

International Protection of Intellectual Property

By Gene M. Grossman and Edwin L.-C. Lai⁰

May 2004

Forthcoming, in *American Economic Review*

We study the incentives that governments have to protect intellectual property in a trading world economy. We consider a world economy with ongoing innovation in two countries that differ in market size and in their capacities for innovation. After describing the determination of national patent policies in a non-cooperative regime of patent protection, we ask, Why is intellectual property better protected in the North than in the South? We also study international patent agreements by deriving the properties of an efficient global regime of patent protection and asking whether harmonization of patent policies is necessary or sufficient for global efficiency. (JEL O34, F13)

During the 1980's and early 1990's, the United States and several European countries expressed strong dissatisfaction with what they deemed to be inadequate protection of intellectual property in many developing countries. The developed countries made the upgrading of intellectual property rights (IPRs) one of their highest priorities for the Uruguay Round of trade talks. Their efforts bore fruit in the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPs), which was approved as part of the Final Act of the Uruguay Round.

TRIPs establishes minimum standards of protection for several categories of intellectual property. For example, in the area of new technology, it essentially harmonizes

⁰Grossman: Department of Economics, Princeton University, Princeton, NJ 08544 (e-mail: grossman@princeton.edu); Lai: Department of Economics and Finance, City University of Hong Kong, Kowloon Tong, Hong Kong (e-mail: edwin.lai@cityu.edu.hk). The authors thank Doug Bernheim, Elhanan Helpman, Robert Lucas, Jr., Giovanni Maggi, Keith Maskus, Lars Svensson, and four anonymous referees for useful suggestions and comments. They are grateful to the U.S. National Science Foundation (grant nos. SES 9904480 and SES 0211748), the Hong Kong Research Grants Council (project nos. CityU 1145/99H and CityU 1235/00H), and the Department of Economics and Finance at City University of Hong Kong for financial support. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

patent policies by requiring countries to grant patents to a broad class of innovations for a minimum of twenty years and to treat foreign and domestic applicants alike. The agreement also imposes new standards of enforcement, including the requirement of border controls, civil and criminal penalties in some circumstances, and expedited procedures for disciplining infringement. But IPRs remain a highly contentious issue in international relations, because many developing countries believe that TRIPs was forced upon them by their economically more powerful trading partners and that this move toward harmonization of patent and other IPR policies serves the interests of the North at the expense of their own.

In a country that is closed to international trade, the design of a system of IPRs poses a clear trade-off to a welfare-maximizing government. By strengthening the protection of intellectual property, a government provides greater incentives for innovation and thus the benefits that come from having more and better products. But, at the same time, it curtails potential competition for firms that have previously innovated and thus limits the benefits that can be realized from existing products. As William D. Nordhaus (1969) argued, the optimal patent policy equates the marginal dynamic benefit with the marginal static efficiency loss.

But in an open economy, the trade-offs are not so clear cut. International trade spreads the benefits of innovation beyond national boundaries. This means that a country does not reap all of the global benefits that come from protecting intellectual property within its borders. Moreover, countries differ in their capacities for innovation due to differences in skill endowments and technical know-how. It is not obvious how a government ought to set its national IPR policy if some of the benefits of its national innovation accrue to foreigners, if its constituents benefit from innovations that are encouraged and take place beyond its boundaries, and if domestic and foreign firms differ in their ability to innovate.

Some previous research has addressed the question of whether a country with a limited capacity to innovate will benefit from extending IPRs to foreign inventors. Judith Chin and Gene M. Grossman (1990) and Alan V. Deardorff (1992) investigated the welfare effects of extending patent protection from a country in which innovation takes place to another country that only consumes the innovative products. These papers treat the investment in R&D as a once-off decision, whereas Elhanan Helpman (1993) models innovation as an ongoing process and associates the strength of the IPR regime

with the flow probability that a given product protected by a patent in the North will be imitated in the South. He evaluates the welfare consequences of marginal changes in the rate of imitation. These papers do not, however, consider the simultaneous choice of IPR protection by trade partners, nor do they discuss what international regime of IPR protection would be globally efficient.¹

In this paper, we study the incentives that governments have to protect intellectual property in a trading world economy. In our formal analysis, we associate the strength of IPR protection with an index variable that captures both the length of a country's patents and the stringency of its enforcement policies. We recognize that, in reality, governments choose packages of policies that include not only these instruments, but others such as patent breadth, limits on patentability, compulsory licensing requirements, copyright and trademark protections, and so on. We believe, however, that our analysis has broader relevance than would be implied by a literal interpretation of the model, because many of the factors that determine optimal patent duration and enforcement would also guide the choice of an optimal package of patent and copyright instruments in a setting in which governments choose from a wider range of policy tools.

We consider a world economy with ongoing innovation in which there are two countries that differ in market size and in their capacities for conducting research and development. Innovators develop the designs for new products, each of which has a limited economic life. Patents, when fully enforced, provide inventors with exclusive rights to produce, sell and distribute their products within a country. We study a regime with national treatment, which means that the same protection is afforded to all inventors irrespective of their nation of origin.

We begin in Section 1 with the case of a closed economy. There we re-examine the trade-off between static costs and dynamic benefits that was first studied by Nordhaus. We derive a neat formula that characterizes the optimal patent policy in a closed economy and discuss the determinants of the optimal strength of patent protection. One interesting finding is that the optimal index of patent protection may be independent of or even decreasing in the size of the economy.

In Section 2, we describe the determination of national policies in a non-cooperative

¹Phillip McCalman (2002) addresses some of these issues in a model of once-off innovation by a single firm in a developed economy. See also Edwin L.-C. Lai and Larry D. Qiu (2003), which we discuss at greater length below.

regime of patent protection. In this, we follow Lai and Qiu (2003), but with some differences that are better described after the model has been developed. After discussing the relationship between our paper and that one, we derive best response functions for the “North” and the “South,” where the North is assumed to have a higher wage than the South, as well as possibly a larger market for innovative products and a greater capacity for innovation. We characterize the nationally optimal policies, compare the incentives for providing protection for intellectual property in an open economy to those that exist in a closed economy, and explain the strategic interactions between countries in the setting of their patent policies.

In Section 3, we ask, Why is IPR protection stronger in the North? If the capacity for R&D is greater in the North than in the South and the market for innovative products is at least as large there, then patent protection will be stronger in the North than in the South in a Nash equilibrium. We explain why relative market size and relative productivity in innovation matter for the relative incentives to protect intellectual property. Patents are a more potent instrument for stimulating innovation in the relatively larger market. And a country that invents a smaller share of the world’s innovative products will find more incentive to ease IPR protection so as to benefit local consumers at the expense of producers.

We study international patent agreements in Section 4. First we derive the properties of an efficient global regime of IPRs. An efficient patent regime is one that provides the optimal aggregate incentives for innovation to inventors throughout the world. These incentives can be achieved by various combinations of patent policies in the two countries, so there is no unique set of policies that is needed for global efficiency. However, different ways of achieving the optimal aggregate incentives have different implications for the distribution of welfare between the North and the South. Among combinations of policies that give the same overall incentives for global research, the North fares better, and the South worse, the stronger is patent protection in the South. An implication of our findings is that harmonization of patent policies is neither necessary nor sufficient for global efficiency. Moreover, starting from a non-cooperative equilibrium with stronger IPR protection in the North than in the South, an efficient agreement calling for harmonization of patent policies benefits the North quite possibly at the expense of the South.

In Section 5, we extend our analysis of both the non-cooperative and cooperative set-

tings to a world with many trading countries. The many-country model is qualitatively similar to the two-country model, although the addition of more countries exacerbates the inefficiencies associated with non-cooperation. Our findings are summarized in Section 6.

I. A Simple Model of Innovation

In this section, we construct a simple model of ongoing innovation. We develop the model for a closed economy and use it to revisit the question of optimal patent policy that was first addressed by Nordhaus (1969). Our model yields a neat formula that characterizes the trade-off between the static costs and dynamic benefits of strengthening patent protection. The discussion of a closed economy lays the groundwork for the more subtle analysis of the international system that we undertake in the sections that follow.

The economy has two sectors, one that produces a homogeneous good and another that produces a continuum of differentiated products. The designs for the differentiated products result from private investments in R&D. Once a good has been invented, it has a finite economic life of length $\bar{\tau}$. That is, a new product potentially provides utility to consumers for a period of length $\bar{\tau}$ from the time of its creation, whereupon its value to consumers drops to zero.

There are M consumers with identical preferences. We shall refer to M as the “size of the market.”² The representative consumer maximizes a utility function of the form

$$(1) \quad U(t) = \int_t^\infty u(z)e^{-\rho z} dz$$

where

$$(2) \quad u(z) = y(z) + \int_0^{n(z)} h[x(i, z)] di,$$

$y(z)$ is consumption of the homogeneous good at time z , $x(i, z)$ is consumption of the i^{th} variety of differentiated product at time z , and $n(z)$ is the measure of differentiated products invented before z that still hold value to consumers at time z . We assume

²In our model, demand for differentiated products does not vary with income. Thus, a rich country need not have a larger market for these goods than a poor country. Nonetheless, we prefer to think of the market for differentiated goods as being larger in the North than in the South. This could be rigorously justified within our model if we were to suppose that differentiated products provide utility only after a threshold level of consumption of the homogeneous goods has been reached. Then, a rich country may have more consumers who surpass the threshold.

that $h'(x) > 0$, $h''(x) < 0$, $h'(0) = \infty$, and $-xh''(x)/h'(x) < 1$ for all x . The third assumption ensures a positive demand for every variety at any finite price. The fourth ensures that any firm producing a differentiated product charges a finite price.

