 $j \in$ {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 $\mathcal{N}$.

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

$$\iff \left(\sum_{k\in\mathcal{N}}M_{kt}P_{kt}\right)\left[\sum_{j\in\mathcal{N}}I_{it}-I_{it}\frac{\pi_i}{C_c^i-C_m^i}\right]=\left(\sum_{j\in\mathcal{N}}I_{jt}\gamma_j\right)M_{it}\left[\frac{C_c^i}{C_c^i-C_m^i}-P_{it}\right] \quad (1)$$

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 if $\frac{\pi_i}{C_c^i-C_m^i}$ and $\frac{C_c^i}{C_c^i-C_m^i}$ are equal across all i, which we shall assume to be the case.

Grouping the terms that involve P_{it} together,

$$\begin{aligned}\left(\sum_{k\neq i}M_{kt}P_{kt}\right)\left[\sum_{j\in\mathcal{N}}I_{it}-I_{it}\frac{\pi_i}{C_c^i-C_m^i}\right] &= -P_{it}\left[M_{it}\left(\sum_{j\in\mathcal{N}}I_{it}-I_{it}\frac{\pi_i}{C_c^i-C_m^i}+\sum_{j\in\mathcal{N}}I_{jt}\gamma_j\right)\right]\\ &\quad+\left(\sum_{j\in\mathcal{N}}I_{jt}\gamma_j\right)M_{it}\left[\frac{C_c^i}{C_c^i-C_m^i}\right]\end{aligned}$$

$$\iff P_{it}=\frac{-\left(\sum_{k\neq i}M_{kt}P_{kt}\right)\left[\sum_{j\in\mathcal{N}}I_{it}-I_{it}\frac{\pi_i}{C_c^i-C_m^i}\right]+\left(\sum_{j\in\mathcal{N}}I_{jt}\gamma_j\right)M_{it}\left[\frac{C_c^i}{C_c^i-C_m^i}\right]}{M_{it}\left(\sum_{j\in\mathcal{N}}I_{it}-I_{it}\frac{\pi_i}{C_c^i-C_m^i}+\sum_{j\in\mathcal{N}}I_{jt}\gamma_j\right)}$$

Divide through by $\sum_{j\in\mathcal{N}}M_{jt}$ and $\sum_{j\in\mathcal{N}}I_{jt}$, and let $\mu_i\equiv\frac{I_{it}}{\sum_{j\in\mathcal{N}}I_{it}}$ and $m_i\equiv\frac{M_{it}}{\sum_{j\in\mathcal{N}}M_{it}}$,

$$P_{it}=\frac{-\left(\sum_{k\neq i}m_{kt}P_{kt}\right)\left[1-\mu_{it}\frac{\pi_i}{C_c^i-C_m^i}\right]+\left(\sum_{j\in\mathcal{N}}\mu_{jt}\gamma_j\right)m_{it}\left[\frac{C_c^i}{C_c^i-C_m^i}\right]}{m_{it}\left(1-\mu_{it}\frac{\pi_i}{C_c^i-C_m^i}+\sum_{j\in\mathcal{N}}\mu_{jt}\gamma_j\right)}$$

$$\iff P_{it}=\frac{\left(\sum_{j\in\mathcal{N}}\mu_{jt}\gamma_j\right)\left[\frac{C_c^i}{C_c^i-C_m^i}\right]}{\left(1-\mu_{it}\frac{\pi_i}{C_c^i-C_m^i}+\sum_{j\in\mathcal{N}}\mu_{jt}\gamma_j\right)}-\frac{\left[1-\mu_{it}\frac{\pi_i}{C_c^i-C_m^i}\right]}{\left(1-\mu_{it}\frac{\pi_i}{C_c^i-C_m^i}+\sum_{j\in\mathcal{N}}\mu_{jt}\gamma_j\right)}\left(\sum_{k\neq i}\frac{m_{kt}}{m_{it}}P_{kt}\right)$$

Assume $\theta_1=\theta_1^i\equiv\frac{C_c^i}{C_c^i-C_m^i}$, $\theta_2=\theta_2^i\equiv\frac{\pi_i}{C_c^i-C_m^i}$ for all i with $\theta_1>1, 0<\theta_2<1$, and $\delta_t\equiv\sum_{j\in\mathcal{N}}\mu_{jt}\gamma_j$,

$$\begin{aligned}P_{it} &= \frac{\delta_t\theta_1}{(1-\mu_{it}\theta_2+\delta_t)}-\frac{[1-\mu_{it}\theta_2]}{(1-\mu_{it}\theta_2+\delta_t)}\left(\sum_{k\neq i}\frac{m_{kt}}{m_{it}}P_{kt}\right) \quad (2)\\ i &= 1,2,.....,N \text{ where } N=\text{total number of countries in }\mathcal{N}.\end{aligned}$$

for any given time t. This assumption of constant θ_1 and θ_2 would be valid if the elasticities of demand in all markets are close[5].

Eqt.(2) provides an estimable structural equation for country i in a system of N equations with N endogenous variables $\{P_{1t}, P_{2t}, ..., P_{Nt}\}$. The set of structural coefficients to be calibrated or estimated include $[\theta_1, \theta_2, \delta]$ for each period. $\mu_i \equiv \frac{I_i}{\sum_{j\in\mathcal{N}} I_i}$ and $m_i \equiv \frac{M_i}{\sum_{j\in\mathcal{N}} M_i}$ refer to the data that measure the innovative capability and market size of country i relative to the aggregate innovative capability and market size of the world. The symbol $\sum_{j\in\mathcal{N}}$ denotes the sum over the data of all countries in the same time period. The variable P_i refers to the strength of patent protection in country i.

At any given t, the set of equations for the N countries (i=1,2,.....,N) can be written in a matrix form as follows:

$$\mathbf{A}_t\mathbf{P}_t = \mathbf{b}_t$$

where $P_t \equiv [P_{1t}, P_{2t}, ..., P_{Nt}]'$ is a $N \times 1$ matrix; A_t is a $N \times N$ matrix with the diagonal elements equal to one and the $(i,k)^{th}$ off-diagonal element equals $\frac{[1-\mu_{it}\theta_2]}{(1-\mu_{it}\theta_2+\delta_t)}\left(\frac{m_{kt}}{m_{it}}\right)$. Also,

$$\mathbf{b}_t = \left[\frac{\delta_t\theta_1}{1-\mu_{1t}\theta_2+\delta_t}, \cdots, \frac{\delta_t\theta_1}{1-\mu_{Nt}\theta_2+\delta_t}\right]'$$

is a $N \times 1$ matrix.

The equilibrium levels of patent protection of the N countries based on the game theoretic model are characterized by:

$$\mathbf{P}_t = \mathbf{A}_t^{-1}\mathbf{b}_t$$

In the real world, instead of observing P_t, we observe the intellectual property right index G_t which is an index that is assumed to be a smooth strict monotone function $h(\cdot)$ of the unobservable P_t plus a measurement error ε_t:

$$\mathbf{G}_t = h(\mathbf{P}_t) + \boldsymbol{\varepsilon}_t$$

[5] θ_1^i and θ_2^i are related to the markups (mp_i), or equivalently, price elasticities (ϵ_i) across countries because $\theta_1^i = 1/\left[1-\left(\frac{\epsilon_i}{\epsilon_i-1}\right)^{1-\epsilon_i}\right]$ and $\theta_2^i = \left(\frac{\epsilon_i}{\epsilon_i-1}\right)^{-\epsilon_i} / \left[1-\left(\frac{\epsilon_i}{\epsilon_i-1}\right)^{1-\epsilon_i}\right]$, with $\epsilon_i = \frac{mp_i}{mp_i-1}$. Based on the values of ϵ_i or mp_i of 14 countries reported by various studies in the literature, the implied θ_1^i and θ_2^i are computed and reported in Appendix A. The results indicate that the values of θ_1^i and θ_2^i across countries are quite similar, which provides some supporting evidence for our assumption that θ_1 and θ_2 are constant across countries. In the model, we calibrate θ_1 and θ_2 using the average values over the 14 countries.

The use of smooth monotone function is particularly relevant here as the observed variable G_t is an index rather than an exact measurement for the intrinsic values of the unobservable variable P_t. Moreover, since ε_t could be correlated with the innovative capability and market size measures μ_{it} and m_{it} due to endogeneity, we make use of their lag values as instruments X_t. Therefore, the moment conditions are given by:

$$E\{\mathbf{G}_t - h(\mathbf{A}_t^{-1}\mathbf{b}_t)|X_t\} = 0 \tag{3}$$

Equation (3) constitutes the basis of the estimations and predictions of this paper. To estimate the unknown structural parameters δ together with the nonparametric function h, we employ the methods of Ramsay (1998) and Ai and Chen (2003). A detailed description of the procedure is provided in Appendix B. The estimated structural coefficients for 1980, 1985, 1990, 1995 and 2000 are reported in Table 2.

3 Estimation of the Structural Model

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).

Equilibrium and endogeneity

The original model of Grossman and Lai (2004) 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 derives the equilibrium for the endogenous $\{P_i\}_{i\in\mathcal{N}}$, which is a function of $\{M_i\}_{i\in\mathcal{N}}$ and $\{I_i\}_{i\in\mathcal{N}}$. In our empirical estimations, we use the average values of the innovative capability and market size data in the preceding three years as the proxies of I_i and M_i respectively. For the sake of robustness check, the estimation results based on the non-smoothed data of market size is also reported. There are two justifications for using the smoothed data of innovative capability and market size. First, because of legislative lag, the patent laws implemented in a certain year are more likely to be determined by the economic conditions in the recent past rather than just by the contemporaneous conditions. Second, taking average can smooth out any spurious time variations and lumpiness in the data in any particular year.

Moreover, it is plausible that stronger patent protection (P_i) increases the contemporaneous market size of the patent sensitive sector and encourages the strengthening of the innovation capability in the economy and hence leads to endogeneity in the M_i and I_i measures in the model. For this reason, the lags of M_i and I_i are used as instruments in the estimation. However, if the error term of the patent equation is serially correlated, using the smoothed data for M_i and I_i would lead to bias as the estimated coefficients would lose their consistency. In order to minimize the bias due to serial correlation, instruments that are sufficiently lagged need to be used. Therefore, we use the four-year-lag of the M_i and I_i variables as instruments in all our estimations. The four-year-lag is the longest lag for which the lag data for the first sample year (1980) is available. We decide not to take further lags because using longer lags will in general lower their correlation with the instrumented variables.

Data for the degree of patent protection

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

An index that is constructed to capture the degree of patent right protection in the real world is given by Ginarte and Park (1997). They later followed up on their 1997 study by updating the index for the year 2000. The Ginarte and Park patent right index (GP) is available for a large number of countries in 1960-2000 at every five-year interval. The Ginarte-Park 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 monotonic function to model the relationship between the observed GP Index and the unobservable variable P_i. However, a number of caveats are in order. First, although the GP index can stand for the degree of patent protection in a country, it is an index rather than an exact measurement for the intrinsic values of the unobservable variable P_t. To a certain extent, it is an ordinal measure and 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 that the observed GP index is a monotonic transformation of the true P. The monotone function is to be estimated using the nonparametric approach of Ramsay (1998) and Ai and Chen (2003).

Second, a possible criticism of the use of the GP Index to stand for the degree of patent rights protection is that P_i should include not just the laws in the book, but also execution of the laws, which the GP index does not capture.[6] Although the 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 also found that the index is strongly correlated with the Global Competitiveness Report (GCR)'s intellectual property rights index, which is based on surveys of opinions of firms and executives about how patent protections are actually implemented in different countries (Park and Wagh, 2002). Despite the difference in the survey method between the GCR and the GP patent rights ratings, a remarkably high simple correlation exists between the two measures. The correlation is about 0.8. This suggests a relatively strong match between the statutory levels of patent protection (i.e., laws on the books) and the perceived levels of protection for intellectual property among practitioners. However, since the GCR rating is only available from mid 1990's onwards, it is considered too short for our purpose and hence we do not employ it in our analysis.

Data for the market size

As M_i refers to the number of consumers of patent-sensitive goods and is a hypothetical variable for which we cannot observe directly, we proxy for it using the total consumption of the patent-sensitive goods Con_i.[7] The UNESCO has data that allows us to calculate Con_i. In the UNESCO dataset, it is defined that

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

[6] See for example Branstetter, Fisman and Foley (2006) for this argument.

