Arbitrage and the Harmonization of Patent Lives

Eric W. Bond and Yih-Ming Lin*

June, 2005

Abstract

We study a North/South model of intellectual property rights protection in which markets are not completely segmented. We examine how the existence of arbitrage affects the incentives of countries to set the length of patent protection. The results show that the North will not have an incentive to completely eliminate arbitrage after patents expire in the South, even when enforcement is costless. The results also show that, if the demand function for an innovation is linear, there exists a pure Nash equilibrium of a non-cooperative game between the North and the South. Furthermore, we demonstrate that the uniform universal standard for IPRs protection will never achieve the global Pareto efficiency when the markets are not perfectly segmented.

JEL Classification: F13; K42; O34

Keywords: Patents; Intellectual Property Rights (IPRs); Arbitrage; Enforcement

*Department of Economics, Vanderbilt University. PO Box 35319, Station B 2301 Vanderbilt Place Nashville, TN, 37235, USA.

1 Introduction

The primary purpose of intellectual property rights (hereinafter IPRs) protection is to provide innovators incentives to develop new technologies, products, and services. IPRs grant innovators monopoly power to make profits on their innovations, which also generates additional consumer surplus. The cost of providing IPRs protection is that it permits innovators to exercise monopoly power over the market, which prevents the benefits of the new products from being enjoyed optimally by consumers. There exists a trade-off between static deadweight loss and dynamic gains from innovation. Therefore, in a closed economy, the existing literature(e.g. Nordhaus (1969), Scherer (1980), and Deardorff (1992)) suggests that IPRs should be granted only for a limited period.

Recently, IPRs protection has gained importance in international trade. Maskus (2000) indicates that IPRs protection has been at the forefront of global economic policy-making, and a number of countries have strengthened their laws and regulations regarding IPRs protection. Especially, the agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPs) was concluded as a part of the foundation of the World Trade Organization (WTO).

Studying IPRs protection in an open economy introduces additional complications, since there are two types of externality involved. One is the terms of trade externality because, for example, it worsens the home country terms of trade on patented products if IPRs protection is granted in the foreign country. The foreign consumer surplus would be diminished by the exercise of monopoly power. The other is the property of public goods since the foreign country benefits from the creation of consumer surplus by the introduction of new innovations in the home country. Thus, an increase in domestic patent length will encourage domestic innovators as well as foreign innovators, but it only negatively affects the domestic consumers' surplus. Therefore, each country hopes that its trading partners can strengthen their level of protection and it can lower its own level of protection, which creates a prisoner's-dilemma-like situation.

Lai and Qiu (2003) utilize a game-theoretical approach to analyze international IPRs protection in a multi-sectoral two-country model . They show that, in Nash equilibrium, developed countries (North) have a stronger incentive to protect IPRs than developing countries (South), which explains why the North chooses a longer IPRs standard than the South. They also find that global welfare increases if developing countries increase their IPRs standards. Grossman and Lai (2004) propose an infinite horizon general equilibrium model in which each firm innovates in every period. They examine the IPRs policy in two countries that differ in market size, capacity, and wage rate. Furthermore, their model demonstrates that harmonization in patent length is not necessary for efficiency. Readjustments of patent protection across countries that leave global profits unaffected will not affect world welfare. In their model, the issue of whether or not to harmonize primarily affects the distribution of the income between the countries.

The existing literature suggests that the South has an incentive to provide IPRs protection whether or not southern firms have capacities to innovate. However, the protection in the South is weaker than in the North. The length of patent protection in developed countries is longer than in developing countries because they have bigger markets and greater capacities for innovation. An important feature of these models is the assumption that markets can be perfectly segmented, so that IPRs protection of different lengths can be maintained without the interference of arbitrage between markets.

In this paper, the assumption of perfect segmentation is relaxed. The existence of arbitrage between international markets is hardly new, as evidenced by the importation of Canadian drugs. The Minnesota Senior Federation, according to Weil (2004), organized its first prescription-drug-buying bus trip to Canada in 1995. The bus picked up seniors in Minneapolis and drove them to Winnipeg Canada, where they could purchase their prescription drugs at reduced prices. The bus still runs eight or nine times a year. Currently, the Federation also runs two Internet pharmacy stores, from which more than 4,000 customers ordered more than two million dollars worth of drugs last year. Weil explains that the Federation's activity has spurned the implementation of similar bus trips in California, Arizona, Maine and other border states.

Some Canadian pharmacies have also been soliciting business in the US, where drug prices are substantially higher. The internet, which is easily accessible, has encouraged this kind of arbitrage behavior. The Wall Street Journal (Spors, 2004) reports that Americans bought approximately $1.1 billion of Canadian drugs last year. Furthermore, some state governments also encourage this kind of arbitrage. For example, Minnesota, New Hampshire, North Dakota and Wisconsin have all constructed Web sites that instruct their residents how to buy relatively cheap prescription drugs from Canadian pharmacies online. (State Legislatures, 2004).

Another example is the trade in illegal compact-discs, videotapes and video games in China prior to the TRIPs agreement. The Economist (1996) reports that China produced

54 million units of CDs in 1994, of which 88 percent of them were pirated. The Chinese bought under 40 million units themselves and the excess was exported. The United States bought 13% of pirated sales by value. Americans bought more pirated music than any other country except Russia.

In the following, we examine how the existence of arbitrage affects the incentives for countries in the setting of patent protections. We first examine the Nash equilibrium of the patent game between countries when the patent agreement does not exist. We then show that the innovating countries are not willing to completely eliminate arbitrage when the patent expires first in non-innovating countries but is still effective in innovating countries, even when enforcement is costless. We also show that the existence of arbitrage makes it more attractive for non-innovating countries to strengthen their IPRs protection. Finally, we investigate the efficient international patent agreement between countries. We show that the existence of arbitrage will result in efficient patent agreements in which differences in patent levels across countries reduce the deadweight loss within the patent system.

The remainder of this paper is organized as follows: In Section 2, we develop a two-country model of intellectual property rights protection in which the markets are not completely segmented. In Section 3, we show how arbitrage affects international IPRs protection and that there exists a pure-strategy Nash equilibrium when arbitrage occurs after the patent protection first expires in the non-innovating country. The efficient international patent agreement is examined in Section 4. Section 5 gives conclusions.