A consumer maximizes utility by purchasing some of all varieties that are not yet obsolete. He chooses $x(i, z)$ so that $h'[x(i, z)] = p(i, z)$ for all i and z , where $p(i, z)$ is the price of variety i at time z . After the consumer makes all of his optimal purchases of differentiated products at time z , he devotes the remainder of his spending to the homogeneous good y . Spending is always positive in the equilibria we describe. This means that the interest rate is constant and equal to ρ , in view of the perfect intertemporal substitutability of good y in the preferences described by (1) and (2).

Manufacturing requires only labor. Any firm can produce good y with a units of labor per unit of output. All known varieties of the differentiated product also can be produced with a units of labor per unit of output.

The design of new varieties requires both labor and human capital. We take $\phi(z) \equiv F[H, L_R(z)] = \{b[L_R(z)/a]^\beta + (1-b)H^\beta\}^{1/\beta}$, where $\phi(z)$ is the flow of new inventions at time z , H is the (constant) stock of human capital, $L_R(z)$ is the amount of labor devoted to R&D, and a is a measure of labor productivity as before. This is, of course, a production function with a constant elasticity of substitution between labor and human capital. We assume that $\beta \leq 1/2$, or equivalently that the elasticity of substitution is less than or equal to two. This assumption is sufficient (but not necessary) to ensure that any patent policy that satisfies the first-order condition for an interior optimum also satisfies the second-order condition. Note that $\dot{n}(z) = \phi(z) - \phi(z - \bar{\tau})$, because the goods that were invented at time $z - \bar{\tau}$ become obsolete at time z .

The government grants the original designer of a differentiated product a patent of length $\tau \leq \bar{\tau}$. The government also chooses the vigor of its enforcement policy, which we index by $\omega \in [0, 1]$. For simplicity, we take ω to be the probability that a non-expired patent is enforced by the government at any moment in time.³ If a live patent is enforced by the government, the holder enjoys exclusive rights to produce and sell the protected product in the local market.

We describe now the static and dynamic equilibrium for an economy that has a patent duration of τ and an enforcement rate of ω . In equilibrium, firms with enforced

³Alternatively, we might interpret ω as being the fraction of the country's territory in which the patent is enforced.

patents behave as monopolies. Each such firm faces an inverse demand curve from each of the M consumers with the form $p(x) = h'(x)$. The firm sets its price so that $(p - aw)/p = -xh''/h'$, where w is the wage rate and x is sales per consumer. This is the usual monopoly-pricing rule whereby the markup over unit cost as a fraction of the price is equal to the inverse demand elasticity. Optimal pricing yields the typical holder of an enforced patent profits of π per consumer and total profits of $M\pi$.

When a live patent is not enforced, competitors imitate the good costlessly. Then the product sells for the competitive price of $p = aw$ and generates no profits for the inventor. Similarly, once a patent expires, the price of the good falls to the competitive level and remains there until the good becomes obsolete. The homogeneous good always carries the competitive price of aw , which, because this good is the numeraire, implies that $w = 1/a$. In writing this condition, we implicitly assume that the economy's labor supply is sufficiently large that some labor remains for production of the homogeneous good after all derived demand for labor for producing differentiated products and conducting R&D has been satisfied.

Labor engages in manufacturing and R&D. The labor employed in manufacturing differentiated goods is just the amount needed to produce the quantities demanded at the equilibrium prices. The allocation of labor to R&D is such that its marginal value product in this activity is equal to the wage rate. Thus,

$$(3) \quad vF_L(H, L_R) = w,$$

where v is the value of a new patent. A patent is worth the discounted value of the expected profits it generates in the time before it expires, or

$$(4) \quad v = \frac{\omega M\pi}{\rho} (1 - e^{-\rho\tau}).$$

We see from (3) and (4) that an increase in either patent length or patent enforcement increases the value of a new patent, thereby drawing additional resources into R&D.

The final equilibrium condition equates savings with investment. Savings are the difference between national income $rH + wL + n_m M\pi$ and aggregate spending E , where r is the return to human capital, L is the aggregate labor supply, and n_m is the number of firms that hold live and enforced patents at a given time. All investment is devoted to R&D. This activity has an aggregate cost of $rH + wL_R$. Thus, we can write the

equilibrium condition as $(rH + wL + n_m M\pi) - E = rH + wL_R$, or

$$(5) \quad E = w(L - L_R) + n_m M\pi.$$

Before proceeding to the government's maximization problem, it is useful to define an index of the strength of IPR protection afforded by the vector of policies (τ, ω) . First, define $T \equiv (1 - e^{-\rho\tau})/\rho$ as the present discounted value of a flow of one dollar from time 0 to time τ . Then define $\Omega \equiv \omega T$. Clearly, (4) can be re-written as $v = M\pi\Omega$. Thus, Ω captures the overall incentives that government policy provides for the creation of new varieties. In what follows we use Ω as a measure of the strength of IPR protection.

Next, we derive an expression for aggregate welfare at date 0, the time at which a new (optimal) patent policy will be set by the government. By assumption, this patent protection applies only to goods introduced after time 0; those introduced beforehand are subject to whatever policy was in effect at the time of their invention.⁴ Our derivation anticipates a stationary equilibrium in which the allocation of labor to R&D and thus the rate of innovation, ϕ , are constant over time. We substantiate this hypothesis in footnote 5 below.

At any moment, each consumer enjoys surplus of $C_m = h(x_m) - p_m x_m$ from his consumption of any good whose live patent is enforced and surplus of $C_c = h(x_c) - p_c x_c$ from any differentiated product that is competitively priced, either because its patent has expired or because it is not enforced. Here, x_m and x_c are the amounts purchased by the consumer from a typical monopoly and competitive supplier, respectively, and p_m and p_c are the respective prices. We distinguish now between goods invented before time 0 and those invented afterward. The former yield some exogenous surplus that is unaffected by the new patent regime. Each good in the latter category yields each of the M consumers a discounted surplus of $C_m\Omega + C_c(\bar{T} - \Omega)$ over its useful life, where $\bar{T} \equiv (1 - e^{-\rho\bar{\tau}})/\rho$. Using (1), (2) and (5), and assuming that new goods are introduced at a constant flow ϕ after time 0, we calculate that aggregate welfare at time 0 is

$$(6) \quad W(0) = \Lambda_0 + \frac{w(L - L_R)}{\rho} + \frac{M\phi}{\rho}(C_m + \pi)\Omega + \frac{M\phi}{\rho}C_c(\bar{T} - \Omega),$$

⁴It would never be optimal for the government to provide patent protection on goods that have already been invented. This would create deadweight loss without any offsetting social benefit. The government might wish to eliminate protection for goods that were invented under a different regime, but we assume that such expropriation of intellectual property would not be legal.

where Λ_0 is the discounted present value of the consumer surplus and profits derived from goods invented before time 0.

We are now ready to derive the optimal patent policy for a closed economy. Note that aggregate welfare depends only on the index of patent protection Ω and not separately on the underlying policies τ and ω . This implies that patent duration and patent enforcement are perfect substitutes as instruments of IPR protection and it justifies our use of Ω as a measure of the strength of the IPR regime. Formally, we maximize $W(0)$ with respect to Ω , after recalling that $\phi = F(H, L_R)$ and that L_R is a function of Ω via (3) and (4).⁵ It is more intuitive, however, to describe the social costs and benefits that derive from a marginal increase in patent protection from a given strength of protection Ω . The cost of strengthening patent protection is that the economy suffers the deadweight loss of $M(C_c - C_m - \pi)$ on each newly invented good for a greater fraction of the product's life. Considering that the constant flow of new products is ϕ per unit time, the total discounted marginal cost on all goods invented after time 0 is

$$\frac{\phi}{\rho} M(C_c - C_m - \pi).$$

The benefit to the economy of strengthening patent policy is that it encourages R&D, which in turn means a greater variety of differentiated products. Each differentiated product yields discounted consumer surplus of $MC_m\Omega$ during times when its patent is active and enforced and surplus of $MC_c(\bar{T} - \Omega)$ during times when it is competitively priced, where in each case the discounting is back to the time of invention. Now if we discount this flow of benefits back to time 0, and multiply by the number of new

⁵Equivalently, we can maximize $\rho W(0)$ over the choice of Ω . Note that C_m , C_c and π do not depend on the strength of patent protection and thus do not depend on Ω . We can combine (3) and (4) to write $M\pi\Omega F_L(H, L_R) = w$, which allows us to solve for the functional relationship between the labor devoted to R&D and the policy variable Ω ; denote it by $L_R(\Omega)$. Then, substituting this expression into (6) and rearranging terms, we can write the maximand as

$$\rho W(0) = \rho\Lambda_0 + w[L - L_R(\Omega)] + MF[H, L_R(\Omega)] [(C_m + \pi - C_c)\Omega + C_c\bar{T}].$$