[7] Similarly, Acemoglu and Linn (2004) use the consumption expenditure data to proxy for the market size.

or

$$\text{Con}_i = \text{Prod}_i + \text{Im}_i - \text{Ex}_i \ .$$

For each country i, we sum up the dollar values of the consumption variable in this dataset for all the patent-sensitive industries 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 lower expenditure even if the number of buyers are the same in all countries. Therefore, without adjusting for price differences across countries, Con_i would overestimate (underestimate) M_i in a high- (low-) price country. To address this concern, we carry out our regressions using the price-adjusted values of Con_i. Price adjustment is done by dividing the nominal Con_i by the PPP conversion factor provided by the World Bank.

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.

We want to find a proxy for the number of patent-sensitive innovations or innovative capability that is comparable across countries. One option is to use some *input measures* — such as the number of scientists and engineers in the country. A larger number signifies more labor can engage in creative innovation. In fact, in Grossman 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 takes the constant elasticity of substitution form (CES). Since the CES function is quite general, the theory is compatible with the use of the number of scientist and engineers as a proxy for I. The data source for the number of scientists and engineers is the UNESCO Statistical Yearbook, supplemented by the World Competitiveness Yearbook of the IMD (Institute for Management Development).

However, as argued in Hall, Jaffe and Trajtenberg (2002), one shortcoming of this measure is that it assumes each scientist and engineer produces the same amount of knowledge output. In the real world, some of the research input may produce "dry holes" which may not contribute much to the "success" in innovative activity . For this reason, one might want to use the number of patents granted as an output measure of the knowledge gained by doing research. In view of this argument, another option to proxy for the number of patent-sensitive innovations or innovative capability is to use an *output measure* — the number of valuable patents (assigned patents) granted to domestic residents. Such measure is also used in McAleer et al. (2007) as a proxy of a country's innovative capability. However, using the simple patent count directly may give a misleading picture as some patents are never put to commercial use, and therefore do not have any significant commercial values. This issue is also discussed in Tsuji (2002), which argues that not all patents are commercially transformed into innovations. Some patents are merely used for their decoy and defence functions and thus should not be counted. For this reason, using the simple count of patents granted in various countries may not reflect the true innovative capabilities of the countries.

To measure the number of patents intended for commercial use, Marinova (1999, 2001) proposes using the number of assigned patents given to residents of the country as a proxy. It is argued that if the inventor of a registered patent has assigned (transferred) the right of use of the patent to someone else (an individual or a company), it is more likely that the legally protected prototype of the patent is intended for commercial exploitation soon. "Unassigned" patents are those for which the inventors have not yet granted the rights of the invention to a legal entity such as a corporation, university, or government agency, or to other individuals. Although an unassigned patent can still be exploited commercially, assigning a patent indicates an explicit intention to use it for commercial purpose. The assignment rate of country j is defined as the ratio of assigned patents with inventors from country j (AP_j) over the total patents granted to country j in the U.S. (TP_j). That is, the assignment rate is defined as $AssignmentRate_j = AP_j/TP_j$.[8] We make use of the assignment rates of various countries to approximate the share of patents granted in different countries that are commercialized. The number of assigned patents ($AssignedPatent_j$) used in the estimation is thus calculated as

$$AssignedPatent_j = Patent_j \times AssignmentRate_j$$

where $Patent_j$ refers to the number of patents granted to domestic residents in country j as re-

[8]For patents with multiple inventors, we classify the country origins of the patent inventors based on the residence of the first inventor. The data source is the US Patent and Trademark Office Database which is available at http://www.uspto.gov/patft/index.html.

ported by the World Intellectual Property Organization (WIPO). As a close inspection of the assigned patent count data indicates that there is much lumpiness in the data, smoothing is necessary to capture the cross-country variations in the longer term. We decide to use the sum of the values over the past three years to minimize the spurious time variations due to the arbitrariness in the granting time of the patents.

In the estimation, we tried both the simple patent count and the assigned patent count as the output measure of innovative capability. The results are largely the same. For this reason, we only present the results based on the assigned patent count. Summary statistics of the data are presented in Appendix C.

3.2 Structural Estimation and Goodness of Fit of the Model

In the non-cooperative game theoretical model of Grossman and Lai (2004), the welfare-maximizing government of each country chooses the level of patent protection. The government balances the benefits of patent protection resulting from the stimulation of innovation incentives with the efficiency costs caused by the curtailment of potential competition among firms.

Our empirical section is to explore to what extent the simultaneous solution of Grossman and Lai (2004)'s non-cooperative game theoretical model can explain the variation of patent protections across countries. According to the Grossman and Lai model discussed in section (2), the levels of patent protection $P_t \equiv [P_{1t}, P_{2t}, ..., P_{Nt}]'$ are determined by the model

$$\mathbf{P}_t = \mathbf{A}_t^{-1}\mathbf{b}_t$$

where P_t is a $N \times 1$ matrix; A_t is a $N \times N$ matrix with the diagonal elements equal to one; the $(i,k)^{th}$ off-diagonal element equals $\frac{[1-\mu_{it}\theta_2]}{(1-\mu_{it}\theta_2+\delta_t)}\left(\frac{m_{kt}}{m_{it}}\right)$; and $b_t = \left[\frac{\delta_t\theta_1}{1-\mu_{1t}\theta_2+\delta_t}, \cdots, \frac{\delta_t\theta_1}{1-\mu_{Nt}\theta_2+\delta_t}\right]'$ is a $N \times 1$ matrix. The variables μ_i and m_i refer to the data that measure the innovative capability of country i relative to the aggregate world innovative capability and the market size of country i relative to market size of the world, respectively. Given that the observed patent protection index (the Ginarte and Park index G_t) can be a monotonic transformation of the theoretical P_t, we assume

$$\mathbf{G}_t = h(\mathbf{P}_t) + \boldsymbol{\varepsilon}_t$$

where $h(\cdot)$ is a smooth strict monotonic function.

In the model, we calibrate the structural parameters θ_1 and θ_2 using the demand elasticities or values of mark-up in the patent-sensitive sectors of 14 countries as reported in the existing literature. We set the values of θ_1 and θ_2 to the average values of the 14 countries, which amount to 1.48 and 0.28 respectively. The remaining unknown structural coefficients δ and the nonparametric function h are jointly estimated using the sieve minimum distance estimator (SMD) of Ai and Chen (2003). The moment conditions used in the estimation are given by

$$E\{\mathbf{G}_t - h(\mathbf{A}_t^{-1}\mathbf{b}_t)|\mathbf{X}_t\} = 0 \tag{4}$$

where X_t refers to the set of instruments. The set of instruments include the four-year-lags of the market size and innovative capability measures.

We employ four sets of variables in the estimation to proxy for the theoretical variables M_i and I_i described in Grossman and Lai (2004)'s model. This is for examining the sensitivity of our results to alternative measures of the theoretical variables. The first two sets of variables are associated with the input measure of innovative capability and the other two sets are associated with the output measure of innovative capability. The input measure of innovative capability refers to the log of the average number of scientists and engineers in the preceding three years and the output measure of innovative capability refers to the log of the total number of assigned patents granted to domestic residents in the preceding three years. For the innovative capability measures, we always smooth the data as there is much lumpiness in the data without the smoothing, especially in the assigned patent variable.

The input and output measures of innovative capability are each matched with the smoothed and non-smoothed measures of market size to give us four different combinations. The smoothed market-size measure refers to the log of the average market size in the preceding three years, and the non-smoothed market-size measure refers to the log of the contemporaneous market size. In our view, taking the average can help to remove the transitory noises and allows us to focus on the medium to long-term pattern. However, the non-smoothed market size data are also used for the sake of comparison.

3.2.1 Estimation results with input measure of innovative capability

Panels (A) and (B) of Table 2 present the estimated coefficients and Pesaran and Smith (1994)'s generalized goodness-of-fit measure R^2 based on the number of scientists and engineers as the input measure of innovative capability and the non-smoothed as well as smoothed market size data respectively.[9] The estimated δ in the pre-TRIPS periods are positive as expected and are significant at the 5% significance level. The comparison of the generalized R^2 before and after the implementation of the TRIPS agreement allows us to compare the prediction performance of the structural model under two significantly different policy regimes, one in which the non-cooperative is believed to hold, and another that the model is less likely to hold[10]. The generalized R^2 shows that the goodness of fit of the non-cooperative game-theoretic model are quite similar for the period 1980 through 1990 before the TRIPS agreement was enforced but display a marked decline in 1995 and 2000. The generalized R^2 amounts to 0.3188, 0.3281 and 0.3274 respectively in the first three years but declines to 0.2112 and 0.1787 in the latter two years.

Figures 1(a) and 1(b) plot the actual patent protection level against the predicted patent protection level based on the input measure of innovative capability together with the non-smoothed and smoothed market size data respectively. The plots for 1995 and 2000 indicate that many of the actual values in the post-TRIPS periods cluster at a high level above the predicted values from the structural model. This provides an evidence that TRIPS agreement has really led to cooperative increase in patent protection beyond the incentive for countries to protect patents non-cooperatively. Table 3 lists the 27 countries (out of a total of 39) whose actual strengths of IPR protection were above the predicted values in 2000.

3.2.2 Estimation results with output measure of innovative capability

Panels (C) and (D) of Table 2 present the estimated δ and the generalized R^2 based on assigned patent count as the output measure of innovative capability. Similar to the case with the input measure of innovative capability, the estimated δ are significant at the 5% significance level and the generalized R^2 shows that the goodness of fit of the non-cooperative game model are about the same in 1980, 1985 and 1990 but drop markedly in 1995 and 2000. The generalized R^2 amounts to 0.2982, 0.2987 and 0.3082 respectively in the first three years but declines to 0.1465 and 0.1761 in the latter two years. Figures 1(c) and 1(d) plot the actual patent protection level against the

[9]It is shown in Pesaran and Smith (1994) that Theil's standard R^2 and $\overline{R}^2$ are inappropriate as a measure of fit and for model selection in the IV context. A generalized R^2 based on the prediction erros is is proposed by Pesaran and Smith for measuring the goodness of fit for IV regressions.

[10]See Keane and Wolpin (2005) for a similar analysis across different policy regimes.

predicted patent protection level based on the output measure of innovative capability together with the non-smoothed and smoothed market-size data, respectively.

[**To the referees: Since Figures 1 (a), (b), (c) and (d) present very similar results, only one of them need to be published.**]

4 Model Validation: Examining the In-Sample and Out-of-Sample Predictive Power of the Structural Model

4.1 Examining the in-sample predictive power of the structural model

We examine whether the structural model is a correct description of the data generating process before the implementation of TRIPS by testing whether the mean squared errors (MSE) is too high for this period. We perform the in-sample prediction of the patent protection in 1980, 1985 and 1990 and infer the statistical significance of their MSE. The null hypothesis is $H_0 : MSE = 0$ (the structural model is a correct description of the data generating process) and the alternative hypothesis is $H_1 : MSE \neq 0$ (the structural model is not a correct description of the data generating process). We calculate the p-values based on Rivers and Vuong (2002)'s method. A small p-value would imply that the MSE is "too high" and the structural model would fail to explain the variation in the patent protection data, while a large p-value would provide supporting evidence for the structural model. The p-values of the estimated MSE are reported in Table 4, and they are larger than 0.05 under all specifications, suggesting that the null hypothesis cannot be rejected.

4.2 Comparing the out-of-sample predictive power before and after the TRIPS agreement was enforced

So far we have checked the regression coefficients for their significance, the in-sample goodness of fit measured by the generalized R^2 and MSE. While the estimated coefficients and in-sample goodness of fit are of interest in their own right, the ultimate utility of the model depends not on the significance of the estimated coefficients but on the usefulness of the model in predicting patent protection across countries. In this section we examine whether the out-of-sample prediction of the structural model is consistent with the empirical observations. However, it should be emphasized that it is not our aim to show that the structural model can explain 100 percent of the variation of patent protections across countries. The objective is rather to offer a first analysis for whether there is a chance that the game-theoretic model is consistent with the empirical observations. To set the reader's expectations straight, we do not claim the game-theoretic model to be comprehensive and cover all factors that affect patent protection. In the structural model, we have abstracted from a number of complications involved in real-world determination of patent protection. In the model, we focus on two major factors, market size and innovative capability, that directly influence the marginal economic costs and benefits of patent protection. Our aim is simply to get an idea of the potential benefits of having a nonlinear structural model designed to explain the determination of patent protection based on solid theoretical grounds.