2 The Model

In this section, we set up a two-country model of IPRs protection in which the markets are not completely segmented. Consider an economy with two countries, named South (s) and North (n), and a variety of firms that develop new products in the North.[1] The model follows Deardorff (1992) and Scotchmer (2004) by considering a two-period model of innovation. It is assumed that there are two sectors, a homogenous good sector and a differentiated products sector. In the first period, firms choose the number of differentiated products to be produced and sold in the two markets in the second period. Individuals are assumed to have identical preferences in each country. Each consumer chooses z_t and $x_t(\omega)$ to maximize his utility which is given by:

$$\int_0^T [\int_0^N u(x_t(\omega))d\omega + z_t]e^{-\rho t}dt,$$

$$subject \;\; to \quad \int_0^N p_t(\omega)x_t(\omega)d\omega + z_t = Y_t, \quad for \;\; all \;\; t,$$

where T is the length of the production period, ω is the index of differentiated goods, N is the measure of new differentiated products in the North, $0 < \rho < 1$ is a discount factor, $x_t(\omega)$ is the consumption of a differentiated good ω at time t, z_t is the consumption of the homogenous good, and Y_t is the current income. We assume that $u' > 0$ and $u'' < 0$. The first order condition yields $x_t(\omega) = x(p_t(\omega))$ where $x = (u')^{-1}$.[2] The demand for the

[1] This assumption is not so restrictive. Braga (1990) indicates that developing countries only hold 1% of existing patents. Chin and Grossman (1990), Diwan, and Rodrick (1991) and Deardorff (1992) make the same assumption in their model. Furthermore, the main results in Grossman and Lai (2004) will not change if their assumption is simplified.

[2] A utility function of this kind greatly simplifies the analysis. Krugman (1992) argues in favor of using this kind of utility function for analyzing political economy issues in a multisectoral general equilibrium

numeraire good is given by:

$$z_t = Y_t - p_t(\omega)x_t(p_t(\omega)).$$

The indirect utility function can be written as:

$$U = \int_0^T \int_0^N s(p_t(\omega))d\omega e^{-\rho t}dt + Y,$$

where $s(p_t(\omega)) = u(x_t(p_t(\omega))) - p_t(\omega)x_t(p_t(\omega))$ is the consumer surplus associated with a representative differentiated product, Y is the present value of income during the production period. The two-period model can be thought of as representing the steady state of an infinite horizon general equilibrium model in which firms innovate in every period and products have an exogenously given useful life, as has been shown by Grossman and Lai (2004).

Suppose that manufacturing requires only labor. Since the homogenous good z is considered as the numeraire, without loss of generality, we assume that the production of the homogeneous good z requires one unit of labor per unit of output and the market for z is competitive, which makes the price of z equal to one. Furthermore, suppose that the production of one unit of all varieties of differentiated products requires c units of labor. Under perfect segmentation, the innovating firms are granted monopoly power over its goods when the patent is still in effect.

framework. He indicates that, in such a framework, partial equilibrium intuition continues to apply in a general equilibrium model. Currently, this is the standard model employed in the political economy literature on trade policy, such as Grossman and Helpman (1994), Mitra (1999), Grossman and Lai (2004).

In this model, it is assumed that markets are imperfectly segmented, so an arbitrageur has incentives to make a profit by exporting differentiated products from the low-price market to the high-price market. Potential arbitrage profits arise from two possibilities. The first is that the manufacturers are price discriminating across markets when patents are still in effect in both markets. This occurs when one manufacturer owns patents or national trademarks in several markets. The producer would rationally set low prices in low-income countries, or in the markets with elastic demand. In contrast, high-income countries, or inelastic-demand markets will face a higher price. The purpose of differential pricing is in order to make a higher profit for the manufacturer. However, since price differences exist between countries, there are incentives for one to purchase goods in a low-price country and resell them in a high-price country if transaction costs are low.

The second possibility occurs when the length of patent protection differs across markets. After the patent expires in one market, imitators can freely produce the products at a constant cost c. The market becomes competitive and the price drops to its marginal production cost, c, so that the manufacturer can not make any profit. Since we focus on the impact of different patent lives across markets, we assume that the demand for each innovation is the same for consumers in the two markets, in the interest of simplicity. The only difference between the two markets is the market size, $x_s(p) = Mx_n(p)$ and $s_s(p) = Ms_n(p)$, where M denotes the relative market size of the southern market.[3] Furthermore, the profit function is assumed to be strictly concave on $[c, \bar{p}]$, where $\pi(c) = \pi(\bar{p}) = 0$.

[3]This assumption follows Grossman and Lai (2004).

Let λ denote the probability that goods which are legally protected from arbitrage are confiscated by the government. This represents the level of enforcement against illegal arbitrage. Therefore, the arbitrageur's expected revenue is $p(1-\lambda)$ for shipping a unit of differentiated product from the South where its unit cost is c. Since arbitrage is subject to free entry, the zero profit condition will force the market price equal to $c/(1-\lambda)$. Therefore, the patent holder will set its price a little lower than $c/(1-\lambda)$ to get the whole market. Thus, if $\lambda = 0$, the market price p will be equal to c. We also can get $p = p_m$, the monopoly price, for $\lambda \geq \bar{\lambda}$ with $p_m = c/(1-\bar{\lambda})$. The market price falls and the volume of consumption increases when λ shrinks. On the other hand, consumers' surplus will increase. Furthermore, it is easy to show that, for $\lambda \in [0, \bar{\lambda}]$, there exists a one-to-one relationship between λ and the market price, $p(\lambda)$, with $p'(\lambda) > 0$.

In order to construct the payoff function for the Northern innovators, we assume that the patent expires first in the South, i.e., $t_n \geq t_s$ where t_i is the length of patent protection in country i.[4] The conditions under which this will hold in the Nash equilibrium will be discussed below. Thus, we can get that $p_t^n(\omega) = p_t^s(\omega) = p_m$ for $t \leq t_s$ and $p_t^n(\omega) = p_t^s(\omega) = c$ for $t > t_n$. For $t \in (t_s, t_n]$, $p_t^s(\omega) = c$ and $p_t^n(\omega) = p(\lambda)$ =min $\{c/(1-\lambda), p_m\}$. The average profits of an innovation during the production period can be written as:

$$\Pi(\tau_s, \tau_n, \lambda, N) = (1+M)\tau_s \pi(p_m) + (\tau_n - \tau_s)\pi(p(\lambda)), \tag{1}$$

where $\tau_n = \frac{1-e^{-\rho t_n}}{1-e^{-\rho T}}$ and $\tau_s = \frac{1-e^{-\rho t_s}}{1-e^{-\rho T}}$ are the shares of the production period in which the product is under IPRs protection in the North and South, respectively. The expression

[4] We will discuss the case of $t_n < t_s$ in the Appendix.