The first-order condition for a maximum requires

$$(C_c - C_m - \pi)MF[H, L_R(\Omega)] = \{MF_L [(C_m + \pi - C_c)\Omega + C_c\bar{T}] - w\} L'_R$$

from which (7) follows. In the appendix we show $\beta \leq 1/2$ is sufficient to ensure that the second-order condition is satisfied at any value of Ω that satisfies the first-order condition (7).

inventions induced by a marginal strengthening of the patent regime, we have the total marginal benefit, which is equal to

$$\frac{1}{\rho} \cdot \frac{d\phi}{dv} \cdot \frac{dv}{d\Omega} \cdot [MC_m\Omega + MC_c(\bar{T} - \Omega)].$$

Using (3) we calculate that

$$\frac{d\phi}{dv} = \gamma \frac{\phi}{v},$$

where γ is the ratio of the elasticity of research output with respect to labor to the elasticity of the marginal product of labor in R&D; i.e., $\gamma \equiv -(F_L)^2 / (FF_{LL})$. The variable γ identifies the responsiveness of innovation to the protections afforded by the patent system. In general, it is a function of L_R and thus indirectly of the strength of patent protection Ω . With the CES research technology, $\gamma = [b/(1-b)(1-\beta)](L_R/aH)^\beta$. For the special case of a Cobb-Douglas technology (which is the limiting case of the CES as $\beta \rightarrow 0$), γ is a constant equal to the ratio of the cost share of labor to the cost share of human capital.

Next, we use (4) to compute that

$$\frac{dv}{d\Omega} = M\pi.$$

Substituting for $d\phi/dv$ and $dv/d\Omega$ in the expression for marginal benefit, and equating the result to the marginal cost, we derive an implicit formula for the optimal index of patent protection. We find that

$$(7) \quad C_c - C_m - \pi = \gamma \left[C_m + C_c \left(\frac{\bar{T} - \Omega}{\Omega} \right) \right]$$

at an (interior) optimal value of Ω .

From (7) we see that the optimal patent protection is stronger, the greater is the useful life of a product (larger $\bar{\tau}$), the more patient are consumers (smaller ρ), and the higher is the responsiveness of innovation to strengthened patent protection.⁶ All of these findings accord well with intuition. One noteworthy feature of (7) is the relationship between market size and the optimal strength of patent protection. In a closed economy, the first-best level of R&D — that which maximizes discounted utility when all goods are competitively priced — typically is an increasing function of market size. This is because

⁶The proof of these statements makes use of the second-order condition, which ensures that the right-hand side of (7) is a declining function of Ω .

innovation is a public good, and the Samuelsonian rule for optimal provision of a public good calls for greater output when the benefits can be spread across more consumers. But the encouragement of innovation by patents achieves only a second best. According to (7), the size of the market M affects the optimal index of patent protection only through its effect on the supply elasticity of innovations. If γ is an increasing function of L_R , as it will be if $1/2 > \beta > 0$, then the optimal Ω is an increasing function of M . But if γ is a decreasing function of L_R , as it will be if $\beta < 0$, then the optimal Ω is a decreasing function of M . In the benchmark Cobb-Douglas case (with $\beta = 0$), γ is independent of L_R and therefore of market size. Then an increase in M enhances the marginal benefit of strengthening patent protection and the marginal cost of doing so in equal proportions. The optimal strength of patent protection in a closed economy with a Cobb-Douglas research technology is invariant to market size.

II. Noncooperative Patent Protection

In this section, we study the national incentives for protection of intellectual property in a world economy with imitation and trade. We derive the Nash equilibria of a game in which two countries set their patent policies simultaneously and noncooperatively. The countries are distinguished by their wage rates, their market sizes, and their stocks of human capital. The last of these proxies for their different capacities for R&D. We shall term the countries “North” and “South,” in keeping with our desire to understand the tensions that surrounded the tightening of IPR protection in the developing countries in the last decade. Keith E. Maskus (2000a, ch.3) has documented an increase in innovative activity in poor and middle-income countries such as Brazil, Korea, and China, so our model of relations between trading partners with positive but different abilities to conduct R&D may be apt for studying the incentives for IPR protection in a world of trade between such nations and the developed economies.⁷ But our model may apply more broadly to relations between any groups of countries that have different wages and different capacities for research. Such differences exist, albeit to a lesser extent than between North and South, in the comparison of countries in Northern and Southern Europe, or the comparison of the United States and Canada. We do not mean the labels North and South to rule out the application of our analysis to these other

⁷He also shows the extent to which patent protection in countries like Mexico, Brazil, Korea, Malaysia, Indonesia and Singapore are dominated by foreign firms, a feature of the data that figures in our analysis.

sorts of relationships.

Before proceeding, we pause to discuss the relationship between the analysis that will be presented in this section and that contained in Lai and Qiu (2003), which covers some similar ground. Lai and Qiu also examine a non-cooperative game between two governments that set national patent policies. They do not, however, model how the resources used in innovation translate into new goods, but rather assume that at each moment in time the different goods that might be invented have different (exogenous) innovation costs. Specifically, the cost of making ϕ innovations is assumed to be an increasing, power function of ϕ . They also assume that demand for each differentiated product is iso-elastic and that the discount rate is zero. With our more general utility function and our explicit treatment of the R&D technology, we are able to explain when and why the countries' patent policies are strategic substitutes or strategic complements based on the technology for innovation—specifically, the substitutability between human capital and mobile resources in the R&D activity. Perhaps more importantly, we are able to provide an intuitive discussion of why and how a government's incentives for protecting intellectual property in an open economy differ from those it would face in a world without trade, whereas Lai and Qiu focus only on the algebra of a single, parameterized example of a trading economy.

A. The Global IPR Regime

The model is a natural extension of the one presented in Section 1. Consumers in the two countries share identical preferences. In each country, the representative consumer maximizes the intertemporal utility function in (1). The instantaneous utility of a consumer in country j now is given by

$$(8) \quad u_j(z) = y_j(z) + \int_0^{n_S(z)+n_N(z)} h[x_j(i, z)] di,$$

where $y_j(z)$ is consumption of the homogeneous good by a typical resident of country j at time z , $x_j(i, z)$ is consumption of the i^{th} differentiated product by a resident of country j at time z , and $n_j(z)$ is the number of differentiated varieties previously invented in country j that remain economically viable at time z . There are M_N consumers in the North and M_S consumers in the South. While we do not place any restrictions on the relative sizes of the two markets at this juncture, we shall be most interested in the case

where $M_N > M_S$.⁸ It does not matter for our analysis whether consumers can borrow and lend internationally or not.

In country j , it takes a_j units of labor to produce one unit of the homogeneous good or to produce one unit of any variety of the differentiated product. New goods are invented in each region according to $\phi_j = F(H_j, L_{Rj}/a_j) = \left[b(L_{Rj}/a_j)^\beta + (1-b)H_j^\beta \right]^{1/\beta}$, where H_j is the human capital endowment of country j , L_{Rj} is the labor devoted to R&D there, and again we take $\beta \leq 1/2$. We assume that $a_N < a_S$, which means that labor is uniformly more productive in the North than in the South.⁹ We also assume that the numeraire good is produced in positive quantities in both countries, so that $w_j = 1/a_j$ for $j = S, N$, and hence $w_N/w_S = a_S/a_N > 1$.

We now describe the IPR regime. In each country, there is *national treatment* in the granting of patent rights. Under national treatment, the government of country j affords the same protection $\Omega_j = \omega_j T_j$ to all inventors of differentiated products regardless of their national origins, where ω_j is the probability that a patent is enforced in country j (or the fraction of country j 's market where a patent is enforced) at any moment in time, $T_j = (1 - e^{-\rho\tau_j})/\rho$, and τ_j is the length of the patents granted by country j . In other words, we assume that foreign firms and domestic firms have equal standing in applying for patents in any country and that all patents are subject to the same enforcement provisions. National treatment is required by TRIPs and it characterized the laws that were in place in most countries even before this agreement.¹⁰ In our model, a patent is an exclusive right to make, sell, use, or import a product for a fixed period of time (see Maskus, 2000a, p.36). This means that, when good i is under patent protection in country j , no firm other than the patent holder or one designated by it may legally produce the good in country j for domestic sale or for export, nor may the good be

⁸We remind the reader that market size is meant to capture not the population of a country, but rather the scale of its demand for innovative products.

⁹In Grossman and Lai (2002), we allowed the relative productivity of labor to vary in different uses; i.e., we allowed for Ricardian comparative advantage across the regions. This feature caused some subtle complications that we ignore here for the sake of simplicity and greater clarity.

¹⁰National treatment is required by the Paris Convention for the Protection of Industrial Property, to which 127 countries subscribed by the end of 1994 and 164 countries subscribe today (see <http://www.wipo.org/treaties/ip/paris/index.html>). There were, however, allegations from firms in the United States and elsewhere that prior to the signing of TRIPs in 1994, nondiscriminatory laws did not always mean nondiscriminatory practice. See Suzanne Scotchmer (2004) for an analysis of the incentives that countries have to apply national treatment in the absence of an enforceable agreement.

legally imported into country j from an unauthorized producer outside the country. We also rule out parallel imports — unauthorized imports of good i that were produced by the patent holder or its designee, but that were sold to a third party outside country j .¹¹ When parallel imports are prevented, patent holders can practice price discrimination across national markets.