In this part of the model validation procedure, we use the rolling window approach to generate the one-step-ahead predictions of the structural models. We first estimate the structural model using the data of 1980 and perform the one-step-ahead out-of-sample prediction for 1985. After that, we estimate the structural model using the data of 1985 and perform out-of-sample prediction for 1990. We continue to roll forward to generate the one-step-ahead predictions for the other periods, assuming that the countries are still playing a non-cooperative game. The predicted values of intellectual property right protection are obtained based on the following equilibrium equations:

$$\begin{aligned} \widehat{\mathbf{G}}_t &= \widehat{h}(\widehat{\mathbf{P}}_t) \\ \text{where } \widehat{\mathbf{P}}_t &= \widehat{\mathbf{A}}_t^{-1}\widehat{\mathbf{b}}_t \end{aligned}$$

Since the TRIPS agreement may introduce a major change from non-cooperative to more cooperative behavior, we expect the goodness of fit of the out-of-sample prediction for 1995 using the estimates from 1990 to be significantly lower. Table 5 presents the sum of squared errors (SSE) of the one-step ahead out-of-sample predictions of the structural model before and after the implementation of the TRIPS agreement. The results indicate that the predictive power of the structural model declines from the pre-TRIPS period to the post-TRIPS period. For example, the SSE based on the input measure of innovative capability and the smoothed market-size data increases markedly by more than 28 percent from an average of 1.19 in the pre-TRIPS period to an average of 1.53 in the post-TRIPS period. The same pattern is present when the output measure of innovative capability or non-smoothed market-size data are used.

In addition, we expect the actual patent protection after the TRIPS agreement was implemented will be in general higher than the predicted patent protection of the non-cooperative game model. To study this, we calculate the percentage of countries which have higher patent protection than the level predicted by the non-cooperative game model ($G_i > \widehat{G}_i^{out-sample}$). The results in Table 6 show that the percentage of countries with $G_i > \widehat{G}_i^{out-sample}$ jumps from an average of 53 percent in the pre-TRIPS periods (1985 and 1990) to an average of 73.1 percent in the post-TRIPS periods (1995 and 2000) when the input measure of innovative capability and non-smoothed market size data is used. Similar jumps are observed when output measure of innovative capability or smoothed market size data is used. This supports the conjecture that TRIPS agreement has led to cooperative increase in patent protection in many countries beyond their incentives to protect patents based on their own self-interest.

4.3 Comparing the Out-of-Sample Predictive Abilities of the Structural Model and Other Alternative Models

The third part of the model validation procedure is to compare the out-of-sample goodness of fit of the non-cooperative game structural model against other competing models. In this section, we assess model validity by running a "horse race" between the structural model and a number of competing models. Existing empirical models in the literature are mostly linear models, eschewing functional nonlinearities. The structural model of Grossman and Lai is one of the first to provide a structural nonlinear functional form to model the global variations in patent protection. While it is understood that a certain amount of discrepancy inevitably exists between the observed data and the model, the best we can do is to identify the most adequate model relatively among a set of alternatives, both in terms of its consistency with the micro-founded optimization conditions and with the data pattern. The set of alternative models that we consider include:

1. Model (a) is a linearized version of our structural model, in which we regress the patent protection index on the innovative capability measure and the market size measure in a linear form. This linear form of the model can be taken as the first order approximation of any higher order structural models in which the optimal degree of patent protection is a function of one or both of these two variables. Examples of structural models in this family include Eicher and Garcia-Peñalosa (2008), in which the optimal degree of patent

enforcement is an increasing function of the innovative capability measured by the share of labor employment in the research sector, and Deng (2007), in which the value of patent rights is an increasing function of the economic size of the country. The linear model is thus very general in the sense that it is not tied to a specific structural model, but its major drawback is that it does not have any structural interpretation.

2. Model (b) is Ginarte and Park (1997)'s linear model, which is motivated by the idea that openness to trade, the degree of market freedom (how well governments abstain from market interference) and the ratio of research and development to GDP have significant bearing on the patent protection of a country. It is conjectured that economies that are more open are more susceptible to American pressure for reform.

3. Model (c) is Maskus (2000)'s linear model, which is motivated by the idea that GDP per capita, the ratio of scientists and engineers to labor share and the origins of the judicial systems (UK colony or French colony) have significant impacts on the patent protection of a country.

It is interesting to start the horse race by contrasting the structural model with the linearized version of it (alternative model (a)). As highlighted by Lucas's critique, structural models are much more appropriate than reduced-form linear models in studying out-of-sample prediction because the slope coefficients should not be some fixed constants but should be functions of the endogenous variables. Comparing the performance of the tightly specified structural model of Grossman and Lai with that of the linear model is like testing the functional form implied by Grossman and Lai's model against the linear approximation to the class of alternative models that are functions of the same factors.

In contrast, the variables included in the models of Ginarte and Park (1997) and Maskus (2000) are chosen largely on the basis of in-sample empirical goodness of fit. While it is expected intuitively that a model with more explanatory variables will provide better in-sample goodness of fit, it is not necessarily preferred if we examine the out-of-sample predictive power of the model. This consideration leads us to the model validation criterion — comparing the out-of-sample predictive performance of the structural model with those of the competing models.

Ginarte and Park (GP) regress the log of the GP patent protection index on the determinants of the patent rights, including the log of real GDP per capita, lagged R&D ratio to GDP

$\left(\frac{R\&D_i}{GDP_i}\right)$, secondary school enrollment ratio ($School_i$), the Barro-Lee (1994) index of political freedom ($PoliticalFree_i$), the Sachs-Warner (1995) dummy variable for an economy's openness to trade ($Openness_i$) and the Johnson-Sheely (1995) index of market freedom ($MarketFree_i$).

$$\begin{aligned} GP_i &= \beta_0 + \beta_1 GDPCap_i + \beta_2 \frac{R\&D_i}{GDP_i} + \beta_3 School_i + \beta_4 PoliticalFree_i \\ &\quad + \beta_5 Openness_i + \beta_6 MarketFree_i + e_i \end{aligned}$$

Ginarte and Park find that there are strong positive impacts from the lagged R&D ratio, trade openness and the index of market freedom on patent rights. Human capital, as proxied by the lagged secondary school enrollment ratio, is positive but is only marginally significant. The political freedom index also turns out to be insignificant. After adding other control variables, the GDP per capita becomes insignificant. This finding leads to the conjecture that using the linear functional form may be subject to high multicollinearity problem. However, once we open the door to the nonlinear world, the number of possibilities is so large that we need to be guided by a structural model which can give a clear guidance on the specific functional form. As Grossman and Lai's model is a structural model that gives us clear direction on the specific nonlinear functional form that should be used, we examine the validity of Grossman and Lai's nonlinear structural model relative to it.

Maskus adds further explanatory variables that account for other influences on patent rights. First, he hypothesizes an inverted-U shape relationship between real GDP per capita and patent protection based on graphical inspection of the data. Second, he adds the size of the economy into the equation as measured by real GDP. It is hypothesized that countries do not begin to provide strong IPRs until they reach a certain size, because there are fixed costs in organizing and administering a patent system. In addition, he adds the proportion of scientists and engineers in the labor force as a proxy of the innovative capability $\left(\frac{S\&E_i}{LabForce_i}\right)$, and two dummies that represent former colonies of UK and France respectively ($UKCol_i$ and $FCol_i$).

$$\begin{aligned} GP_i &= \alpha_0 + \alpha_1 GDPCap_i + \alpha_2 (GDPCap_i)^2 + \alpha_3 School_i \\ &\quad + \alpha_4 Openness_i + \alpha_5 GDP_i + \alpha_6 \frac{S\&E_i}{LabForce_i} \\ &\quad + \alpha_7 UKCol_i + \alpha_8 FCol_i + e_i \end{aligned}$$

Their estimation results indicate that the scientists and engineers ratio and the two dummies for UK and France have significantly positive impact on the strength of patent protection. However, market size has no detectable impact on patent rights, which suggests GDP itself is not a determinant of IPRs reform. The school enrollment ratio and trade openness also turn out to be insignificant.

Table 5 presents the SSE of the one-step ahead out-of-sample predictions of the structural model and the alternative models for the periods 1985-2000. In each of the year 1980, 1985, 1990 and 1995, we estimate the structural model and employ the estimated coefficients to perform one-step ahead predictions of patent protection in 1985, 1990, 1995 and 2000 respectively. When the input measure of innovative capability is used, the structural model consistently outperforms the linear model. In our judgement, this finding provides us with some confidence in the nonlinear structural model as an explanation of patent protection choices of different countries. However, the result is mixed when the output measure of innovative capability is used.

When the nonlinear structural model is compared with Maskus's model as well as Ginarte and Park's model, Ginarte and Park's model performs uniformly better in all years. Sign test results in Table 7 confirm that the MSE of the Ginarte and Park model is significantly lower than that of the structural model in 1990, while Wilcoxin's signed rank test results indicate that the MSE of Maskus and Ginarte and Park's models are significantly lower than that of the nonlinear structural model in both 1985 and 1990 at the 5% significance level.[11] While these findings by and large provides some supporting evidence to Maskus and Ginarte and Park's variable choice, it does not necessarily indicate that the game-theoretic model has to be abandoned. They just call for further theoretical research in this area to incorporate the factors in Ginarte and Park's model to the structural model. The success of the Ginarte and Park model leads us to believe that factors like trade openness and institutional setting are relevant in the global determination of IPR protection.

5 Conclusions

This paper examines the structural model of Grossman and Lai (2004) by analyzing its consistency with the cross sectional and time series pattern of the data. As the structural model is a non-cooperative game model, it is expected to be more applicable in a policy regime before any

[11]Details of the three statistical tests (the Rivers-Vuong test, sign test and Wilcoxin's signed rank test) are provided in Appendix D.

cooperative agreement is implemented. A comparison of the in-sample generalized R^2 before and after the implementation of the TRIPS agreement in 1994 indicates the structural model fits the data significantly better prior to 1994. The generalized R^2 ranges from 0.2796 to 0.3430 in the period 1980 through 1990, but declines substantially in 1995 and 2000. The same pattern is robust to the use of the input and output measures of innovative capability.

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 substantially from around 57 percent in the pre-TRIPS periods to over 75 percent in the post-TRIPS periods. This confirms our expectation that countries were more cooperative in their post-TRIPS patent policies.

In this paper, we also assess the validity of the model by running a "horse race" between Grossman and Lai's structural model and a number of competing models, which include (i) a linearized version of the structural model, (ii) Maskus's model and (iii) Ginarte and Park's model. The comparison of the MSE across the models indicate that the nonlinear structural model improves over the linearized version when the input measure of innovative capability is used. However, the nonlinear structural model does not outperform Maskus's model or Ginarte and Park's model. Overall, Ginarte and Park's model performs uniformly better in all years. The success of the Ginarte and Park model calls for the extension of the Grossman and Lai model to incorporate additional factors like trade openness and institutional setting. This is an important extension that is beyond the scope of this paper but it merits further investigation.