$\pi(p_m)$ is the (monopoly) profit of a new product per period in the North, while $\pi(p(\lambda))$ is the firm profit when the Northern government chooses λ as the enforcement level to prevent arbitrage after the patent expires in the South.

The demand function and the profit for a particular enforcement level for a representative innovation are illustrated in Figure 1. In Figure 1, $\Delta(\lambda)$ is the social deadweight loss and $s(\lambda)$ is the consumer's surplus. Since $p'(.)$ is positive, we can see that a decrease in λ will increase the consumer's surplus and reduce social deadweight loss. The benefit of an innovation is the monopolistic profits in the two countries during the life of the patent plus the profit in the North after the patent expires in the South.

According to the definition of τ_i, there exists a one-to-one relationship between t_i and τ_i. For the sake of simplicity, we designate τ_i as the length of patent protection in the following model. The goal of the government is to maximize the welfare of its country by choosing its policy tools, the length of the patent(τ_i) and λ. The welfare of country i is the sum of firms' profits in the two countries and consumers' surplus minus the cost of innovation and enforcement. Therefore, the Northern government's objective function is:

$$W_n(\tau_s, \tau_n, \lambda, N) = (\tau_s(1+M)\pi(p_m) + (\tau_n - \tau_s)\pi(p(\lambda)))N \tag{2}$$

$$+(\tau_s s(p_m) + (1-\tau_n)s(c) + (\tau_n - \tau_s)s(p(\lambda)))N - C(N) - G(\lambda).$$

The first part of the right hand side (hereinafter RHS) of equation (2) is the manufacturer's profit in both markets during the time when the patent has not expired. The profit from the Southern market is $\tau_s M\pi(p_m)$. While the Southern patent has not expired, firms can make a monopoly profit π in the North. When the Southern patent expires, but the Northern patent remains in effect, firms can only make $\pi(p(\lambda))$. Therefore, producers can

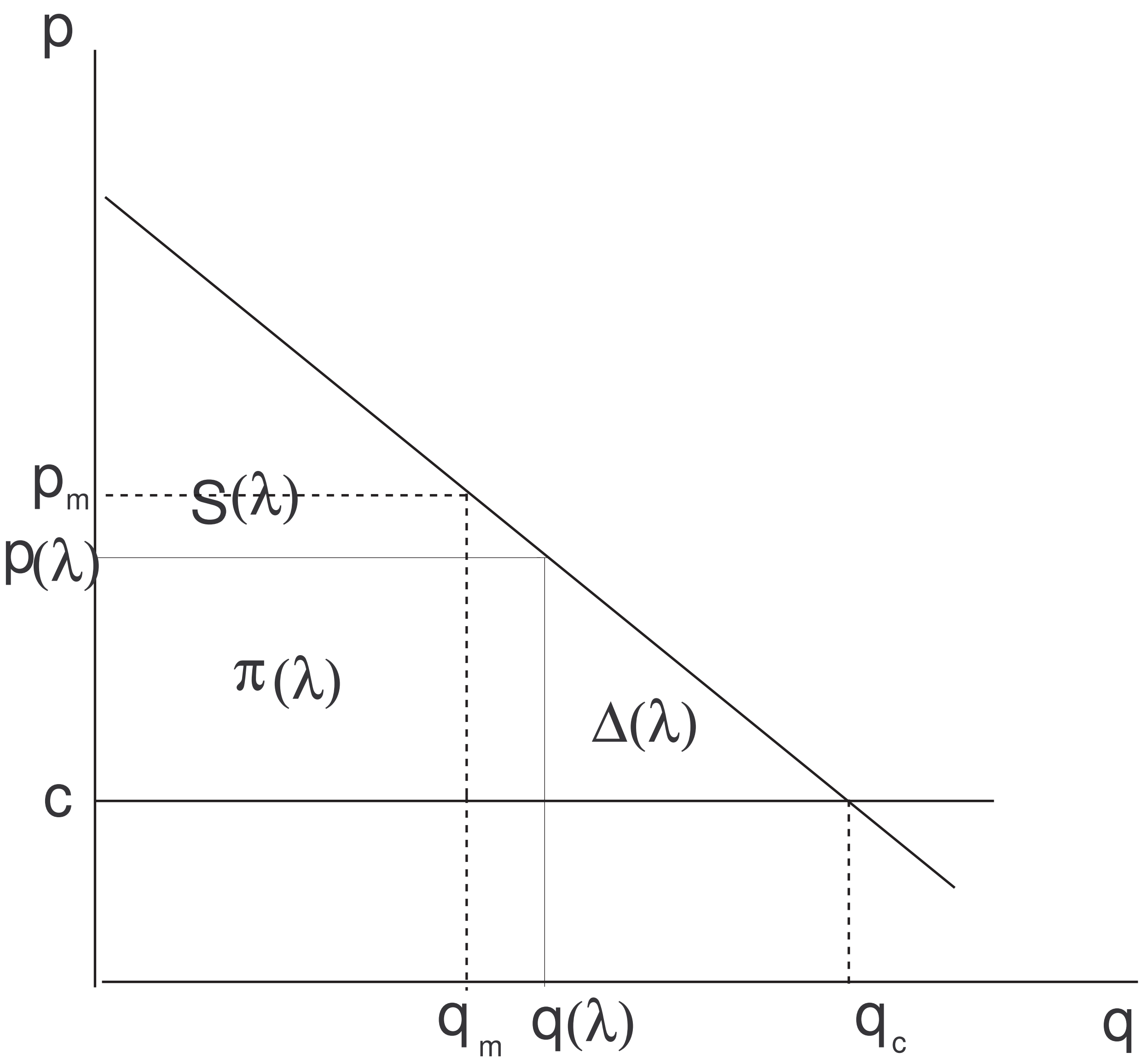

Figure 1: Profit, $\pi(\lambda)$, consumer's surplus, $s(\lambda)$ and social deadweight loss, $\Delta(\lambda)$.

make $[\tau_s \pi(p_m) + (\tau_n - \tau_s)\pi(p(\lambda))]$ from the Northern market.