We solve the Nash game in which the governments set their patent policies once-and-for-all at time 0. These patents apply only to goods invented after time 0; goods invented beforehand continue to receive the protections afforded at their times of invention. So long as the governments cannot remove protections that were previously granted, the economy has no state variables that bear on its choice of optimal patent policies at a given moment in time. This means that the Nash equilibrium in once-and-for-all patents is also a sub-game perfect equilibrium in the infinitely repeated game in which the governments can change their patent policies periodically, or even continuously. Of course, the repeated game may have other equilibria in which the governments base their current policies on the history of prior actions. We do not investigate such equilibria with tacit cooperation here, but rather postpone our discussion of cooperation until Section 4.

Let us describe, for given patent strengths Ω_N and Ω_S , the life cycle of a typical differentiated product. During an initial phase after the product is introduced, the inventor holds an active patent in both countries which is only partially enforced. The patent holder earns an expected flow of profits of $\omega_N M_N \pi$ from sales in the Northern market and an expected flow of profits of $\omega_S M_S \pi$ from sales in the Southern market, where π is earnings per consumer for a monopoly selling a typical brand. Notice that monopoly profits per consumer are the same for sales in both markets, because consumers share identical preferences. Also, they do not depend on where a good was invented or where it is produced, because the productivity gap between the countries exactly offsets

¹¹The treatment of parallel imports under TRIPs remains a matter of legal controversy. Countries continue to differ in their rules for territorial exhaustion of IPRs. Some countries, like Australia and Japan, practice international exhaustion, whereby the restrictive rights granted by a patent end with the first sale of the good anywhere in the world. Other countries or regions, like the United States and the European Union, practice national or regional exhaustion, whereby patent rights end only with the first sale within the country or region. Under such rules, patent holders can prevent parallel trade. See Maskus (2000b) for further discussion.

the wage differential.¹² Each Northern consumer realizes a flow of expected surplus of $\omega_N C_m + (1 - \omega_N) C_c$ from his purchases of the good, where C_m is the surplus that a consumer derives from purchases of a good produced at a cost of $w_j a_j = 1$ and sold at the monopoly price p_m and C_c is the surplus he derives from a product sold for the competitive price of $p_c = 1$. Similarly, a Southern consumer realizes an expected flow of consumer surplus of $\omega_S C_m + (1 - \omega_S) C_c$ from his purchases of the good.

After a while, the patent will expire in one country. For concreteness, let's say that this happens first in the South. Then the good will be legally imitated by competitive firms producing there, for sales in the local (Southern) market. The imitators will not, however, be able to sell the good legally in the North, because the live patent there, if enforced, affords protection from such infringing imports. When the patent expires in the South, the price of the good falls permanently to $w_S a_S = 1$, and the original inventor ceases to realize profits in that market. The flow of consumer surplus in the South rises to $M_S C_c$.

Eventually, the inventor's patent expires in the North. Then the Northern market can be served completely by competitive firms producing in either location. At this time, the price of the good in the North falls to $p_c = 1$ and households there begin to enjoy the higher flow of consumer surplus $M_N C_c$. The original inventor loses his remaining source of monopoly income. Finally, after a period of length $\bar{\tau}$ has elapsed from the moment of invention, the good becomes obsolete and all flows of consumer surplus cease.

B. The Best Response Functions

We are now ready to derive the best response functions for the two governments. The best response expresses the strength of patent protection that maximizes a country's aggregate welfare as a function of the given patent policy of its trading partner. Consider the choice of Ω_S by the government of the South. This country bears two costs from strengthening its patent protection slightly. First, it expands the fraction of goods previously invented in the South on which the country suffers a static deadweight loss of $M_S(C_c - C_m - \pi)$. Second, it augments the fraction of goods previously invented in the North on which its consumers realize surplus of $M_S C_m$ instead of $M_S C_c$. Notice that the profits earned by Northern producers in the South are not an offset to this latter

¹²In Grossman and Lai (2002), where we allowed for comparative technological advantages across different uses of labor, we were forced to consider separately situations in which direct foreign investment is and is not a possibility. However, our conclusions were qualitatively similar for the two cases.

marginal cost, because they accrue to patent holders in the North. The marginal benefit that comes to the South from strengthening its patent protection reflects the increased incentive that Northern and Southern firms have to engage in R&D. If the welfare-maximizing Ω_S is positive and less than \bar{T} , then the marginal benefit *per consumer* of increasing Ω_S must match the marginal cost, which implies

$$(9) \quad \phi_S(C_c - C_m - \pi) + \phi_N(C_c - C_m) = \frac{\gamma_S \phi_S + \gamma_N \phi_N}{v} M_S \pi [C_m \Omega_S + C_c(\bar{T} - \Omega_S)],$$

where $v = (M_S \Omega_S + M_N \Omega_N) \pi$ is the value of a new patent and γ_j is the responsiveness of innovation in region j to changes in the value of a patent (in elasticity form).

Similarly, in the North, the marginal benefit of strengthening patent protection must match the marginal cost at any interior point on the best response curve. The marginal cost in the North is different from that in the South, because the North's national income includes the profits earned by Northern patent holders but not those earned by Southern patent holders. The marginal benefit differs too, because the effectiveness of patent policy as a tool for promoting innovation varies according to the importance of a country's market in the aggregate profits of potential innovators and because the surplus from a typical product over its lifetime depends upon a country's patent regime. The condition for the best response of the North, analogous to (9) above, is

$$(10) \quad \phi_S(C_c - C_m) + \phi_N(C_c - C_m - \pi) = \frac{\gamma_S \phi_S + \gamma_N \phi_N}{v} M_N \pi [C_m \Omega_N + C_c(\bar{T} - \Omega_N)].$$

Noting that $\gamma_S = \gamma_N = \gamma$,¹³ the two best response functions can be written similarly

¹³The fact that the two supply elasticities γ_S and γ_N are equal despite the differences in human capital endowments, in employment, and in labor productivity is a property of the CES research technology. It follows from the observation that

$$\gamma_i = \frac{b}{(1-b)(1-\beta)} \left(\frac{L_{Ri}}{a_i H_i} \right)^\beta$$

and $v F_L(L_{Ri}/a_i, H_i) = w_i$, or

$$\frac{vb}{a_i} \left[b + (1-b) \left(\frac{L_{Ri}}{a_i H_i} \right)^{-\beta} \right]^{\frac{1-\beta}{\beta}} = \frac{1}{a_i}.$$

Combining the two, we find $\gamma_S = \gamma_N = \gamma$, where

$$\gamma = \frac{b}{1-\beta} \left[\left(\frac{1}{bv} \right)^{\frac{\beta}{1-\beta}} - b \right]^{-1}.$$

as

$$(11) \quad C_c - C_m - \mu_i \pi = \gamma \frac{M_i \Omega_i}{M_S \Omega_S + M_N \Omega_N} \left[C_m + C_c \left(\frac{\bar{T} - \Omega_i}{\Omega_i} \right) \right] \quad \text{for } i = S, N,$$

where $\mu_i = \phi_i / (\phi_S + \phi_N)$ is the share of world innovation that takes place in country i . This form of the best response function facilitates a comparison of the incentives that a government has for protecting intellectual property in a world with trade compared to those that exist when there is no trade, as expressed in (7). On the left-hand side of (11), the government of an open economy considers only a fraction of the profits that flow to patent holders to be an offset to the static cost of patent protection. On the right-hand side, the ability of an open economy to stimulate innovation with a given change in patent protection is a fraction of what it is in a closed economy, because inventors earn only part of their discounted profits within the country's borders. Both of these forces point to weaker patent protection in an open economy than would be optimal in the absence of trade. Against this, possibly, is the difference between the supply elasticities for innovation in the closed and open economies; the presence of a foreign country offering protection for innovators may increase the responsiveness of innovation to home patent policy if γ is an increasing function of L_R . However, with the CES research technology γ is in fact a *non-increasing* function of L_R whenever $\beta \leq 0$; i.e., when the elasticity of substitution between human capital and labor is less than or equal to one. It follows that¹⁴

Proposition 1 *Let the research technology be $\phi_i = [b[L_{Ri}/a_i]^\beta + (1-b)H_i^\beta]^{1/\beta}$ in country i , for $i = S, N$. If $\beta \leq 0$, patent protection is weaker in each country in any Nash equilibrium than it would be if the country were closed to international trade.*

If the research technology in each country takes a Cobb-Douglas form (i.e., $\beta = 0$), the supply elasticity γ is a constant equal to $b/(1-b)$. Moreover, $\mu_i = H_i/(H_S + H_N)$

¹⁴Suppose the government of an open economy were to choose the autarky strength of patent protection. The marginal cost of strengthening protection would be greater in the open compared to the closed economy, since $\mu_i < 1$. And, since $\gamma'(L_R) \leq 0$ and $M_i \Omega_i / (M_S \Omega_S + M_N \Omega_N) < 1$, the marginal benefit from strengthening the protection would be smaller in the open as compared to the closed economy. Thus, the marginal cost would exceed the marginal benefit in the open economy, which means that the government would have reason to reduce the index of patent protection from the autarky level.

for any CES research technology.¹⁵ Thus, both μ_i and γ are independent of the patent policies in the Cobb-Douglas case. It follows from (11) that the best response functions are linear and downward sloping in this case, and that the best response function for the South is steeper than that for the North, when the two are drawn in (Ω_S, Ω_N) space.