Appendix A: Demand elasticities and mark-up in the patent sensitive sector

The parameters θ_1^i and θ_2^i are related to the markups (mp_i), or equivalently, price elasticities (ϵ_i) across countries as $\theta_1^i = 1/\left[1 - \left(\frac{\epsilon_i}{\epsilon_i - 1}\right)^{1-\epsilon_i}\right]$ and $\theta_2^i = \left(\frac{\epsilon_i}{\epsilon_i - 1}\right)^{-\epsilon_i} / \left[1 - \left(\frac{\epsilon_i}{\epsilon_i - 1}\right)^{1-\epsilon_i}\right]$, with $\epsilon_i = \frac{mp_i}{mp_i - 1}$. Based on the values of ϵ_i or mp_i of 14 countries reported by various studies in the literature, the implied θ_1^i and θ_2^i are computed and reported below. The results indicate that the values of θ_1^i and θ_2^i across countries are quite similar, which provides supporting evidence for the assumption that θ_1 and θ_2 are constant across countries.

countries	studies	ϵ_i in the patent sensitive sector	θ_1^i	θ_2^i
Australia	Joaquim et al. (1996)	-5.53	1.51	0.26
Belgium	Joaquim et al. (1996)	-7.48	1.52	0.26
Canada	Morrison (1994)	-5.46	1.51	0.26
	Beccarello (1996)	-10.09	1.55	0.24
	Joaquim et al. (1996)	-6.67	1.52	0.26
Denmark	Joaquim et al. (1996)	-6.38	1.52	0.26
Finland	Joaquim et al. (1996)	-4.48	1.49	0.27
France	Beccarello (1996)	-4.49	1.49	0.27
	Joaquim et al. (1996)	-6.73	1.52	0.26
Germany	Beccarello (1996)	-2.51	1.44	0.30
	Joaquim et al. (1996)	-4.85	1.50	0.27
	Möller (2001)	-1.34	1.37	0.34
Italy	Beccarello (1996)	-5.23	1.50	0.27
	Joaquim et al. (1996)	-14.72	1.55	0.24
Japan	Beccarello (1996)	-9.20	1.53	0.25
	Joaquim et al. (1996)	-5.21	1.51	0.27
	Nishimura (1999)	-10.26	1.54	0.25
Netherland	Joaquim et al. (1996)	-6.06	1.51	0.26
Norway	Joaquim et al. (1996)	-6.99	1.52	0.26
Sweden	Joaquim et al. (1996)	-6.56	1.52	0.26
UK	Beccarello (1996)	-3.62	1.47	0.28
	Joaquim et al. (1996)	-12.54	1.54	0.25
	Möller (2001)	-0.6	1.26	0.43
US	Hall (1988)	-1.57	1.39	0.33
	Norrbin (1993)	-1.85	1.41	0.32
	Haskel et al. (1995)	-2.59	1.44	0.30
	Roeger (1995)	-3.32	1.47	0.28
	Beccarello (1996)	-2.97	1.46	0.29
	Joaquim et al. (1996)	-8.49	1.53	0.25
	Möller (2001)	-1.18	1.36	0.35
Average			1.48	0.28

Appendix B: Estimating the Smooth Monotone Functions

In this section, we describe how to estimate the smooth monotone function $h(\cdot)$. By using the result of Ramsay (1998), we can set the smooth monotone function as

$$h(p) = \beta_0 + \beta_1 \int_0^p \exp\{\sum_{j=1}^{J_N} c_j Q_j(s)\} ds$$

where β_0, β_1, c_1, ... ,c_k are the real parameters, $p \in [0,1]$ and the linear combinations of $\{q_j(p) = d(Q_j(p))/dp\}_{j=1}^{\infty}$ that span the space of square Lebesgue integrable functions. In Ai and Chen's (2003) terminology, such set of basis functions form a linear sieve for the function $w(p)$ in Ramsay (1998). $J_N \to \infty$ slowly as $N \to \infty$. The sieve minimum distance (SMD) estimator of Ai and Chen (2003) is used in this paper to jointly estimate the coefficients of the structural model and the unknown monotonic function $h(\cdot)$. The SMD estimator can be interpreted as the nonparametric version of GMM-IV. While the SMD estimator is based on the nonparametric conditional moments restrictions, the GMM-IV is based on the parametric counterpart.

For illustration, Figures 2 (a)-(d) show the shapes of the estimated smooth monotone functions under different specifications of the data. One phenomenon that is worth noticing is that the monotone functions significantly flattened out in the middle range (when $\widehat{\mathbf{P}}_t$ is between around 0.4 and 0.8) during the post-TRIPS periods (1995 and 2000). This indicates that the structural model's prediction ($\widehat{\mathbf{P}}_t$) becomes less effective in the prediction of the Ginarte-Park patent protection index G_t in the post-TRIPS periods.

Implementation of the SMD estimation

Let

$$m(\Theta) = E[\mathbf{G}_t - h_\theta(\mathbf{A}_t \mathbf{b}_t)|X_t]$$

where $\Theta = (h, \theta)$ and X_t are the instrumental variables. The SMD estimator is obtained by solving

$$\inf_{(\theta,h)} E[m(\Theta)^T \Sigma^{-1} m(\Theta)]$$

where Σ is a positive definite weighting matrix.

- Step 1: Take the weighting matrix Σ to be an identity matrix and get an initial consistent

SMD estimator $\hat{\Theta} = \{\hat{\beta}_0, \hat{\beta}_1, \{\hat{c}_j\}, \hat{\delta}\}$ by solving

$$\min_{\beta_0,\beta_1,\{c_j\},\delta} m_t^T m_t$$

- Step 2: Compute the optimal weighting matrix Σ using:

$$E_t = \mathbf{G}_t - \hat{h}(\hat{\mathbf{A}}_t^{-1}\hat{\mathbf{b}}_t)$$

and then let W_t be the upper diagonal elements (including the diagonal) of $\left(E_t E_t^T\right)_{N\times N}$.

If we run the regression

$$W_t = P(X_t)\Psi + \eta'$$

then the upper diagonal elements (including the diagonal) of the optimal weighting matrix $\hat{\Sigma}(\hat{\Theta})$ is given by $\hat{W}_t$. Here the columns of $P(X_t)$ are the corresponding tensor product spline basis of X_t mentioned in Ai and Chen (2003, p.1799).

- Step 3: Get $\tilde{\Theta}$ by solving

$$\min_{\beta_0,\beta_1,\{c_j\},\delta} \hat{m}^T \hat{\Sigma}(\hat{\Theta})^{-1}\hat{m}$$

where $\tilde{\Theta}$ is the corresponding SMD estimator.

Properties of the SMD estimators

Consider another set of basis functions (tensor product splines in our case) $\{\psi_k(x_1, x_2)\}_{k=1}^{\infty}$ which again spans the space of square Lebesgue integrable functions. The following approximation is applied:

$$\mathbf{G}_t - h(\mathbf{A}_t^{-1}\mathbf{b}_t) = \Psi_{N\times K_n}\Gamma + \eta$$

where η is of mean 0 and $K_n \to \infty$ slowly. Then,

$$\hat{m} = \Psi(\Psi^T\Psi)^{-1}\Psi^T[\mathbf{G}_t - h(\mathbf{A}_t^{-1}\mathbf{b}_t)]$$

is a good estimate of $E\{G_t - h(A_t^{-1}b_t)|X_t\}$. The minimum distance estimator is simply

$$\min_{\beta_0,\beta_1,\{c_j\},\delta} \hat{m}^T\hat{\Sigma}^{-1}\hat{m}$$

where $\hat{\Sigma}$ is a consistent estimator of the variance of η. As established in Ai and Chen (2003), the estimators are consistent, efficient and asymptotically normally distributed.

Basis functions

According to Ramsay (1998), we could simply choose

$$q_j(p) = I\{p \in (\xi_{j-1}, \xi_j]\}$$

where $\xi_0 = 0 < \xi_1 < \cdots < \xi_{J_n} < \xi_{J_n+1} = 1$ and $I\{\cdot\}$ is the indicator function. $\{\xi_j\}_{j=1}^{J_N}$ are the knots (they refer to the quantiles of $\hat{P}_t$). We choose $J_n = 3$. In that case,

$$\begin{aligned} Q_j(p) = \int_0^p I\{s \in (\xi_{j-1}, \xi_j]\}ds &= \begin{cases} \xi_j - \xi_{j-1} & \text{if } p > \xi_j \\ p - \xi_{j-1} & \text{if } p \in (\xi_{j-1}, \xi_j] \\ 0 & \text{if } p \leq \xi_{j-1} \end{cases} \\ = \min((p - \xi_{j-1})_+, \xi_j - \xi_{j-1}) \end{aligned}$$

Therefore,

$$h(p) = \beta_0 + \beta_1 \int_0^p \exp\{\sum_{j=1}^{J_n} c_j \min((s - \xi_{j-1})_+, \xi_j - \xi_{j-1})\}ds.$$

For $p \leq \xi_1$,

$$h(p) = \beta_0 + \beta_1 \int_0^p \exp\{c_1 s\}ds = \beta_0 + \beta_1 \frac{1}{c_1}[\exp\{c_1 p\} - 1]$$

For $p \in (\xi_1, \xi_2]$,

$$\begin{aligned} h(p) &= \beta_0 + \beta_1 [\int_0^{\xi_1} \exp\{c_1 s\}ds + \int_{\xi_1}^p \exp\{c_1\xi_1 + c_2(s - \xi_1)\}ds \\ &= \beta_0 + \beta_1 \frac{1}{c_1}[\exp\{c_1\xi_1\} - 1] + \beta_1(\exp\{c_1\xi_1\})\frac{1}{c_2}\exp\{c_2(p - \xi_1)\} \\ &= \beta_0 + \beta_1 \frac{1}{c_1}[\exp\{c_1\xi_1\} - 1] + \beta_1 \frac{1}{c_2}\exp\{c_1\xi_1 + c_2(p - \xi_1)\} \end{aligned}$$

In general, $p \in (\xi_{j-1}, \xi_j]$

$$\begin{aligned} &h(p) \\ &= \beta_0 + \beta_1 \frac{1}{c_1}[\exp\{c_1\xi_1\} - 1] + \beta_1 \frac{1}{c_2}\exp\{c_1\xi_1 + c_2(\xi_2 - \xi_1)\} \\ &+ \beta_1 \frac{1}{c_3}\exp\{c_1\xi_1 + c_2(\xi_2 - \xi_1) + c_3(\xi_3 - \xi_1)\} \\ &\dots \\ &+ \beta_1 \frac{1}{c_j}\exp\{c_1\xi_1 + c_2(\xi_2 - \xi_1) + \cdots + c_{j-1}(\xi_{j-1} - \xi_{j-2}) + c_j(p - \xi_{j-1})\} \end{aligned}$$

Also, we take additive quadratic splines to approximate $E[\mathbf{G}_t - h(\mathbf{A}_t^{-1}\mathbf{b}_t)]$. That is, the Ψ basis functions are

$$x_{1t}, x_{1t}^2, \{(x_{1t} - \zeta_{1,j})_+^2\}_{j=1}^{K_N}, x_{2t}, x_{2t}^2, \{(x_{2t} - \zeta_{2,j})_+^2\}_{j=1}^{K_n}$$

where $\zeta_{1,j}$ and $\zeta_{2,j}$ are the corresponding quantiles of x_{1t} and x_{2t}. K_n is chosen to be 3.

Appendix C: Summary Statistics of the Data

		1980	1985	1990	1995	2000
GP index[12]	mean	0.425	0.443	0.477	0.633	0.741
(0: lowest,	std. err.	0.204	0.216	0.242	0.228	0.177
1: highest)	max.	0.870	0.935	0.935	0.975	0.975
	min.	0.117	0.117	0.117	0.215	0.255
Market size	mean	56.76	66.76	90.13	116.82	158.62
(million USD)	std. err.	148.29	164.47	216.65	265.88	373.51
	max.	833.19	904.29	1238.2	1500.9	2092.4
	min.	0.356	0.391	0.477	0.500	0.490
Scientists and	mean	5.468	5.97	6.74	7.319	8.051
engineers	std. err.	13.832	15.385	18.169	18.965	20.592
(in 10,000)	max.	62.971	74.103	94.182	96.65	113.12
	min.	0.0033	0.0034	0.0034	0.0036	0.0036
Assigned Patents	mean	0.4648	0.5634	0.7835	0.9246	1.4228
(in 10,000)	std. err.	2.0244	2.2834	3.0855	3.8051	5.6448
	max.	12.487	13.788	18.328	22.932	34.155
	min.	0.0001	0.0001	0.0001	0.0001	0.0001
observations		39	39	39	39	39

[12]The GP index is taken to be the Ginarte and Park index (which originally ranges between 0 and 5) divided by 5. Hence $GP_{it} \in [0, 1]$.

Appendix D: Hypothesis tests on the difference of MSE

In Table 7, we perform three tests to evaluate the statistical significance of the difference in the MSE across models. The three tests include the Rivers and Vuong (2002) test, the sign test and the Wilcoxin's signed-rank test. In the three tests, the null hypothesis is that the MSE of the structural model and the alternative model are equal. The alternative hypothesis is that the MSE are unequal.

$$H_0 \; : \; MSE_{structural} = MSE_{alternative}$$
$$H_1 \; : \; MSE_{structural} \neq MSE_{alternative}$$

Rejecting the null hypothesis indicates that the MSE of the two models are significantly different. The three tests can also be applied to do one-sided hypothesis testing, with the p-value being half of the p-value of a two-sided test.

The Rivers-Vuong Test

Rivers and Vuong (2002) generalizes Vuong (1989)'s asymptotically normal tests for model selection by allowing for model selection criteria other than the models' likelihoods such as the mean squared errors of prediction. This test is used to compare competing models based on the MSE of their out-of-sample prediction. The test statistic is given as

$$T_n = \frac{\sqrt{n}}{\widehat{\sigma}_n} \left\{ Q_n^1 \left(\widehat{\gamma}_n^1 \right) - Q_n^2 \left(\widehat{\gamma}_n^2 \right) \right\} \sim N(0,1)$$

where $Q_n^1 \left(\widehat{\gamma}_n^1 \right)$ and $Q_n^2 \left(\widehat{\gamma}_n^2 \right)$ are the lack-of-fit criteria for model 1 and 2 respectively, and $\widehat{\sigma}_n$ is the asymptotic standard error of the difference in lack-of-fit.