The second part of the North's welfare is the consumers' surplus. The Northern consumers' surplus is $s(p_m)$ before the Southern patent expires, $s(c)$ after the Northern patent expires, and $s(p(\lambda))$ when the patent expires in the South but remains in effect in the North, with an enforcement level, λ.

The third part of RHS of equation (2) is the innovation cost. $C(N)$ denotes the cost of innovation required to generate N innovations in the North. Suppose that $C(.)$ is strictly convex in N. This model follows Grossman and Lai (2004), and considers that increasing marginal costs of producing innovations can arise from the existence of a sector-specific human capital input employed in the R&D sector along with mobile labor.

The final part, $G(\lambda)$, is the government cost of resources devoted to preventing arbitrage. $G'(.)$ is supposed to be positive which implies that the enforcement cost is increasing in enforcement level, λ.

Similarly, we also can obtain the welfare for the South:

$$W_s(\tau_s, N) = M(\tau_s s(p_m) + (1 - \tau_s)s(c))N. \tag{3}$$

Equation (3) indicates that the welfare of the non-innovating country only comes from its consumers' surplus, as the South's buyers consume the differentiated goods invented in the North.

3 The Non-Cooperative Equilibrium

In this section, we analyze the Nash equilibrium of this patent-length setting game. We assume that the sequence of decision is as follows: First, based on the welfare functions, governments simultaneously set their patent length, τ_i, and the level of enforcement, λ.[5] Then, the innovators in the North decide how many innovations they plan to invent and sell their products in the two markets. Since we are interested in the interaction between the two governments, the equilibrium concept employed here is Nash equilibrium.

In stage 2, whether or not a firm decides to develop an innovation is determined by the marginal profit and the marginal cost. Firms will invest in R&D up to the point where the cost of introducing an additional innovation is equal to the benefit from the innovation, $C'(N^*) = \Pi$. Therefore, the optimal number of innovations can be expressed as a function of the enforcement level and the two patent lengths, $N^* = N^*(\tau_s, \tau_n, \lambda)$.

Totally differentiating $C'(N^*(\tau_s, \tau_n, \lambda)) = \Pi(\tau_s, \tau_n, \lambda, N^*(\tau_s, \tau_n, \lambda))$ yields:

$$\frac{dN^*}{N^*} = \frac{[(\pi(1+M) - \pi(p))d\tau_s + \pi(p)d\tau_n + (\tau_n - \tau_s)\pi'(p)p'(.)d\lambda]\gamma}{\Pi}, \tag{4}$$

where $\gamma = C'(N^*)/(N^* C''(N^*))$ is the elasticity of innovation with respect to an increase in the profit from innovation. An increase in patent length in either market will increase the profits from innovation. Thus, it provides innovators incentives to increase the amount of innovation. Furthermore, strengthening the enforcement level also will increase the

[5] It is not reasonable if the North chooses λ first and then two countries choose τ_n and τ_s. If the two countries choose τ_n and τ_s first, it does not affect the main conclusion of this paper in Proposition 1, in which the innovating country does not have an incentive to completely eliminate arbitrage after the patent expires in the non-innovating country.

number of inventions because the profits returned from an innovation improve.

Substituting $N^*(\tau_s, \tau_n, \lambda)$ into equation (2), the North's welfare function can be rewritten as

$$\tilde{W}_n(\tau_s, \tau_n, \lambda) = W_n(\tau_s, \tau_n, \lambda, N^*(\tau_s, \tau_n, \lambda)). \tag{5}$$

Similarly, the South's welfare function can be rewritten as

$$\tilde{W}_s(\tau_s, \tau_n, \lambda) = W_s(\tau_s, N^*(\tau_s, \tau_n, \lambda)). \tag{6}$$

In stage 1, the North's government chooses the enforcement level and patent length. Taking the derivative of equation (5) with respect to λ, τ_n, we can obtain the following first order conditions:

$$\begin{aligned}
\frac{\partial \tilde{W}_n}{\partial \lambda} &= \frac{\partial W_n}{\partial \lambda} + \frac{\partial W_n}{\partial N^*}\frac{\partial N^*}{\partial \lambda} \\
&= (\tau_n - \tau_s)(\pi'(p(\lambda)) + s'(p(\lambda)))p'(\lambda)N^* \qquad (7)\\
&+ (\tau_s s(p_m) + (1-\tau_n)s(c) + (\tau_n - \tau_s)s(p(\lambda)))\frac{(\tau_n - \tau_s)\pi'(p(\lambda))p'(\lambda)N^*\gamma}{\Pi} - G'(\lambda) = 0,
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial \tilde{W}_n}{\partial \tau_n} &= \frac{\partial W_n}{\partial \tau_n} + \frac{\partial W_n}{\partial N^*}\frac{\partial N^*}{\partial \tau_n} \\
&= (\pi(p(\lambda)) + s(p(\lambda)) - s(c))N^* \qquad (8)\\
&+ (\tau_s s(p_m) + (1-\tau_n)s(c) + (\tau_n - \tau_s)s(p(\lambda)))\pi(p(\lambda))\frac{N^*\gamma}{\Pi} = 0.
\end{aligned}$$

In the South, the government has one decision to make. The necessary condition is:

$$\begin{aligned}
\frac{\partial \tilde{W}_s}{\partial \tau_s} &= \frac{\partial W_s}{\partial \tau_s} + \frac{\partial W_s}{\partial N^*}\frac{\partial N^*}{\partial \tau_s} \\
&= (s(p_m) - s(c))MN^* \\
&+ M(\tau_s s(p_m) + (1-\tau_s)s(c))\frac{(\pi(1+M) - \pi(p(\lambda)))N^*\gamma}{\Pi} = 0. \qquad (9)
\end{aligned}$$

In order simplify our analysis, we assume that enforcement against illegal arbitrage is costless (i.e. $G(.) = 0$), in section 3 and section 4.[6] The pure Nash equilibrium can be solved by equations (7), (8) and (9). Next, we characterize the properties of the Nash equilibrium.

Equation (7) shows that the marginal social welfare effects of increasing enforcement level contain two effects. The first effect is, given N, the marginal social deadweight loss caused by the increased level of enforcement. The second effect is the social surplus generated by the new products resulting from the increased profit in the North market. Thus, the optimal enforcement policy equates the marginal cost of increased enforcement, which is the increased deadweight loss, to the marginal benefit.