More generally, the best response functions need not be linear, but they must be downward sloping whenever $\beta \leq 0$; i.e., when the elasticity of substitution between human capital and labor in designing new products is less than or equal to one. Thus, the patent policies of the two countries are strategic substitutes in such circumstances. To understand the strategic interdependence between the governments in choosing their policies, consider the choice of patent protection by the South. Suppose the North were to strengthen its patent protection; i.e., to increase Ω_N . This would shrink the fraction of total discounted profits that an innovator earns in the South and so, *ceteris paribus*, reduce the responsiveness of global innovation to patent policy in the South. Moreover, the increase in Ω_N would draw labor into R&D in the North and South. If $\beta < 0$, the elasticity of innovation with respect to patent value would fall. The South would find that its market is relatively less important to potential innovators and that these innovators are less responsive to its patent policy. For both reasons, the marginal benefit to the South of strengthening its patent protection would fall and so the government would respond to the increase in Ω_N with a reduction in patent length or an easing of enforcement.

A situation of strategic complementarity (i.e., upward-sloping best response function) can arise only if the supply elasticity of R&D rises as the size of the research sector expands ($\beta > 0$) and then only if it rises sufficiently much to compensate for the decline in relative importance of a country's market that results when its trading partner strengthens its patent protection. It is straightforward to show that the two best response functions must slope in the same direction at any point of intersection. Thus, if the two patent policies are strategic complements in one country, they are strategic

¹⁵Note that

$$\phi_i = H_i \left[b \left(\frac{L_{Ri}}{a_i H_i} \right)^\beta + (1 - b) \right]^{\frac{1}{\beta}}.$$

From the fact that $vF_L(L_{Ri}/a_i, H_i) = w_i$, we have that $L_{Ri}/a_i H_i$ takes on a common value in the two countries; see footnote 13. It follows that ϕ_i is proportional to H_i , with the same factor of proportionality in both countries.

complements in both.

Returning to the case with $\beta \leq 0$, it is easy to show using (11) and $d\gamma/d\Omega_i \leq 0$ that the best response curve for the South must have a slope that is everywhere greater in absolute value than M_S/M_N , while the best response curve for the North must have a slope that is everywhere smaller in absolute value than M_S/M_N .¹⁶ It follows that the curve for the South must be steeper than that for the North at any point of intersection. This guarantees uniqueness of the Nash equilibrium and ensures stability of the policy setting game.

We summarize the most important findings in this section as follows.

Proposition 2 *Let the research technology be $\phi_i = \left[b[L_{Ri}/a_i]^\beta + (1 - b)H_i^\beta \right]^{1/\beta}$ in country i , for $i = S, N$. If $\beta \leq 0$, then the two patent policies are strategic substitutes in both countries and there exists a unique and stable Nash equilibrium of the policy setting game.*

III. Why is Patent Protection Stronger in the North?

Governments in the North typically provide stronger patent protection than their counterparts in the South.¹⁷ In this section, we identify sufficient conditions under which patent protection in the North will be stronger than that in the South in the Nash equilibrium of a noncooperative policy game. Our goal here is to understand the reasons why the North may have a greater incentive to protect IPRs than the South. We shall also examine how the equilibrium patent policies respond to changes in the endowments of human capital and to changes in the size of the market in each region.

We organize our discussion of the national differences in equilibrium policy choices around the following proposition.¹⁸

¹⁶We have not discussed the shape of the best response functions where they hit the axes or where the constraint that $\Omega_i \leq \bar{T}$ begins to bind. The best-response curve of the South becomes vertical if it hits the vertical axis at a point below $\Omega_N = \bar{T}$. It also becomes vertical if the South's best response is \bar{T} for some positive value of Ω_N . Similarly, the best-response curve for the North becomes horizontal if either it hits the horizontal axis before $\Omega_S = \bar{T}$ or if the North's best response is \bar{T} for some positive value of Ω_S . Thus, the best response curve for the South must be steeper than that for the North at any point of intersection, even if these additional segments of the best response functions are taken into account.

¹⁷See, for example, Juan C. Ginarte and Walter G. Park (1997) who have constructed an index of patent rights and have shown that this index is highly correlated with per capita GDP.

¹⁸Lai and Qiu (2003) provide conditions under which the Nash equilibrium patent duration is longer

Proposition 3 *Suppose $M_N > M_S$ and $H_N > H_S$. Then $\Omega_N \geq \Omega_S$ in any Nash equilibrium of the patent policy game. Moreover, $\Omega_N > \Omega_S$ unless $\Omega_S = \bar{T}$.*

The proposition is readily proved using the expressions for the best response functions in (11).¹⁹ First, recall that with a CES research technology, $\mu_i = H_i/(H_S + H_N)$. Thus, $H_N > H_S$ implies $\mu_N > \mu_S$. The left-hand side of (11) is a decreasing function of μ_i . If we cancel the terms that are common to the two best response functions, the remaining expression on the right-hand side is an increasing function of M_i and a decreasing function of Ω_i . It follows that if $\mu_N > \mu_S$ and $M_N > M_S$, a pair of policies can be mutual responses only if $\Omega_N > \Omega_S$.

Our answer to the question in the section heading is that the North has a larger market for innovative goods and a greater capacity to conduct R&D. Why do these characteristics induce the Northern government to provide stronger patent protection in a noncooperative equilibrium than its counterpart in the South? The reasons are somewhat subtle.

Recall from the discussion in Section 1 that having a large market is not per se a reason for a government to grant stronger patent protection. The optimal patent strength in a closed economy can in fact be independent of or even decreasing with market size, because both the marginal benefit of stronger protection and the marginal costs of the associated distortions are proportional to M for given γ , and the supply elasticity may remain the same or even decline as more resources are employed in R&D. The role of market size in generating different incentives for the governments has to do, instead, with the relative effectiveness of the countries' policy instruments. If M_N is larger than M_S , innovative firms earn a majority of their profits in the North. Then, a given change in Ω_N will generate a larger response of global innovation than would the same change in Ω_S . Since patents generate deadweight loss in the country that affords the protection, the country that can more effectively stimulate innovation with a given strengthening of its patent protection will have an incentive to provide stronger protection, all else equal.

in the North than in the South. Recall, however, that their model is considerably less general than ours. Moreover, their conditions include that the North invents more goods at every moment than the South, whereas our Proposition 3 refers only to the primitives of the model.

¹⁹There are some details involving corner solutions with $\Omega_S = 0$ or $\Omega_N = \bar{T}$ that we leave to the interested reader.

In our model, the endowment of human capital proxies for the capacity to conduct R&D. With $H_N > H_S$, a majority of the world's research is carried out in the North. As a consequence, a majority of the world's profits from innovative products accrue to residents of the North. In the North, the marginal cost of strengthening patent protection reflects the attendant loss in consumer surplus on all protected products less the profits that are captured by Northern producers. Similarly, the marginal cost of strengthening patents in the South reflects the loss of consumer surplus there less the profits captured by Southern producers. But since the Northern producers earn a majority of the profits, the offset to marginal cost is larger in the North than in the South. Accordingly, the government of the North has less of a temptation to ease patent protection than that of the South.

We turn next to the comparative static properties of the model. For this, we concentrate on the case in which the best response functions are downward sloping, which necessarily arises when (but not only when) $\beta \leq 0$.

Consider first the factor endowments. An equiproportionate change in H_S and H_N has no effect on μ_S or μ_N , and thus no effect on the best response functions or the Nash equilibrium. Policy outcomes change only when there is a change in the *relative* endowments of human capital in the two countries. Suppose H_S/H_N rises. This increases the share of innovation that occurs in the South (μ_S) and reduces the share in the North (μ_N). From (11) we see that the South's best response curve shifts to the right while the North's best response curve shifts downward. In Figure 1, we depict the original best response functions by SS and NN , respectively, and the shifted curves by $S'S'$ and $N'N'$. The equilibrium moves from E to E' , with a reduction in the strength of patent protection in the North and an increase in patent protection in the South. This result is consistent with the Ginarte and Park (1997) finding that patent rights are positively correlated in a cross-national sample with secondary school enrollment rates and with the share of R&D in GDP.