Using the empirical mean square error (MSE) as the lack-of-fit criteria, we have

$$T_n = \frac{\sqrt{n}}{\widehat{\sigma}_n} \left\{ MSE_{2n} - MSE_{1n} \right\} \sim N(0,1)$$

Following McCracken and Sapp (2005), we can approximate $\widehat{\sigma}_n$ by $\sqrt{\sum_{i=1}^{N} \widehat{d}_{i,out}^2}$ where $\widehat{d}_{i,out} = \widehat{e}_{1,out}^2 - \widehat{e}_{2,out}^2$.

The Sign Test

Consider two forecasts, $\left\{\widehat{y}_i^A\right\}_{i=1}^N$ and $\left\{\widehat{y}_i^B\right\}_{i=1}^N$ of the data series $\{y_i\}_{i=1}^N$. Let the associated forecast errors be $\left\{e_i^A\right\}_{i=1}^N$ and $\left\{e_i^B\right\}_{i=1}^N$. The loss function $g\left(y_t, \widehat{y}_i^j\right)$ is a function of the forecast errors. That is, $g\left(y_i, \widehat{y}_i^j\right) = g\left(e_i^j\right)$ for $j = A, B$. The null hypothesis is a zero median loss differential: $med(g(e_i^A) - g(e_i^B)) = 0$.

The test statistic is constructed as follows: assuming that the loss-differential series is i.i.d., the number of positive loss-differential observations in a sample of size N has the binomial distribution with parameters N and $\frac{1}{2}$ under the null hypothesis. The test statistic is therefore simply

$$S_1 = \sum_{i=1}^{N} I_+\left(d_i\right)$$

where

$$\begin{aligned} I_+\left(d_i\right) &= 1 \qquad \text{if } d_i > 0 \\ &= 0 \qquad \text{otherwise} \end{aligned}$$

Significance may be assessed using a table of cumulative binomial distribution. In large samples, the studentized version of the sign-test statistic is standard normal:

$$S_{1a} = \frac{S_1 - 0.5N}{\sqrt{0.25N}} \overset{a}{\sim} N\left(0,1\right)$$

Wilcoxin's Signed-Rank Test

A related distribution-free procedure that requires symmetry of the loss differential (but can be more powerful than the sign test in that case) is Wilcoxin's signed-rank test. Assume that the loss-differential series is i.i.d., the test statistic is

$$S_2 = \sum_{i=1}^{N} I_+\left(d_i\right) rank\left(|d_i|\right)$$

which is the sum of the ranks of the absolute values of the positive observations. The exact finite-sample critical values of the test statistic are invariant to the distribution of the loss differential —

it needs only be zero-mean and asymmetric — and have been tabulated. Moreover, its studentized version is asymptotically standard normal,

$$S_{2a} = \frac{S_2 - \frac{N(N+1)}{4}}{\sqrt{\frac{N(N+1)(2N+1)}{24}}} \overset{a}{\sim} N\left(0,1\right)$$

References

[1] Acemoglu, Daron and Joshua Linn (2004), "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry", the Quarterly Journal of Economics, MIT Press 119(3), 1049-1090.

[2] Ai, Chunrong and Xiaohong Chen (2003), "Efficient Estimation of Models with Conditional Moment Restrictions Containing Unknown Functions," Econometrica 71(6), 1795-1843.

[3] Barro, Robert J. and J. H. Lee (1994), "Sources of Economic Growth", Carnegie-Rochester Conference Series on Public Policy, 40, 1-46.

[4] Beccarello, Massimo (1996), "Time Series Analysis of Market Power: Evidence from G-7 Manufacturing", International Journal of Industrial Organization, 15, 123-136.

[5] Branstetter, Lee, Raymond Fisman and C. Fritz Foley (2006), "Do Stronger Intellectual Property Rights Increase International Technology Transfer? Empirical Evidence from U.S. Firm-Level Panel Data", Quarterly Journal of Economics, 121(1), 321-349.

[6] Deng, Y. (2007), "Private Value of European Patents," European Economic Review 51, 1785-1812.

[7] Eicher, Theo and Cecilia Garcia-Peñalosa (2008), "Endogenous Strength of Intellectual Property Rights: Implications for Economic Development and Growth," European Economic Review 52, 237-258.

[8] Ginarte, Juan C. and Walter G. Park (1997), "Determinants of Patent Rights: A Cross-National Study," Research Policy 26, 283-301.

[9] Grossman, Gene M. and Edwin L.-C. Lai (2004), "International Protection of Intellectual Property," American Economic Review 94 (5), 1635-1653.

[10] Hall, Robert E. (1988), "The Relation between Price and Marginal Cost in U.S. Industry", the Journal of Political Economy 96(5), 921-947.

[11] Hall, Robert E., A. Jaffe and M. Trajtenberg (2002), "Market Value and Patent Citations: A First Look," Development and Comp Systems 0201001, EconWPA.

[12] Haskel, Jonathan, Christopher Martin and Ian Small (1995), "Price, Marginal Cost and the Business Cycle", Oxford Bulletin of Economics and Statistics 57(1), 25-41.

[13] Joaquim, Oliveira Martins, Stefano Scarpetta and Dirk Pilat (1996), "Mark-up Pricing, Market Structure and the Business Cycle", OECD Economic Studies No.27.

[14] Johnson, B. T. and T. P. Sheehy (1995), "The Index of Economic Freedom", Washington: Heritage Foundation.

[15] Keane, Michael P. and Kenneth I. Wolpin (2005), "Exploring the usefulness of a non-random holdout sample for model validation: Welfare Effects on Female Behavior", mimeo, Yale University.

[16] Lee, Jeong-Yeon and Edwin Mansfield (1996), "Intellectual Property Protection and U.S. Foreign Direct Investment," The Review of Economics and Statistics 78(2), 181-186.

[17] Lerner, Josh (2002), "150 Years of Patent Protection," American Economic Review Papers and Proceedings, 92: 221-225.

[18] Marinova, D. (1999), "Patent Data Models: Study of Technological Strengths of Western Australia," In: Proceedings of the IASTED International Conference on Applied Modelling and Simulation. Cairns, Australia, 118-123.

[19] — (2001), "Eastern European Patenting Activities in the USA," Technovation 21, 571-584.

[20] Maskus, Keith E. (2000), Intellectual Property Rights in the Global Economy, Washington D.C.: The Institute for International Economics.

[21] Maskus, Keith E. and Denise Eby-Konan (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.

[22] Maskus, Keith E. and Mohan Penubarti (1995), "How Trade-related are Intellectual Property Rights?" Journal of International Economics 39, 227-248.

[23] McAleer, Michael, Felix Chan and Dora Marinova (2007), "An Econometric Analysis of Asymmetric Volatility: Theory and Application to Patents," Journal of Econometrics 139, 259-284.

[24] McCalman, Phillip (2001), "Reaping What you Sow: An Empirical Analysis of International Patent Harmonization," Journal of International Economics 55, 161-185.

[25] McCracken, Michael W. and Stephen G. Sapp (2005), "Evaluating the Predictability of Exchange Rates Using Long-Horizon Regressions: Mind your p's and q's," Journal of Money, Credit and Banking, 37(3), 473-494.

[26] Möller, Joachim (2001), "Income and Price Elasticities in Different Sectors of the Economy: an Analysis of Structural Change for Germany, the UK and the USA", ed. by Thijs ten Raa and Ronald Schettkat, The growth of service industries: The Paradox of Exploding Costs and Persistent Demand, Edward Elgar: Cheltenham, UK and Northampton, MA, USA.

[27] Morrison C.J. (1994), "The Cyclical Nature of Markups in Canadian Manufacturing: A Production Theory Approach", Journal of Applied Econometrics 9(3), 269-282.

[28] Nishimura, Kiyohiko, Yasushi Ohkusa and Kenn Ariga (1999), "Estimating the mark-up over marginal cost: a panel analysis of Japanese firms 1971-1994", International Journal of Industrial Organization 17, 1077-1111.

[29] Norrbin, Stefan C (1993), "The Relation Between Price and Marginal Cost in U.S. Industry: A contradiaction", Journal of Political Economy 101(6), 1149-1164.

[30] Park, Walter G. and Smita Wagh (2002), "Index of Patent Rights," in Economic Freedom of the World: 2002 Annual Report, Chapter 2.

[31] Pesaran, M. Hashem and Richard J. Smith (1994), "A Generalized R^2 Criterion for Regression Models Estimated by the Instrumental Variables Method", Econometrica 62(3), 705-710.

[32] Ramsay, J.O. (1998), "Estimating Smooth Monotone Functions," Journal of the Royal Statistical Society, Series B (Statistical Methodology) 60(2), 365-375.

[33] Rivers, D. and Q. Vuong (2002), "Model Selection Tests for Nonlinear Dynamic Models", Econometrics Journal 5, 1-39.

[34] Roeger, Werner (1995), "Can Imperfect Competition Explain the Difference between Primal and Dual Productivity Measures? Estimates for U.S. Manufacturing", Journal of Political Economy 103(2), 316-330.

[35] Sachs, Jeffrey and Andrew Warner (1995), "Economic Reform and the Process of the Global Integration", Brookings Papers on Economic Activity 1, 1-95.

[36] Smith, Pamela J. (1999), "Are Weak Patent Rights a Barrier to U.S. Exports?" Journal of International Economics 48, 151-177.

[37] Tsuji, Y.S. (2002), "Organizational Behavior in the R&D process based on patent analysis: strategic R&D management in a Japanese Electronics firm," Technovation 22(7), 417-425.

[38] Vuong, Q. (1989), "Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses," Econometrica 57, 307-333.

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

Structural Parameter Values and Goodness of Fit					
θ_1	1.48				
θ_2	0.28				
	prior to TRIPS			after TRIPS	
coefficients	1980	1985	1990	1995	2000
(A) with input measure of innovative capability and non-smoothed market size data					
estimated δ	0.7695	0.1991	0.0572	0.1093	0.3825
p-value	(0.0000)*	(0.0000)*	(0.0000)*	(0.0000)*	(0.0000)*
generalized R^2	0.3188	0.3281	0.3274	0.2112	0.1787
(B) with input measure of innovative capability and smoothed market size data					
estimated δ	0.1074	0.4738	0.1051	0.4224	0.1709
p-value	(0.0000)*	(0.0000)*	(0.0000)*	(0.0000)*	(0.0000)*
generalized R^2	0.3187	0.3430	0.3019	0.2128	0.1248
(C) with output measure of innovative capability and non-smoothed market size data					
estimated δ	0.9626	0.7965	0.6598	0.0194	0.1542
p-value	(0.0000)*	(0.0000)*	(0.0000)*	(0.0000)*	(0.0000)*
generalized R^2	0.2982	0.2987	0.3082	0.1465	0.1761
(D) with output measure of innovative capability and smoothed market size data					
estimated δ	0.4876	0.9191	0.0703	0.4891	0.4444
p-value	(0.0000)*	(0.0000)*	(0.0000)*	(0.0000)*	(0.0000)*
generalized R^2	0.2982	0.3100	0.2796	0.1498	0.1081

Note: θ_1 and θ_2 are calibrated based on the demand elasticities of the patent-sensitive sectors of 14 countries as reported in Appendix A. The calibrated θ_1 and θ_2 are taken as the average values over the 14 countries .
The numbers in parentheses are the p-values. "*" means the estimated coefficient is significant at the 5% significance level.
Input measure of innovative capability refers to the number of scientists and engineers.
Output measure of innovative capability refers to the number of assigned patents.
Smoothed data of M_i refers to the average market size values in the three preceding years.

Table 2: Parameter estimates and goodness of fit of the structural model

Countries with actual IPR protection above and below the in-sample predicted values from the structural model during 2000	
$G_{i,00} \geq \widehat{G}_{i,00}^{in-sample}$	$G_{i,00} < \widehat{G}_{i,00}^{in-sample}$
Australia	Colombia
Austria	Costa Rica
Canada	Egypt
Chile	Guatemala
Cyprus	India
Ecuador	Jordan
Fiji	Malaysia
Finland	Mexico
France	Pakistan
Germany	Peru
Greece	Thailand
Hong Kong	Venezuela
Iceland	
Ireland	
Japan	
Korea, South	
Malta	
Netherlands	
Norway	
Philippines	
Portugal	
Singapore	
Spain	
Sweden	
Turkey	
United Kingdom	
United States	
Note: The results are based on the input measure of innovative capability and the non-smoothed market size data. The results based on the output measure of innovative capability and the smoothed market size data are similar.	