The condition for the optimal level of enforcement can be rewritten as:

$$\Delta'(p(\lambda)) = \frac{\pi'(p(\lambda))[\tau_s s(p_m) + (1-\tau_n)s(c) + (\tau_n - \tau_s)s(p(\lambda))]\gamma}{\tau_s \pi(1+M) + (\tau_n - \tau_s)\pi(p(\lambda))}, \tag{10}$$

where $\Delta(p(\lambda)) = s(c) - \pi(p(\lambda)) - s(p(\lambda))$, the social deadweight loss when the price of a differentiated product is $p(\lambda)$.

Note that since $p(\bar{\lambda}) = p_m$ for all $\lambda \geq \bar{\lambda}$, we get that $\pi'(p(\lambda))$ is close to zero as λ goes to $\bar{\lambda}$, which means the marginal benefit of enforcement goes to 0 as λ goes to $\bar{\lambda}$. Moreover, $\Delta'(p(\lambda))$ is positive and not close to zero as the λ converges to zero. Thus, the left hand side (hereinafter LHS) of equation (10) is greater than the RHS at $\bar{\lambda}$. This means that, in the neighborhood of $\bar{\lambda}$, a small decrease in λ will increase social welfare in the North for any (τ_s, τ_n), which guarantees that the North's optimal enforcement policy will be less

[6] It is surprising that we will find that the North will not choose to completely eliminate arbitrage even though enforcement is costless.

than $\bar{\lambda}$. On the other hand, we examine equation (10) at $\lambda = 0$ (and thus $p(\lambda = 0) = c$). According to Figure 2, we can show that $\frac{d\Delta(p(\lambda))}{d\lambda}$ is positive and close to zero when $p(\lambda)$ is close to c. Furthermore, $\pi'(p(\lambda))$ is positive when $p \in [c, p_m)$. Thus, The RHS of equation (10) is greater than the LHS when $\lambda = 0$. Since $\tilde{W}_n$ is a continuous function of λ, we can therefore conclude that there exists an interior optimal $\lambda^*(\tau_s, \tau_n) \in (0, \bar{\lambda})$ for any (τ_s, τ_n) such that the first order condition for λ is satisfied.

Figure 2 illustrates the existence of an optimal level of enforcement to prevent arbitrage that is smaller than $\bar{\lambda}$,[7] which means the price is lower than the monopoly price in the North after the patent expires in the South. This gives us the following proposition.

Proposition 1: *If the markets are not perfectly segmented, the North does not have an incentive to completely eliminate arbitrage after the patent expires in the South, even though enforcement is costless.*

Notice that a corner solution at $\lambda = 0$ is possible if the marginal enforcement cost is large enough. Furthermore, Proposition 1 will not be affected if enforcement is not costless.

Next, we investigate the relationship between τ_s and τ_n when the price $p(\lambda)$ is between the monopoly price p_m and the competitive price c. Assume that γ is constant.[8] The op-

[7] If we rearrange equation (10), we can find that Figure 2 has a unique intersection if $\frac{\Delta'(p(\lambda))}{\pi'(p(\lambda))}$ is increasing in λ.

[8] This assumption follows Grossman and Lai (2004) and is satisfied if the cost function has this form, $C(N) = a + bN^{\alpha}$ with $\alpha > 1$, where a, b and α are constant. We can obtain that γ is equal to $\frac{1}{\alpha-1}$ which is a constant.

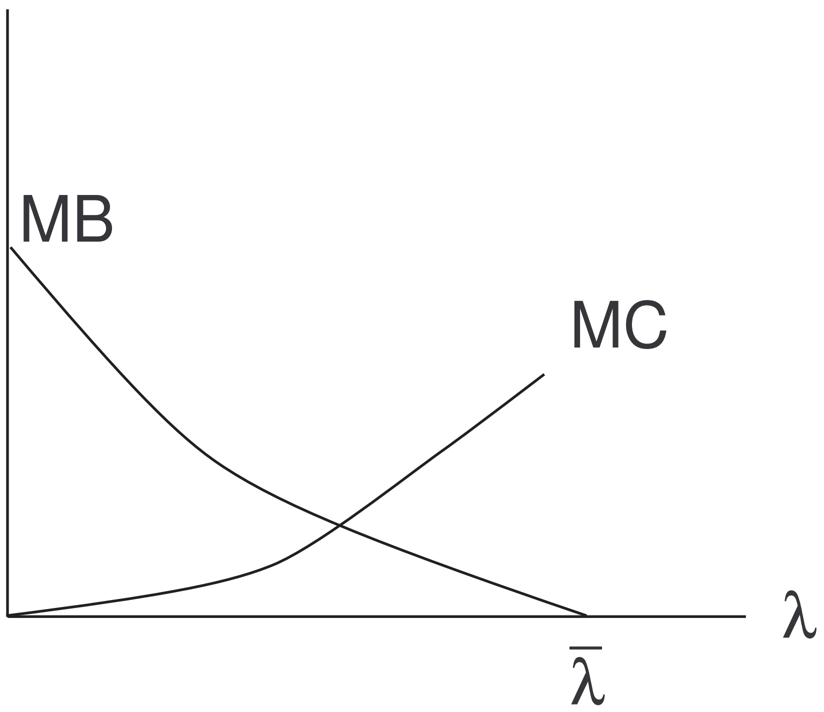

Figure 2: Marginal benefit and marginal cost of increasing enforcement level, λ.

timal choice of patent length in the North at an interior solution will satisfy the condition that equation (8) is equal to zero. Thus, we obtain

$$\Delta(p(\lambda)) = (\tau_s s(p_m) + (1-\tau_n)s(c) + (\tau_n - \tau_s)s(p(\lambda)))\pi(p(\lambda))\frac{\gamma}{\Pi}. \tag{11}$$

The LHS of equation (11) is the marginal social deadweight loss caused by a patent length extension in the North. The RHS of equation (11) is the increase in the consumers' surplus caused by the new innovation, which occurred due to the additional patent length extension. Based on equation (11), we obtain the following results:

Proposition 2: *If* $\tau_s \leq A_n(\lambda) = \frac{\gamma s(c)}{\gamma(\Delta(p_m)+\pi(p_m))+(1+M)\frac{\Delta(p(\lambda))}{\pi(p(\lambda))}}$, *then the North's optimal choice of patent length on* $[\tau_s, 1]$ *is:*

$$\tau_n = Min\{\frac{\gamma s(c) - [\frac{((1+M)\pi(p_m)-\pi(p(\lambda)))\Delta(p(\lambda))}{\pi(p(\lambda))} + \gamma(\Delta(p_m)+\pi(p_m)-\Delta(p(\lambda))-\pi(p(\lambda)))]\tau_s}{(\gamma+1)\Delta(p(\lambda))+\gamma\pi(p(\lambda))}, 1\}. \tag{12}$$

$A_n(\lambda)$ is the intersection of the North's *"conditional best response function"*[9] and the 45 degree line, which can be solved by setting $\tau_s = \tau_n = A_n(\lambda)$ in equation (11).