We turn to the effects of market size. If M_S and M_N grow equiproportionately, the term $M_i\Omega_i/(M_S\Omega_S + M_N\Omega_N)$ on the right-hand side of (11) is not affected at the initial values of Ω_S and Ω_N . Then, if $\beta = 0$ (Cobb-Douglas research technology), γ also is constant, and there is no effect on patent policy in either country. However, if $\beta < 0$, the extra resources that are drawn into R&D reduce the supply elasticity of innovation with respect to the value of a patent. Then the index of patent protection falls in both

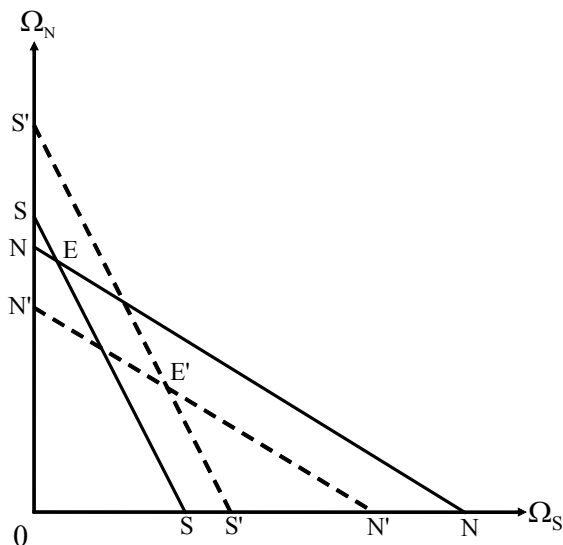


Figure 1: Comparative Static Effects of an Increase in H_S/H_N

countries. This is similar to our finding for a closed economy that, when $\beta < 0$, the optimal strength of IPR protection shrinks when the market for differentiated products expands.

Next consider an expansion in the size of the Southern market with no change in market size in the North. If $\beta = 0$, γ is constant, and an increase in M_S has qualitatively the same effects as an increase in μ_S ; these effects are shown in Figure 1, where we see that the strength of patent protection in the South grows while that in the North shrinks. However, if $\beta < 0$, the increase in M_S reduces γ at the initial values of Ω_S and Ω_N . Relative to the situation depicted in Figure 1, there is a further downward shift in NN and an offsetting leftward shift in SS . Indeed, if the supply elasticity of innovation falls by enough, the SS curve might even shift to its left relative to its initial location before the market expansion. In such circumstances Ω_S might fall as M_S grows.

IV. International Patent Agreements

In this section, we study international patent agreements.²⁰ We begin by character-

²⁰See also McCalman (2002), who discusses globally efficient patent policies in his two-country extension of the Nordhaus (1969) model. Lai and Qiu (2003) consider whether the joint welfare of the two countries would be increased if the South were to extend its patents so as to be equal in length to those chosen by the North in a Nash equilibrium.

izing the combinations of patent policies that are jointly efficient for the two countries.²¹ Then we compare the Nash equilibrium outcomes with the efficient policies, to identify changes in the patent regime that ought to be effected by an international treaty. Finally, we address the issue of policy harmonization. By that point, we will have seen that harmonization is neither necessary nor sufficient for global efficiency. We proceed to investigate the distributional properties of an agreement calling for harmonized patent policies and ask whether both countries would benefit from such an agreement in the absence of some form of direct compensation.

A. Efficient Patent Regimes

We shall begin by showing that the sum of the welfare levels of the two countries depends only on a measure Q of the *overall* protection afforded by the international patent system. This means that the same aggregate world welfare level can be achieved with different combinations of Ω_S and Ω_N that imply the same overall level of protection. One particular level of Q —call it Q^* —maximizes the sum of the countries' welfare levels. For a wide range of distributions of world welfare, efficiency is achieved by setting the individual patent policies so that the overall index of patent protection is Q^* .

In particular, let $Q = M_S\Omega_S + M_N\Omega_N$. This measure of global patent protection weighs the degree of patent protection in each country by the size of the country's market. A firm that earns a flow of expected profits of $\omega_S M_S \pi$ for a period of length τ_S in the South and a flow of expected profits of $\omega_N M_N \pi$ for a period of τ_N in the North earns a total discounted sum of expected profits equal to $Q\pi$. Thus, Q governs the allocation of resources to R&D in each country, regardless of the particular combination of patent policies in the separate countries.

Consider the choice of patent policies Ω_N and Ω_S that will take effect at time 0 and apply to goods invented thereafter. The expressions for the countries' gross welfare levels at time 0 are analogous to those for a closed economy, as recorded in equation (6). The

²¹Ours is a constrained efficiency, because we assume that innovation must be done privately and that patents are the only policies available to encourage R&D. We do not, for example, allow the governments to introduce R&D subsidies, which if feasible, might allow them to achieve a given rate of innovation with weaker patents and less deadweight loss.

aggregate welfare in country i , discounted to time 0, is given by

$$(12) \quad W_i(0) = \Lambda_{i0} + \frac{w_i(L_i - L_{Ri})}{\rho} + \frac{M_i(\phi_S + \phi_N)}{\rho} [\Omega_i C_m + (\bar{T} - \Omega_i) C_c] + \frac{\phi_i}{\rho} \pi (M_S \Omega_S + M_N \Omega_N), \text{ for } i = S, N,$$

where Λ_{i0} is the fixed amount of discounted surplus that consumers in country i derive from goods that were invented before time 0.

Summing the expressions in (12) for $i = S$ and $i = N$, we find that

$$(13) \quad \rho [W_S(0) + W_N(0)] = \rho (\Lambda_{S0} + \Lambda_{N0}) + w_S(L_S - L_{RS}) + w_N(L_N - L_{RN}) + (M_S + M_N) \bar{T} (\phi_S + \phi_N) C_c - Q (\phi_S + \phi_N) (C_c - C_m - \pi)$$

Since $v_S = v_N = \pi Q$, L_{RS} and L_{RN} are functions of Q .²² The same is true of ϕ_S and ϕ_N . It follows that different combinations of Ω_S and Ω_N that yield the same value of Q also yield the same level of aggregate world welfare.²³

If international transfer payments are feasible, then a globally efficient patent regime must have $M_S \Omega_S + M_N \Omega_N = Q^*$, where Q^* is the value of Q that maximizes the right-hand side of (13).²⁴ Notice that a range of efficient outcomes can be achieved without the need for any international transfers. By appropriate choice of Ω_N and Ω_S , the countries can be given any welfare levels on the efficiency frontier between that which they would achieve if $\Omega_S = 0$ and $\Omega_N = Q^*/M_N$ and that which they would achieve if $\Omega_S = Q^*/M_S$ and $\Omega_N = 0$.²⁵

²²In country i , the allocation of labor to research is determined by

$$\pi Q F_L(L_{Ri}/a_i, H_i) = 1/a_i.$$

²³This result is anticipated by a similar one in McCalman (2002), who studied efficient patent agreements in a partial equilibrium model of cost-reducing innovation by a single, global monopolist.

²⁴The first-order condition for maximizing $\rho[W_S(0) + W_N(0)]$ implies

$$C_c - C_m - \pi = \gamma \left\{ C_m + C_c \left[\frac{(M_S + M_N) \bar{T} - Q^*}{Q^*} \right] \right\}.$$

The second-order condition is satisfied at $Q = Q^*$ when $\beta \leq 1/2$.

²⁵This statement ignores the ceiling on patent lengths imposed by the finite economic life of differentiated products. A more precise statement is that a range of distributions of maximal world welfare can be achieved by varying Ω_S between $\Omega_S = \max\{0, (Q^* - M_N \bar{T})/M_S\}$ and $\min\{Q^*/M_S, \bar{T}\}$

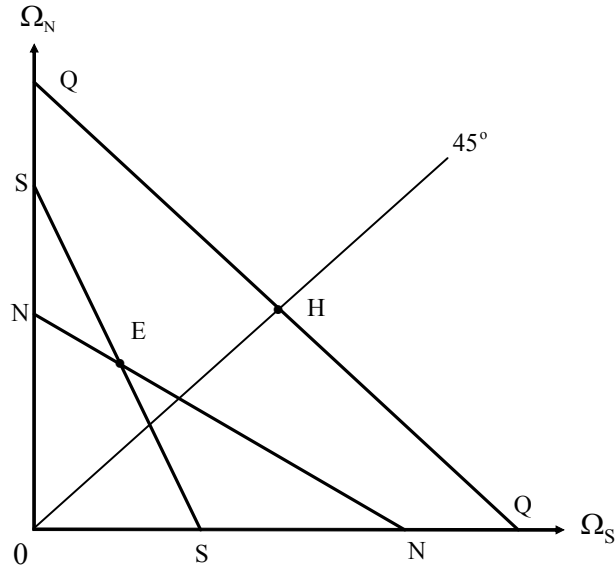


Figure 2: Comparison of Nash Equilibrium and an Efficient Patent Regime

Although aggregate world welfare does not vary with the national policies ω_i and τ_i as long as $M_S\Omega_S + M_N\Omega_N = Q^*$, the countries fare differently under the alternative combinations of policies that can be used to achieve global efficiency unless compensating transfers take place. In particular, the welfare of the North increases and that of the South decreases as Ω_S is increased and Ω_N is decreased in such a way as to keep the weighted sum constant. It follows that, absent any international transfer payments, the countries have a strong conflict of interest over the terms of an international patent agreement.

B. Pareto-Improving Patent Agreements

How do the efficient combinations of patent policies compare to the policies that emerge in a noncooperative equilibrium? The answer to this question — which informs us about the likely features of a negotiated patent agreement — is illustrated in Figure 2. The figure depicts the best response functions and the efficient policy combinations on the same diagram.