Table 3: Countries with actual IPR protection above the in-sample predicted values in 2000

Hypothesis tests on the MSE of the in-sample predictions of the structural model in the pre-TRIPS periods $H_0 : MSE = 0$ (the structural model is a correct description of the data generating process) $H_1 : MSE \neq 0$ (the structural model is not a correct description of the data generating process)			
(A) with input measure of innovative capability and non-smoothed data			
	$\widehat{MSE}_{80}$	$\widehat{MSE}_{85}$	$\widehat{MSE}_{90}$
value p-value	0.0286 (0.5164)	0.03117 (0.5124)	0.0388 (0.5016)
(B) with input measure of innovative capability and smoothed data			
	$\widehat{MSE}_{80}$	$\widehat{MSE}_{85}$	$\widehat{MSE}_{90}$
value p-value	0.0287 (0.5164)	0.0305 (0.5079)	0.0399 (0.4966)
(C) with output measure of innovative capability and non-smoothed data			
	$\widehat{MSE}_{80}$	$\widehat{MSE}_{85}$	$\widehat{MSE}_{90}$
value p-value	0.0290 (0.5155)	0.0329 (0.5044)	0.0389 (0.4918)
(D) with output measure of innovative capability and smoothed data			
	$\widehat{MSE}_{80}$	$\widehat{MSE}_{85}$	$\widehat{MSE}_{90}$
value p-value	0.0290 (0.5156)	0.0323 (0.4966)	0.0402 (0.4872)
Note: The p-values are calculated based on the Rivers-Vuong test statistic. "*" indicates that the null hypothesis is rejected at the 5 percent significance level.			

Table 4: In-sample model validation: hypothesis tests on the significance of the in-sample MSE in the pre-TRIPS periods

SSE of One-Step-Ahead Out-of-Sample Predictions (with input measure of innovative capability)				
	prior to TRIPS		after TRIPS	
Model	1985	1990	1995	2000
structural (non-smoothed market size data)	1.0307	1.3214	1.5548	1.1306
structural (smoothed market size data)	0.9839	1.3969	1.6236	1.4225
linear (non-smoothed market size data)	1.2374	1.6261	4.2342	1.6819
linear (smoothed market size data)	1.2325	1.6462	4.2005	3.0507
Maskus	0.9033	0.8643	–	–
Ginarte and Park	0.5112*	0.4785*	–	–

SSE of One-Step-Ahead Out-of-Sample Predictions (with output measure of innovative capability)				
	prior to TRIPS		after TRIPS	
Model	1985	1990	1995	2000
structural (non-smoothed market size data)	0.8912	1.3518	1.3225	1.1450
structural (smoothed market size data)	0.9574	1.2751	1.3146	1.0019
linear (non-smoothed market size data)	0.9671	0.8159	1.2119	0.7398
linear (smoothed market size data)	1.0171	0.7333	1.2679	0.6922
Maskus	0.9033	0.8643	–	–
Ginarte and Park	0.5112*	0.4785*	–	–

Note: SSE stands for the sum of squared errors.
"*" indicates the model with the lowest SSE among all models in the same period.
Since the data for Maskus's and Ginarte and Park's models are not all available in 1995 and 2000, the estimation for these two years are skipped.

Table 5: Out-of-sample model validation: one-step-ahead predictions of various models

(A) using input measure of innovative capability and non-smoothed market size data			
$G_{i,85} > \widehat{G}_{i,85}^{out-sample}$	$G_{i,90} > \widehat{G}_{i,90}^{out-sample}$	$G_{i,95} > \widehat{G}_{i,95}^{out-sample}$	$G_{i,00} > \widehat{G}_{i,00}^{out-sample}$
56.4%	51.3%	66.7%	79.5%
(B) using input measure of innovative capability and smoothed market size data			
$G_{i,85} > \widehat{G}_{i,85}^{out-sample}$	$G_{i,90} > \widehat{G}_{i,90}^{out-sample}$	$G_{i,95} > \widehat{G}_{i,95}^{out-sample}$	$G_{i,00} > \widehat{G}_{i,00}^{out-sample}$
61.5%	51.3%	64.1%	89.7%
(C) using output measure of innovative capability and non-smoothed market size data			
$G_{i,85} > \widehat{G}_{i,85}^{out-sample}$	$G_{i,90} > \widehat{G}_{i,90}^{out-sample}$	$G_{i,95} > \widehat{G}_{i,95}^{out-sample}$	$G_{i,00} > \widehat{G}_{i,00}^{out-sample}$
60.0%	60.0%	68.6%	80.0%
(D) using output measure of innovative capability and smoothed market size data			
$G_{i,85} > \widehat{G}_{i,85}^{out-sample}$	$G_{i,90} > \widehat{G}_{i,90}^{out-sample}$	$G_{i,95} > \widehat{G}_{i,95}^{out-sample}$	$G_{i,00} > \widehat{G}_{i,00}^{out-sample}$
62.8%	57.1%	68.6%	80.0%

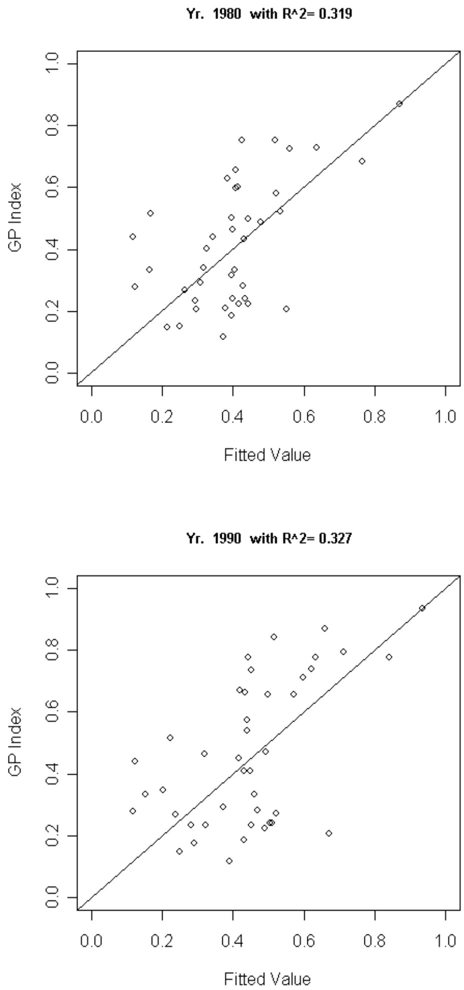
Table 6: Percentage of countries with actual patent protection above the out-of-sample predicted patent protection in various years

Tests on the null hypothesis of equal MSE in the one-step-ahead predictions of the structural model and various alternative models in the pre-TRIPS period								
(A) with input measure of innovative capability, market size data is non-smoothed								
alternative models	$\widehat{MSE}_{structural} - \widehat{MSE}_{alternative}$		Rivers-Vuong's test statistic		Sign test statistic		Wilcoxin's signed rank test statistic	
	1985	1990	1985	1990	1985	1990	1985	1990
linear	-0.0050	-0.0073	-0.2535	-0.2902	0.3123	1.3887	1.1014	2.1006
			(0.7998)	(0.7716)	(0.7548)	(0.1649)	(0.2707)	(0.0357)*
Maskus	0.0031	0.0108	0.0847	0.2112	1.8740	1.3887	2.6564	2.4882
			(0.9324)	(0.8327)	(0.0609)	(0.1649)	(0.0079)*	(0.0128)*
Ginarte and Park	0.0126	0.0201	0.3225	0.3775	1.8864	2.5724	2.7012	3.5047
			(0.7470)	(0.7057)	(0.0592)	(0.0101)*	(0.0069)*	(0.0005)*
(B) with input measure of innovative capability, market size data is smoothed								
linear	-0.0060	-0.0059	-0.2683	-0.2354	1.2493	0.7715	1.3995	1.4754
			(0.7884)	(0.8138)	(0.2115)	(0.4404)	(0.1617)	(0.1401)
Maskus	0.0019	0.0126	0.0556	0.2331	1.2493	1.0801	2.3713	2.5882
			(0.9556)	(0.8156)	(0.2115)	(0.2801)	(0.0177)*	(0.0096)*
Ginarte and Park	0.0115	0.0218	0.3107	0.3804	1.5434	2.5724	2.5473	3.3509
			(0.7560)	(0.7036)	(0.1227)	(0.0101)*	(0.0109)*	(0.0008)*
(C) with output measure of innovative capability, market size data is non-smoothed								
linear	-0.0021	0.0137	-0.0044	0.2857	0.0500	0.3202	0.2592	1.1233
			(0.9647)	(0.7750)	(0.6171)	(0.7488)	(0.7955)	(0.2613)
Maskus	-0.0003	0.0124	-0.0089	0.2397	2.1667	1.4599	2.7572	2.7409
			(0.9928)	(0.8104)	(0.0303)*	(0.1443)	(0.0058)*	(0.0061)*
Ginarte and Park	0.0105	0.0223	0.2749	0.4383	1.8864	2.5724	2.6328	3.3509
			(0.7833)	(0.6611)	(0.0592)	(0.0101)*	(0.0085)*	(0.0008)*
(D) with output measure of innovative capability, market size data is smoothed								
linear	-0.0016	0.0138	-0.0311	0.2582	0.1667	0.6405	0.1178	1.3327
			(0.9751)	(0.7661)	(0.8676)	(0.5218)	(0.9062)	(0.1826)
Maskus	0.0013	0.0105	0.0402	0.1864	1.5000	1.7844	2.4901	2.7264
			(0.9678)	(0.8521)	(0.1336)	(0.0744)	(0.0128)*	(0.0064)*
Ginarte and Park	0.0103	0.0204	0.3029	0.3592	1.2004	2.5724	2.4106	3.2141
			(0.7618)	(0.7194)	(0.2299)	(0.0101)*	(0.0159)*	(0.0013)*
Note: The numbers in parentheses are the p-values. "*" indicates that the null hypothesis is rejected at the 5 percent significance level.								

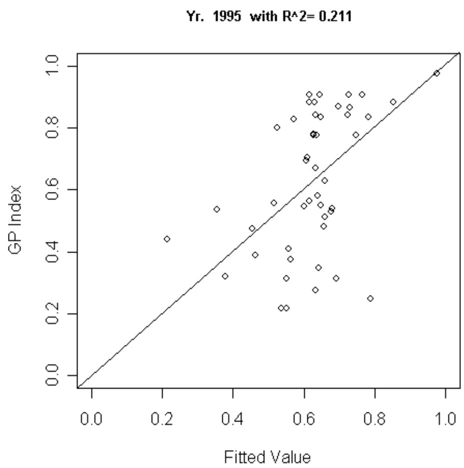
Table 7: Out-of-sample model validation: tests on the null hypothesis of equal MSE in the one-step ahead out-of-sample predictions of various models

Figure 1: (a) Actual and predicted patent protection levels (using input measure of innovative capability and non-smoothed market size data). [Note to referees: Only one of Figures 1(a)-(d) need to be published.]