By using equation (4) and rearranging equation (11), we can solve the conditional best response function, in which τ_n is a linear function of τ_s. Proposition 2 also indicates that the optimal patent length in the North is negatively related to the South's patent length, which means the South's patent life is a strategic substitute for the North's patent life.

Next, we examine how the North responds to relaxing the assumption of perfect segmentation between markets. Equation (11) can be rearranged and rewritten as:

$$\frac{\Delta(p(\lambda))}{\pi(p(\lambda))} = \frac{[\tau_s s(p_m) + (1-\tau_n)s(c) + (\tau_n-\tau_s)s(p(\lambda))]\gamma}{\tau_s\pi(1+M) + (\tau_n-\tau_s)\pi(p(\lambda))}. \tag{13}$$

[9]The *"conditional best response function"* is the reaction function given a $\lambda \in (0, \bar{\lambda})$. This is not the real best response function since λ is also an endogenous variable in this model.

In order to simplify this analysis, we would like to make the following assumption.

Assumption 1: *For all* $\lambda \in [0, \bar{\lambda}]$, $\frac{\Delta(p(\lambda))}{\pi(p(\lambda))}$ *is increasing in* λ.

Assumption 1 rules out the cases wherein demand function is "too convex" in p. It is just a technical assumption which is necessary for our proof and is easily satisfied. For example, the linear demand function satisfies this assumption.

Equation (13) can be illustrated in Figure 3. In the case of the constraint of that $\tau_n \geq \tau_s$, the RHS of equation (13) is decreasing in τ_n and λ. If λ increases, then the RHS of equation (13) will shift down. Moreover, the LHS of equation (13) will shift up if λ goes up. Thus, by Figure 3, it shows that for any given τ_s, the patent length of the North is decreasing in λ if Assumption 1 is satisfied.

Propositions 1 and 2 imply that, for any given τ_s, the innovating country is willing to reduce its market price and give a longer patent life if the assumption of perfect segmentation is relaxed. This result is similar to the result in Gilbert and Shapiro (1990). For given the South's patent length, τ_s, allowing some arbitrage is an alternative way to lower the monopoly power of innovators in the North. Therefore, it can be interpreted as increasing the North's patent breadth with extending the North's patent length.[10] Thus, the North's welfare can be improved.

[10] In Gilbert and Shapiro (1990), the patent breadth has many different definitions. They indicates that any definition of breadth involves the idea that a broader patent allows the innovator to make a higher flow rate of profits during its patent life. Thus, they just simplify the breadth as the profit flow. Since potential arbitrage will make the flow rate of profits decline when the North patent is still in force but not in the South, we can consider the existence of potential arbitrage as another form of patent breadth.

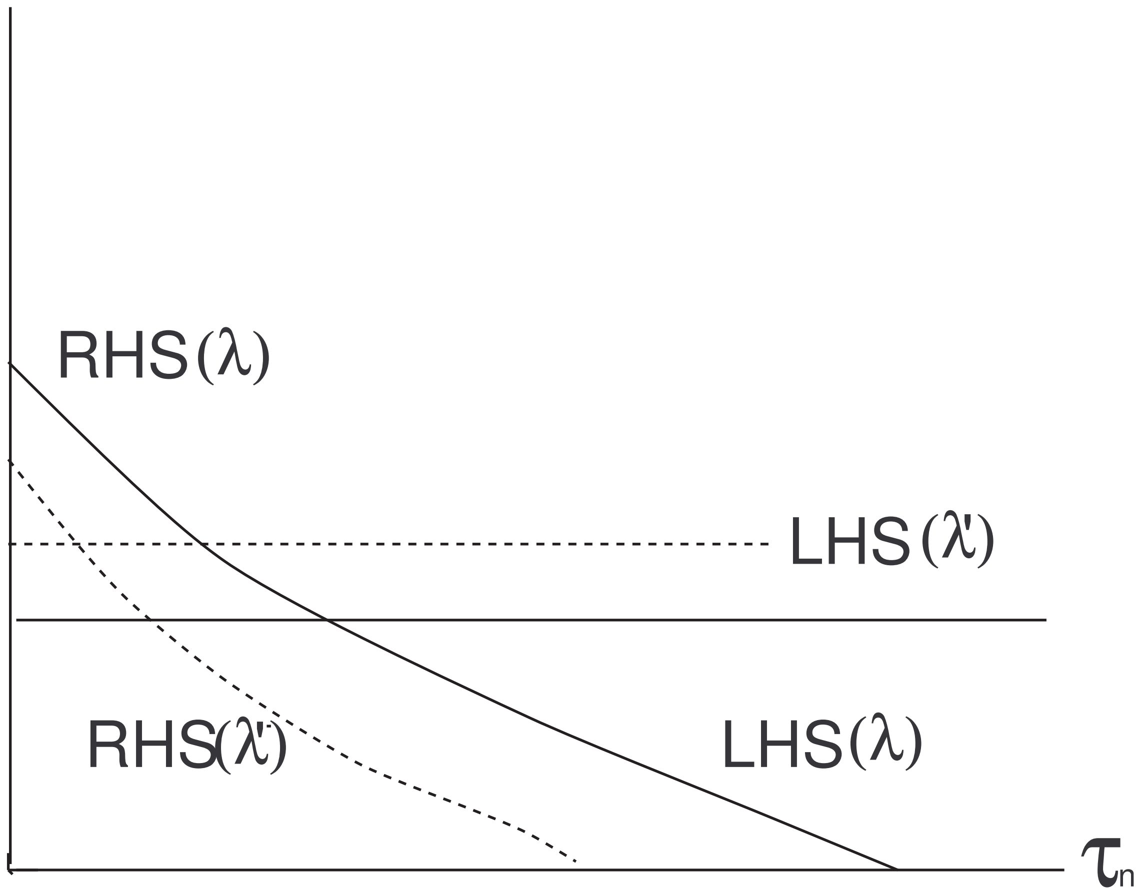

Figure 3: Right hand side (RHS) and left hand side (LHS) of equation (16) where $\lambda' > \lambda$.