In the figure, the efficient policy combinations are depicted by the line QQ .²⁶ We while varying Ω_N between $\Omega_N = \min\{Q^*/M_N, \bar{T}\}$ and $\max\{0, (Q^* - M_S\bar{T})/M_N\}$ in such a way that $M_S\Omega_S + M_N\Omega_N = Q^*$.

²⁶If international transfer payments are infeasible, the set of Pareto efficient policy combinations

show this line being situated to the right of the SS curve and above the NN curve, which is a general feature of our model. The reasons are clear. Starting from a point on the South's best response function, a marginal strengthening of IPR protection in the South increases world welfare. Such a change in Southern policies has only a second-order effect on welfare in the South, but it conveys two positive externalities to the North. First, it provides extra monopoly profits to Northern innovators, which contributes to aggregate income there. Second, it enhances the incentives for R&D, inducing an increase in both ϕ_S and ϕ_N . The extra product diversity that results from this R&D creates additional surplus for Northern consumers.

By the same token, a marginal increase in the strength of Northern patent protection from a point along NN increases world welfare. Such a change in policy enhances profit income for Southern firms and encourages additional innovation in both countries. It follows, of course, that the QQ line must lie outside the Nash equilibrium. We record our finding in

Proposition 4 *Let (Ω_S, Ω_N) be an interior equilibrium in the noncooperative policy game and let (Ω_S^*, Ω_N^*) be any efficient combination of patent policies. Then $M_S \Omega_S^* + M_N \Omega_N^* > M_S \Omega_S + M_N \Omega_N$.*

The proposition implies that, starting from any interior Nash equilibrium, an efficient patent treaty must strengthen patent protection in at least one country. It also implies that the treaty will strengthen global incentives for R&D and induce more rapid innovation in both countries.

C. Harmonization

Commentators sometimes claim that it is desirable to have universal standards for intellectual property protection and for many other national policies that affect international competition. The arguments for harmonization are not always clear, but they seem to be based on a desire for global efficiency. Yet it is hardly obvious why efficiency should require identical policies in countries at different stages of economic development. In this section, we examine the aggregate and distributional effects of international harmonization of patent policies. It should be noted that TRIPs essentially harmonizes

includes the segment of the vertical axis above its intersection with QQ and extending as far as the point $(0, \bar{T})$ and the segment of the horizontal axis to the right of its intersection with QQ and extending to $(\bar{T}, 0)$.

patent policy by requiring all governments to grant patents of at least twenty years from date of filing and by requiring minimum standards of effective enforcement.

As should be apparent from the preceding discussion, harmonization of patent policies is neither necessary nor sufficient for global efficiency, regardless of whether international transfer payments are feasible or not. A regime of harmonized policies will only be efficient if the common index of patent protection in the two countries is such that $Q = Q^*$. And any combination of patent policies that provides the proper global incentives for R&D will be efficient, no matter whether the patent policies in the two countries are the same or not.

If patent protection is stronger in the North than in the South in an initial Nash equilibrium, then harmonization might be achieved either by a unilateral strengthening of patent laws in the South or by a combination of policy changes in the two countries. A unilateral increase in Ω_S is bound to harm the South (absent any side payments), because the equilibrium policy package is a best response by the South to the North's choice of patent laws and any unilateral deviation from a country's best response is, by definition, damaging to its interests.²⁷ As for harmonization that might be achieved through a combination of policy changes, we focus on a treaty that would achieve global efficiency. Such a treaty is represented by point H in Figure 2. Efficient harmonization surely requires a strengthening of patent protection in the South, since $\Omega_N > \Omega_S$ at E and QQ lies outside this point. If $\beta \leq 0$, it also requires a strengthening of patent protection in the North.²⁸ If $M_N \geq M_S$ and $H_N \geq H_S$, the North definitely gains from efficient harmonization.²⁹ However, the South may be worse off at point H than in

²⁷See also Lai and Qiu (2003), who consider the welfare effects of harmonizing IPR protection at the standard that would be chosen by the North in a non-cooperative equilibrium. In a model of once-off investment in R&D, they show that such a change in the South's policy from the Nash equilibrium level would benefit the North by more than it would harm the South.

²⁸First, we note that when $\beta \leq 0$, point H lies above the intersection of the NN curve with the vertical axis. This can be seen by substituting $\Omega_N = \Omega_S$ in the first-order condition for maximizing $\rho [W_S(0) + W_N(0)]$ and comparing the resulting expression for $\Omega_N = Q^*/(M_N + M_S)$ with the expression for Ω_N that comes from (11) when $\Omega_S = 0$. Then, since the NN curve is downward sloping when $\beta \leq 0$, the fact that it starts below point H implies that the extent of patent protection in the North is greater at point H than it is at point E .

²⁹If $M_N \geq M_S$ and $H_N \geq H_S$, the common index of patent policy that maximizes the welfare of the North is greater than the common index of patent policy that maximizes aggregate world welfare. Therefore, the North gains from a unilateral increase in Ω_S that brings the Southern patent policy into

the Nash equilibrium at point E , unless some form of compensation is provided by the North. In general, the larger are M_N/M_S and H_N/H_S , the more likely it is that the South would lose from efficient harmonization.

Summarizing, we have

Proposition 5 *Suppose $M_N \geq M_S$, $H_N \geq H_S$, and $\beta \leq 0$. Then efficient harmonization requires a strengthening of patent protection in both countries. The North necessarily gains from efficient harmonization, while the South may gain or lose.*

We conclude that harmonization has more to do with distribution than with efficiency, and that incorporation of such provisions in a treaty like TRIPs might well benefit the North at the expense of the South.³⁰

V. Patent Policy with Many Countries

In this section, we extend our analysis to a trading world with many countries. Our main finding is that adding countries exacerbates the free-rider problem that plagues the noncooperative policy equilibrium. Small countries are inclined to allow others to provide the incentives for innovation so as to avoid the deadweight losses in their home markets. In the limit, as the number of countries grows large and each one is small in relation to the world economy, the unique Nash equilibrium has universal patents of strength zero. Then, a patent treaty is critical for creating incentives for private innovation.

We assume that there are J countries, and that country i has market size M_i , human capital endowment H_i , and labor productivity $1/a_i$. The research technology in country i is $\phi_i = F(H_i, L_{Ri}/a_i) = [b[L_{Ri}/a_i]^\beta + (1-b)H_i^\beta]^{1/\beta}$, with $\beta \leq 1/2$. All consumers share the preferences given in (8).

Suppose that there is no cooperation between nations in setting their patent policies. In country i , either $\Omega_i = 0$ and the marginal cost of providing the first bit of patent protection exceeds the marginal benefit, $\Omega_i = \bar{T}$ and the marginal benefit of providing conformity with the Nash equilibrium policy in the North, and further gains from an increase in the common policy until $Q = Q^*$.

³⁰McCalman (2001) estimates the income transfers implicit in TRIPs and finds that international patent harmonization benefits the United States at the expense of the developing countries as well as Canada, the United Kingdom and Japan. He does not, however, include in his calculations the benefits that countries derive from the global increase in innovative activities.

the last bit of patent protection exceeds the marginal cost, or $0 < \Omega_i < \bar{T}$ and the marginal benefit of strengthening patent protection equals the marginal cost. Equality between marginal benefit and marginal cost implies

$$(14) \quad C_c - C_m - \mu_i \pi = \frac{M_i}{Q} \gamma [\Omega_i C_m + C_c (\bar{T} - \Omega_i)] ,$$

where $Q = \sum_j M_j \Omega_j$ measures the strength of global patent protection in the Nash equilibrium.

Observe first that as $\mu_i \rightarrow 0$, the left-hand side of (14) approaches $C_c - C_m$; a small country captures virtually none of the monopoly profits from innovative products, so the marginal cost of a patent per consumer and product is the difference between the competitive and monopoly levels of consumer surplus. But as $M_i \rightarrow 0$, the right-hand side of (14) approaches zero, because a small country provides innovators with virtually none of their global profits and so worldwide innovation is hardly responsive to a change in such a country's patent policy. It follows that a small country will set its index of patent protection equal to zero in a Nash equilibrium.

If all countries choose positive patent strengths that are less than \bar{T} , equation (14) holds for every i . Then we can sum (14) across the J countries, which gives

$$(15) \quad J(C_c - C_m) - \pi = \gamma \left[C_m - C_c + \frac{C_c \left(\sum_j M_j \right) \bar{T}}{Q} \right] .$$

Then, for a given size of the world market, Q depends only on the *number* of countries J and not on the distribution of consumers and human capital across countries. Moreover, if $\beta \leq 0$, Q is a declining function of J ; the greater is the number of countries, the weaker are the global incentives for innovation in a noncooperative equilibrium. As the number of countries grows large (holding constant the size of the world market), the aggregate incentives for innovation approach zero.³¹ Evidently, the free-rider problem becomes increasingly severe as the number of independent decision makers in the world economy expands.