Periods prior to the enforcement of TRIPS



Periods after the enforcement of TRIPS



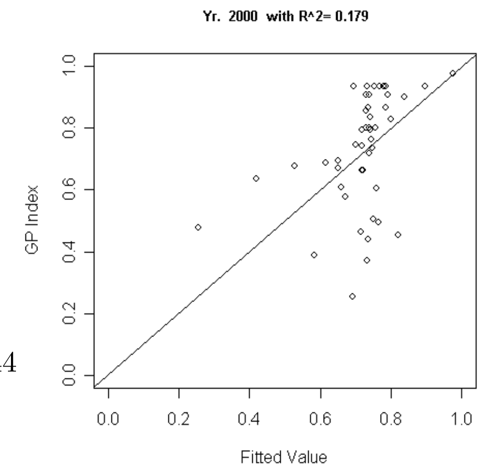
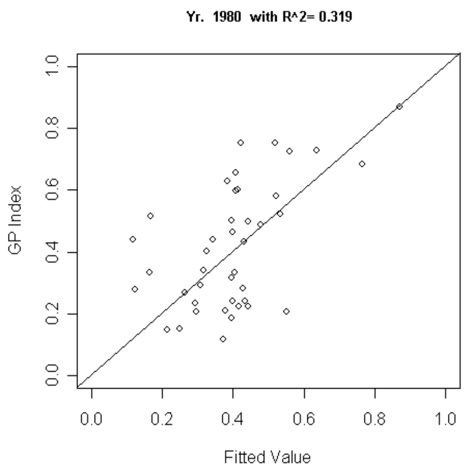
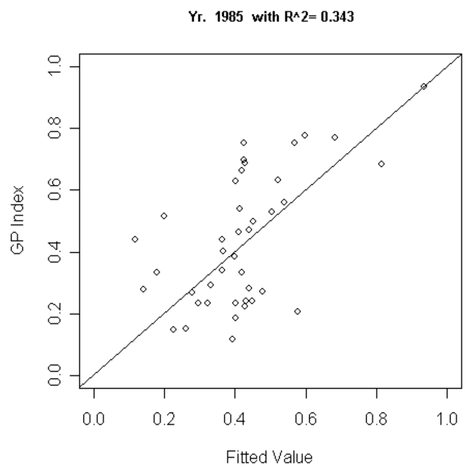
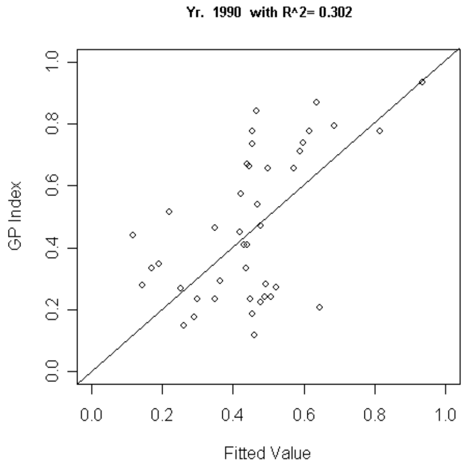


Figure 1: (b) Actual and predicted patent protection levels (using input measure of innovative capability and smoothed market size data)

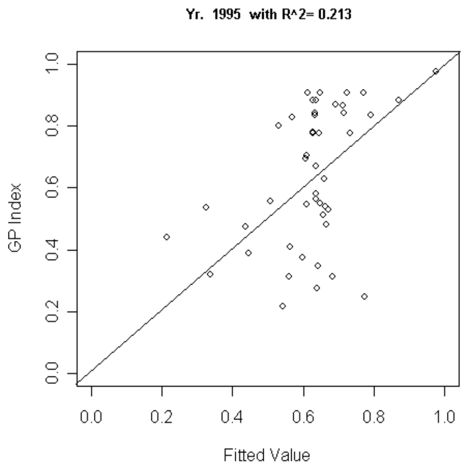
Periods prior to the enforcement of TRIPS







Periods after the enforcement of TRIPS



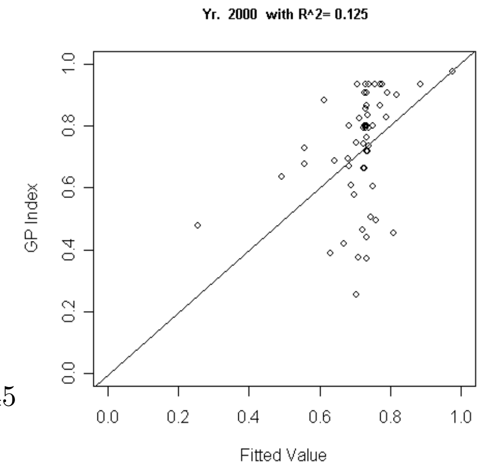
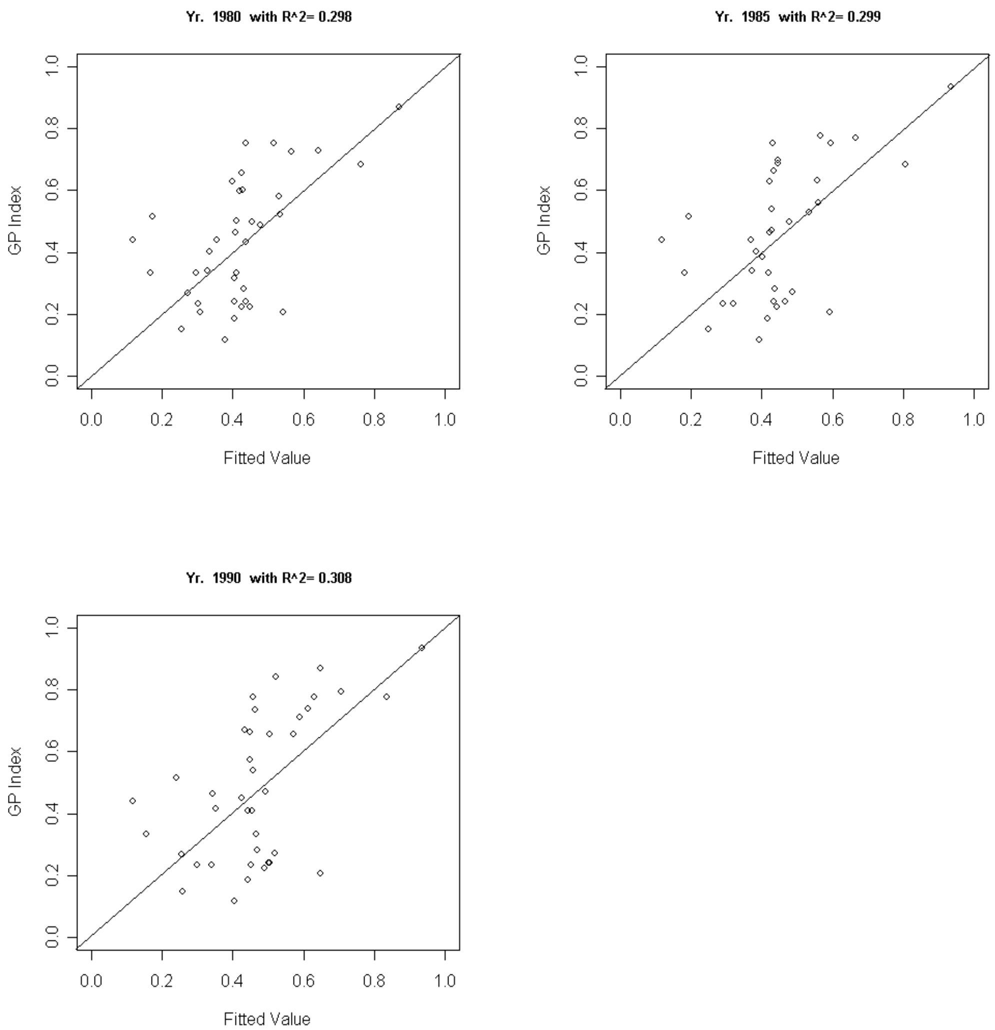
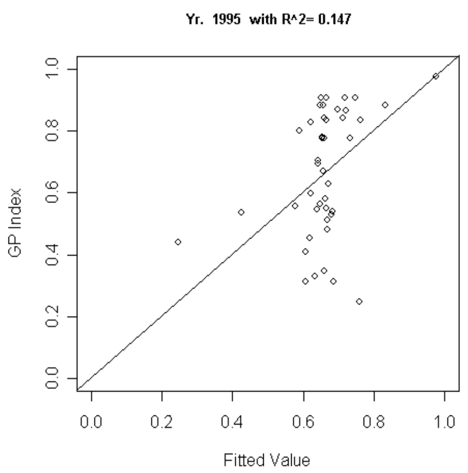


Figure 1: (c) Actual and predicted patent protection levels (using output measure of innovative capability and non-smoothed market size data)

Periods prior to the enforcement of TRIPS



Periods after the enforcement of TRIPS



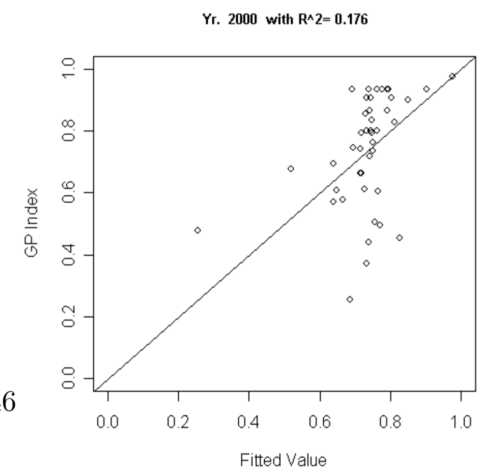
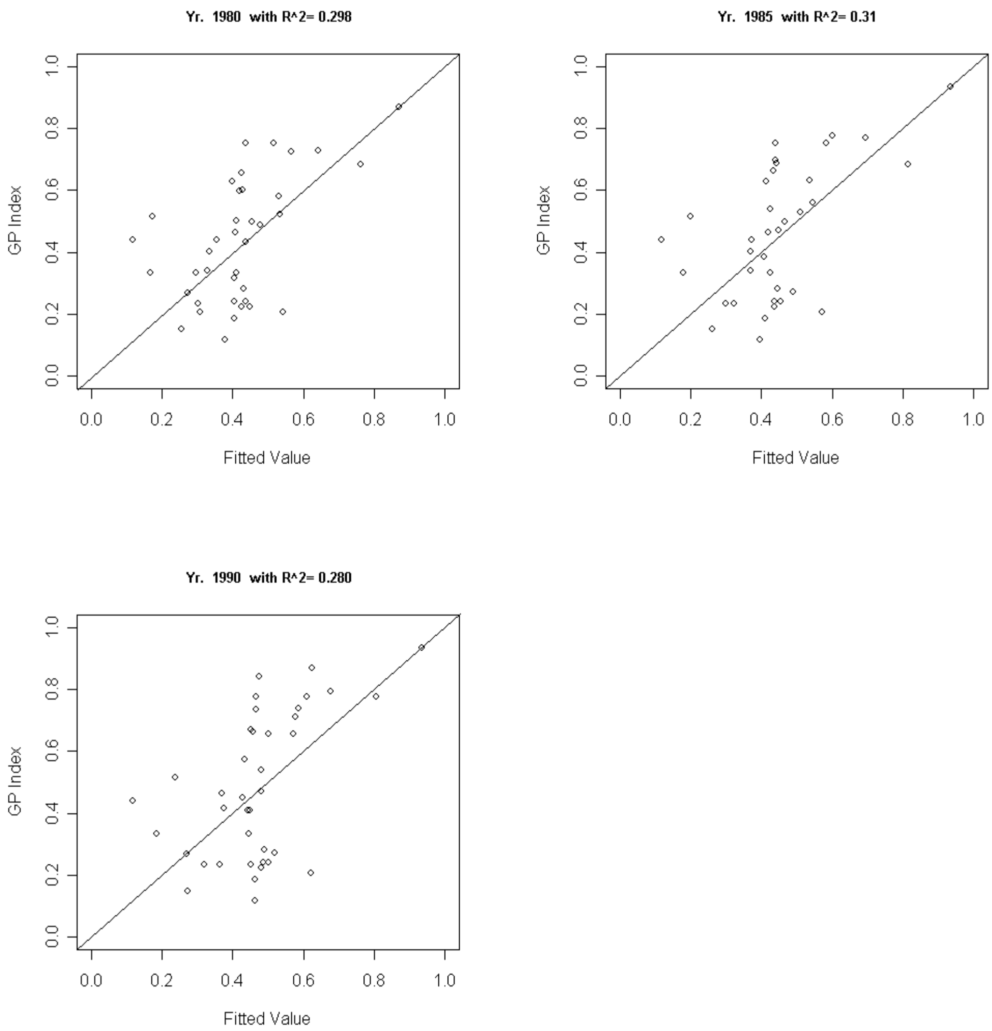


Figure 1: (d) Actual and predicted patent protection levels (using output measure of innovative capability and smoothed market size data)

Periods prior to the enforcement of TRIPS



Periods after the enforcement of TRIPS

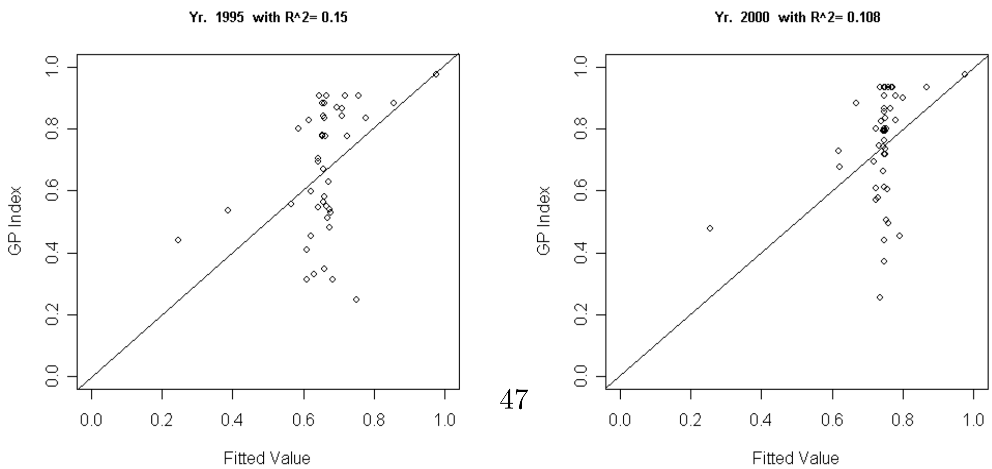


Figure 2: (a) Estimated smooth monotone functions h (using input measure of innovative capability and non-smoothed market size data) . [Note to referees: Only one of Figures 2(a)-(d) need to be published.]