We will next examine how the conditional best response function of the South is affected by the existence of arbitrage. If an interior solution exists, equation (9) can be rewritten as

$$\Delta(p_m) + \pi(p_m) = (\tau_s s(p_m) + (1-\tau_s)s(c))\frac{(\pi(1+M) - \pi(p(\lambda)))\gamma}{(1+M)\tau_s\pi(p_m) + (\tau_n - \tau_s)\pi(p(\lambda))}. \tag{14}$$

Similar to equation (11), the LHS of equation (14) is the marginal social deadweight loss that the South will incur if it increases its patent length. The RHS of equation (14) is the marginal social benefit from the additional patent extension.

Equation (14) illustrates a trade-off for the South between dynamic welfare gain and static efficiency losses similar to that for the North. However, there are two differences between equations (11) and (14). The first is that the static welfare cost of increasing patent protection in the South has an effect similar to raising its term of trade, which includes the monopoly profit and the social deadweight loss, since its domestic price increases. The static welfare cost of extending patent protection in the North contains an additional effect; the social deadweight loss due to the monopoly power and enforcement level.

The second difference is that the extension of patent protection in the North will just raise the profits of the new products in the Northern market with arbitrage. An increase in the South's patent life will make an impact on the profits of innovation in both the South and the North without arbitrage. Thus, the change in the South's patent protection still could have a large impact on innovation even though the South is a very small market.

Based on equation (14), we have the following results:

Proposition 3: *If* $\tau_n \geq A_s(\lambda) = \frac{\gamma s(c)}{(\Delta(p_m)+\pi(p_m))\gamma + \frac{(\Delta(p(\lambda))+\pi(p(\lambda)))(1+M)\pi(p_m)}{(1+M)\pi(p_m)-\pi(p(\lambda))}}$, *then the South's optimal choice of patent length on* $[0, \tau_n]$ *is*

$$\tau_s = Max\{0, \frac{\gamma s(c) - [\frac{\pi(p(\lambda))(\pi(p_m)+\Delta(p_m))}{(1+M)\pi(p_m)-\pi(p(\lambda))}]\tau_n}{(1+\gamma)(\Delta(p_m)+\pi(p_m))}\}. \tag{15}$$

$A_s(\lambda)$ can be solved by letting $\tau_n = \tau_s = A_s(\lambda)$ in equation (14). The conditional best response function, equation (15), can also be derived from the first order condition, which shows that the optimal patent length in the South is also negatively related to the Northern patent length. This means the two patent lengths are strategic substitutes for each other.

Next, we will examine the existence of a Nash equilibrium in this non-cooperative patent-setting game between the North and the South. In Proposition 2 and Proposition 3, we derive the conditional best response function of the North and South, in the case of $\tau_n \geq \tau_s$. It should be noted that these results will not be sufficient to establish that this pair is a Nash equilibrium. It is only shown that these patent lives are a best response for the North on $[\tau_s, 1]$ and for the South on $[0, \tau_n]$. We, therefore, need to show that the North has no incentive to undercut its patent life and the South will not set $\tau_s \in [\tau_n, 1]$.

We prove the existence of a Nash equilibrium by the following lemmas:

Lemma 1: *If the demand curve is linear, the Southern market size* M *is smaller than one, and* γ *is constant, there exists a unique pair* (τ_n, τ_s) *satisfying equations (12) and (15) with* $\tau_n \geq A_n(\lambda)$ *and* $\tau_s \leq A_s(\lambda)$.

The proofs of all Lemmas are in the Appendix. We can conclude that τ_n and τ_s are continuous functions of λ, respectively. The continuity of τ_n and τ_s is important for us to prove the existence of Nash equilibrium. Thus, we can rewrite $(\tau_n, \tau_s) = (\tau_n(\lambda), \tau_s(\lambda))$.

Lemma 2: *If the demand curve is linear, the South's market size M is smaller than one and γ is constant, then there does not exist any point of intersection between the two conditional best response functions which involves $\tau_n < \tau_s$.*

Lemma 2 demonstrates that there cannot exist a Nash equilibrium in which $\tau_n < \tau_s$.

Lemma 3: *If the demand curve is linear, the South's market size M is smaller than one and γ is constant, then*

$$(i)\frac{\partial \tilde{W}_n}{\partial \tau_n} > 0 \quad for \quad \tau_s \in [0, A_s(\lambda)],$$
$$(ii)\frac{\partial \tilde{W}_s}{\partial \tau_s} < 0 \quad for \quad \tau_n \in [A_n(\lambda), 1].$$

in the case of $\tau_n < \tau_s$.

Lemma 3 ensures that the solution in Lemma 1 will be a Nash equilibrium. Therefore, we obtain the following result.

Proposition 4: *If the demand curve is linear, the South's market size M is smaller than one and γ is constant, there exists a Nash equilibrium $(\lambda^*, \tau_n^*, \tau_s^*)$ satisfying equations (10), (12), and (15) in this patent-setting game.*

Lemma 1 shows that, for any given $\lambda \in [0, \bar{\lambda}]$, there exists a unique solution for the two conditional best response functions, denoted $(\tau_n(\lambda), \tau_s(\lambda))$, while imposing the constraint that $\tau_n \geq \tau_s$, in which $\tau_n(\lambda)$ and $\tau_s(\lambda)$ are continuous in λ. Next, we need to show that there exists a λ^* which maximizes the social welfare of the North.

First, we show that there does not exist a corner solution in this problem. According Proposition 1 and Figure 2, we obtain that $\frac{\partial \tilde{W}_n(\tau_n(\lambda^*),\tau_s(\lambda^*),\lambda^*)}{\partial \lambda} > 0$ when $\lambda = 0$ and $\frac{\partial \tilde{W}_n(\tau_n(\lambda^*),\tau_s(\lambda^*),\lambda^*)}{\partial \lambda} < 0$ when $\lambda = \bar{\lambda}$. Thus, there does not exist any corner solution in this problem.