Finally, note that the requirements for global efficiency do not depend on the number of countries. Again, the sum of all national welfare levels is a function of the aggregate

³¹Suppose Q were to approach a finite number as $J \rightarrow \infty$. Then γ would approach a finite number as well, and the right-hand side of (15) would be finite. But the left-hand side of (15) approaches infinity as $J \rightarrow \infty$.

world incentive for innovation. This sum is maximized when

$$(16) \quad C_c - C_m - \pi = \gamma \left[C_m - C_c + \frac{C_c \left(\sum_j M_j \right) \bar{T}}{Q^*} \right].$$

Thus, if international compensation is possible, an efficient global patent treaty will have $\sum_j M_j \Omega_j = Q^*$, where Q^* is solved from (16). Notice that Q^* must exceed Q , the aggregate patent protection in the Nash equilibrium. Even if international compensation is not feasible, an efficient agreement will have $\sum_j M_j \Omega_j = Q^*$ for a range of distributions of world welfare.

VI. Conclusions

We have developed a simple model of endogenous innovation and have used it to study the incentives that governments face in choosing their patent policies. Our model features a familiar trade-off between the static benefits of competitive pricing and the dynamic benefits of increased innovation. For a closed economy, we derived a simple formula for the optimal strength of patent protection that relates the deadweight loss induced by a marginal strengthening of IPR protection to the surplus that results from the extra innovation.

In an open economy, differences in market size and differences in capacity for R&D generate national differences in optimal patent policies. We focused on policies that are applied with *national treatment*; that is, regimes that require equal protection for foreign and domestic applicants. A country's optimal patent policy is found by equating the sum of the extra deadweight loss that results from strengthening the IPR protection granted to domestic firms and the extra consumer surplus loss that results from expanding the fraction of imported goods that are subject to monopoly pricing with the benefits that flow from providing greater incentives for innovation to firms worldwide. A country's optimal IPR protection depends on the policies set by its trading partner, because the strength of foreign patent rights affects the responsiveness of global innovation to a change in a country's own patent policies.

We found that having a larger market for innovative products typically enhances a government's incentive to grant stronger patent rights. Also, a government's relative incentive to protect IPRs typically increases with its relative endowment of human capital. In a noncooperative equilibrium, patent protection will be stronger in the North than in the South if the North has a larger market for innovative products and a greater capacity for R&D.

Starting from a Nash equilibrium, countries can benefit from negotiating an international patent agreement. A treaty can ensure that national policies reflect the positive externalities that flow to foreign residents when a country tightens its patent laws. To achieve (constrained) efficiency, an international agreement must strengthen aggregate world patent protection relative to the Nash equilibrium. Harmonization of patent policies is neither necessary nor sufficient for the efficiency of the global IPR regime. If patent policies are harmonized at an efficient level, the move from a Nash equilibrium typically will benefit the North but possibly harm the South.

Our conclusions are essentially the same for a world with more than two countries. Countries with larger markets and more human capital will provide stronger IPRs in a noncooperative equilibrium than those with smaller markets and less human capital. Indeed, a country that is small in relation to the world economy has no incentive whatsoever to grant patents. The greater is the number of independent countries, the more severe is the free-rider problem inherent in the setting of national patent policies. Thus, the value of an international patent agreement grows with the number of independent sovereign decision makers.

Our analysis can be extended to more general environments. For example, in an earlier version of this paper (Grossman and Lai, 2002), we allowed for cross-national differences in relative labor productivity in the two industries. With comparative advantage in production, the productivity gap in the industry that produces differentiated products may not be offset by the gap in relative wages. Then the production costs for innovative products will be higher in one region or the other. This can create an asymmetry in the life cycle of a new good depending upon whether patents are longer in the North or in the South. We showed how such an asymmetry may generate multiple equilibria in the policy game.

Another possible extension would allow different preferences in different countries. With different demands, the marginal cost of strengthening IPR protection will vary around the globe. Then differences in the elasticities of demand for innovative products will be another factor that affects the governments' relative incentives for granting long patents or providing strict enforcement. Moreover, asymmetries in demand would be reflected in the characteristics of a globally efficient IPR regime. An efficient regime would equalize across countries the marginal deadweight loss associated with providing a given push to global innovation. Efficiency requires stronger patent protection in countries that have more inelastic demands for innovative products, all else the same.

Appendix

In this appendix we show that, for a closed economy, when Ω solves (7) and $\beta \leq 1/2$, the second-order condition for an optimal patent policy is satisfied. Similar calculations ensure that $\beta \leq 1/2$ is sufficient for the second-order condition to be satisfied for the best response given by (11) for an open economy.

Let us rewrite the first-order condition (7) as

$$\frac{\gamma}{\Omega} \left\{ C_c \bar{T} - \Omega \left[(C_c - C_m) + (C_c - C_m - \pi) \frac{1}{\gamma} \right] \right\} = 0.$$

Since $\gamma/\Omega > 0$, the term in curly brackets must vanish at any local extremum. We will show that at any such point the term in curly brackets is a decreasing function of Ω ; i.e., that

$$(A1) \quad - \left[(C_c - C_m) + (C_c - C_m - \pi) \frac{1}{\gamma} - \Omega \frac{(C_c - C_m - \pi) d\gamma(\Omega)}{\gamma^2 d\Omega} \right] < 0.$$

This means that any point satisfying the first-order condition is a local welfare maximum. Since the welfare function is continuous and differentiable, it follows that there can be at most one local extremum point, and that the value of Ω that generates this point is the unique welfare-maximizing patent policy.

It is straightforward to calculate that

$$\gamma \equiv - \left[\frac{(F_L)^2}{F F_{LL}} \right] = \frac{b}{(1-b)(1-\beta)} \left(\frac{L_R}{aH} \right)^\beta$$

for the CES research technology. Meanwhile, $vF_L = w = 1/a$ and $v = M\pi\Omega$ imply that $F_L = 1/aM\pi\Omega$, or that

$$\frac{b}{a} \left[b + (1-b) \left(\frac{aH}{L_R} \right)^\beta \right]^{\frac{1-\beta}{\beta}} = \frac{1}{aM\pi\Omega}$$

in the CES case. Using these two equations, we can express γ as a function of Ω ; we find that $\gamma = \left(\frac{b}{1-\beta} \right) \left[(bM\pi\Omega)^{\frac{-\beta}{1-\beta}} - b \right]^{-1}$.

We can now compute $d\gamma(\Omega)/d\Omega$, and substitute the resulting expression into the left-hand side of (A1), which then becomes

$$\Psi = - \left[C_c - C_m - (1-\beta)(C_c - C_m - \pi) + (C_c - C_m - \pi) (bM\pi\Omega)^{\frac{-\beta}{1-\beta}} \left(\frac{1-2\beta}{b} \right) \right].$$

But $\beta \leq 1/2$ ensures that

$$\begin{aligned}\Psi &\leq -[C_c - C_m - (1 - \beta)(C_c - C_m - \pi) + (1 - 2\beta)(C_c - C_m - \pi)] \\ &= -[C_c - C_m - \beta(C_c - C_m - \pi)] \\ &< 0.\end{aligned}$$

So the second-order condition is satisfied when $\beta \leq 1/2$ and the Ω that solves (7) – if it exists – is the unique welfare-maximizing patent policy.

References

- [1] Chin, Judith and Grossman, Gene M. "Intellectual Property Rights and North-South Trade," in R.W. Jones and A.O. Krueger, eds., *The Political Economy of International Trade*, Cambridge MA: Basil Blackwell Publishers, 1990, pp. 90-107.
- [2] Deardorff, Alan V. "Welfare Effects of Global Patent Protection." *Economica*, February 1992, 59 (issue 233), pp. 35-51.
- [3] Ginarte, Juan C. and Park, Walter G. "Determinants of Patent Rights: A Cross-National Study." *Research Policy*, October 1997, 26(3), pp. 283-301.
- [4] Grossman, Gene M. and Lai, Edwin L.-C. "International Protection of Intellectual Property." National Bureau of Economic Research Working Paper No. 8794, Cambridge MA, January 2002.
- [5] Helpman, Elhanan "Innovation, Imitation, and Intellectual Property Rights." *Econometrica*, November 1993, 61(6), pp. 1247-1280.
- [6] Lai, Edwin L.-C. and Qiu, Larry D. "The North's Intellectual Property Rights Standard for the South?" *Journal of International Economics*, January 2003, 59(1), pp. 183-209.
- [7] Maskus, Keith E. *Intellectual Property Rights in the Global Economy*, Washington D.C.: The Institute for International Economics, 2000a.
- [8] Maskus, Keith E. "Parallel Imports." *The World Economy*, September 2000b, 23(9), pp. 1269-1284
- [9] McCalman, Phillip "Reaping What you Sow: An Empirical Analysis of International Patent Harmonization." *Journal of International Economics*, October 2001, 55(1), pp. 161-185.
- [10] McCalman, Phillip "National Patents, Innovation and International Agreements." *Journal of International Trade and Economic Development*, March 2002, 11(1), pp. 1-14.
- [11] Nordhaus., William D. *Invention, Growth and Welfare: A Theoretical Treatment of Technological Change*. Cambridge MA: The M.I.T. Press, 1969.

- [12] Scotchmer, Suzanne “The Political Economy of Intellectual Property Treaties.”
Journal of Law, Economics, and Organization, 2004, 20 (2), pp.415-437.