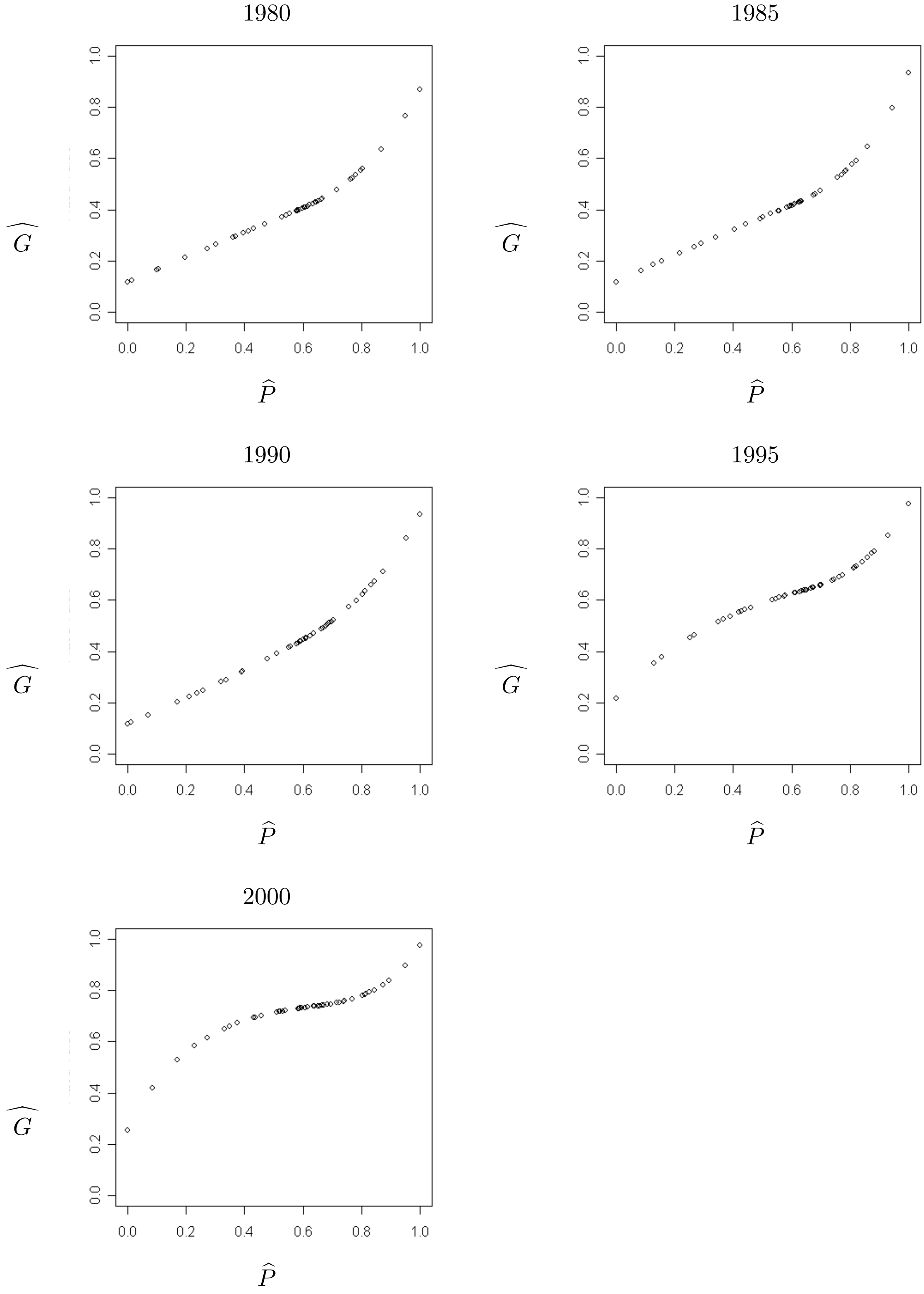


Figure 2: (b) Estimated smooth monotone functions h (using input measure of innovative capability and smoothed market size data)

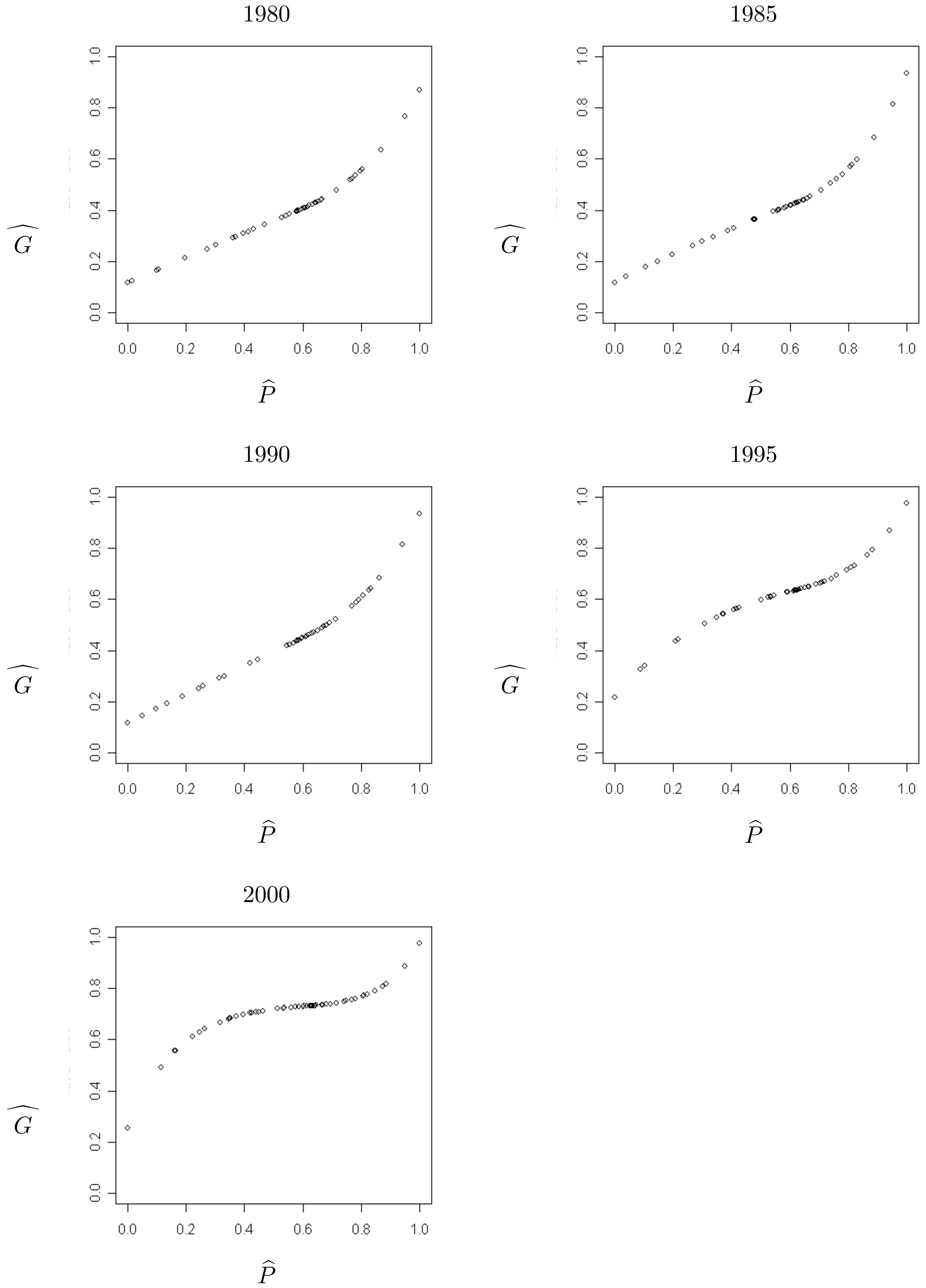


Figure 2: (c) Estimated smooth monotone functions h (using output measure of innovative capability and non-smoothed market size data)

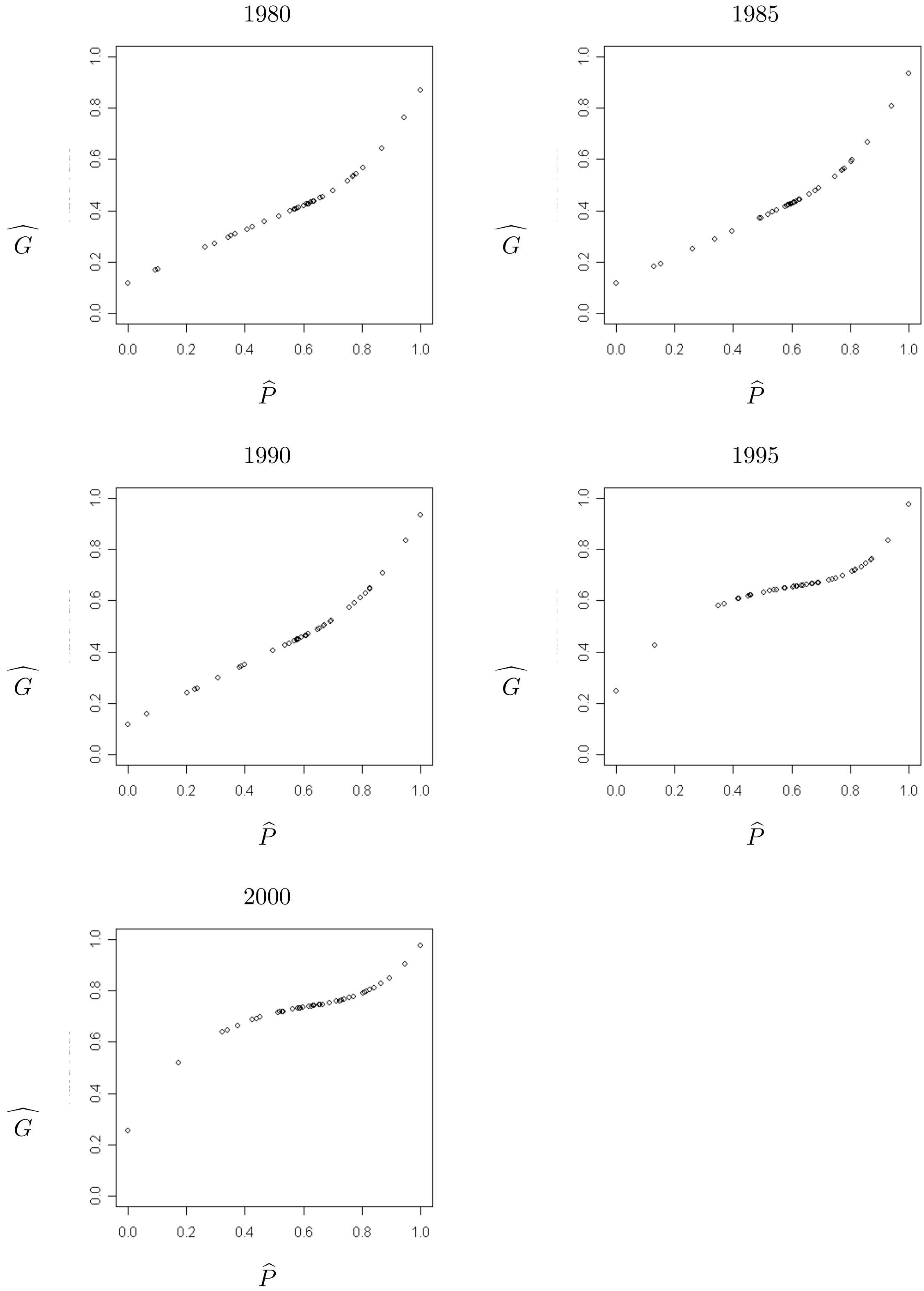


Figure 2: (d) Estimated smooth monotone functions h (using output measure of innovative capability and smoothed market size data)