Next, we need to show that there exists a λ^* such that

$$\frac{\partial \tilde{W}_n(\tau_n(\lambda^*), \tau_s(\lambda^*), \lambda^*)}{\partial \lambda} = 0.$$

We rearrange equation (7) and rewrite it as

$$\frac{\Delta'(p(\lambda))}{\pi'(p(\lambda))} = \frac{[\tau_s s(p_m) + (1-\tau_n)s(c) + (\tau_n - \tau_s)s(p(\lambda))]\gamma}{\tau_s \pi(1+M) + (\tau_n - \tau_s)\pi(p(\lambda))}. \tag{16}$$

Since τ_n and τ_s are continuous functions of λ, this means that the RHS of equation (16) is continuous in λ. Since $\Delta'(p(\lambda)$ and $\pi'(p(\lambda))$ are also continuous in λ, the LHS of equation (16) is also continuous in λ for $\lambda \in [0, \bar{\lambda})$. Since the LHS of equation (16) is close to 0 as λ is close to zero, the LHS is less than the RHS of equation (16). Furthermore, the LHS is approaching infinity as λ is close to $\bar{\lambda}$ but the RHS is not. Thus, by the intermediate value theorem, we can conclude the existence of λ^*.

Lemma 2 shows that it is impossible for a Nash equilibrium in which $\tau_n < \tau_s$ to exist. Lemma 3 ensures that $(\tau_n(\lambda^*), \tau_s(\lambda^*), \lambda^*)$ is a Nash equilibrium. This completes the proof of Proposition 4.

If we compare the Nash equilibria with complete segmented market and with relaxing the assumption of complete segmentation, we find that the North will increase the patent protection as well as the South. However, we do not conclude which country extends longer. In the following condition, we obtain an unambiguous result.

Corollary: *If, in a Nash equilibrium, the North's optimal patent length is its economic life, (i.e., $\tau_n = 1$), in the case of perfectly segmented markets, then the gap of optimal patent length between the two markets will be shorter in a Nash equilibrium in the case of markets that are not completely segmented.*

This Corollary follows from Propositions 1, 2, 3 and 4. Suppose that (τ_s^e, τ_n^e) is the Nash equilibrium when $\lambda \geq \bar{\lambda}$. It is hypothesized that $\tau_s^e < 1$ and $\tau_n^e = 1$. If the assumption of perfect segmentation is relaxed, Proposition 4 indicates that there exists a Nash equilibrium , $(\tau_s^*, \tau_n^*, \lambda^*)$. By Proposition 1, it implies that $\lambda^* < \bar{\lambda}$. Thus, Proposition 2 demonstrates that the conditional best response function of the North will shift up and Proposition 3 indicates that the conditional best response function of the South will shift to the right. Therefore, it is easy to show that τ_s^* is greater than τ_s^e and $\tau_n^* \leq 1$. This completes the proof of the Corollary.

4 Efficient International Agreements

The Nash equilibrium outcome can be compared with the efficient agreement that would be chosen if the policies were chosen to maximize the sum of the North and the South welfare. Such an outcome would arise if the North and the South were able to commit to an agreement on patent lives, with lump sum transfers being made between countries to achieve the desired distribution of the income between countries. In the following, we will examine how total welfare is affected by the potential for arbitrage if the markets are not perfectly segmented. We will show that harmonization will never achieve global efficiency in this model.

We prove our argument by a two-step approach. First, like equation (7), we will show that the globally socially optimal λ will be smaller than $\bar{\lambda}$. This means that relaxing enforcement away from perfect enforcement will improve global efficiency. Next, we will show that, for any given λ and N, increasing τ_n and reducing τ_s will improve the global welfare.

Since there are only two countries in this economy, the world welfare is sum of the North and South welfare. Thus, combining equations (2) and (3), we obtain the world welfare function

$$\begin{aligned} W(\tau_n, \tau_s, \lambda, N(\tau_n, \tau_s, \lambda)) &= W_n + W_s \\ &= (\tau_s(1+M)\pi(p_m) + (\tau_n - \tau_s)\pi(p(\lambda)))N + (\tau_s s(p_m) + (1-\tau_n)s(c) \\ &+ (\tau_n - \tau_s)s(p(\lambda)))N + M(\tau_s s(p_m) + (1-\tau_s)s(c))N - C(N). \quad (17) \end{aligned}$$

In order to maximize total welfare, the central planner can choose three variables, λ, τ_n, and τ_s. Taking the derivative with respect to λ, we obtain the following

$$\begin{aligned} &(\tau_n - \tau_s)(\pi'(\lambda) + s'(p(\lambda)))p'(\lambda)N \\ &+ \frac{(\tau_s s(p_m) + (1-\tau_n)s(c) + (\tau_n - \tau_s)s(\lambda))(\tau_n - \tau_s)\pi'(p(\lambda))p'(\lambda)\gamma}{\Pi}. \end{aligned} \quad (18)$$

Similar to equation (7), since $\pi'(p_m)$ is zero and $s'(p_m)$ is positive, equation (18) is positive when $\lambda = \bar{\lambda}$ (and thus the market price is p_m). This means that the optimal $\lambda < \bar{\lambda}$ (and so the optimal p is smaller than p_m).

Next, we will show that for any given $p(\lambda)$ and N, increasing τ_n and reducing τ_s will improve global welfare. According to equation (4), the number of new products depends on the marginal cost and the marginal revenue of a new innovation. Suppose that the number of innovations and $p(\lambda)$ are fixed. We can derive the following:

$$\frac{d\tau_s}{d\tau_n}|_{N=\bar{N}} = -\frac{\pi(p(\lambda))}{M\pi(p_m) + (\pi(p_m) - \pi(p(\lambda)))}.$$

This shows that, for any given $p(\lambda)$ and N, the patent protections of the two countries go in the opposite direction if the marginal revenue remains unchanged.

Since $\frac{dW}{d\tau_n}|_{N=\bar{N}} = \frac{\partial W}{\partial \tau_n} + \frac{\partial W}{\partial \tau_s}\frac{\partial \tau_s}{\partial \tau_n}|_{N=\bar{N}}$, we get the following result:

$$\begin{aligned}
\frac{dW}{d\tau_n}|_{N=\bar{N}} &= N(\pi(p(\lambda) + s(p(\lambda)) - s(c)) \\
&- (\Delta(p(\lambda) - (1+M)\Delta(p(\lambda))))\frac{N\pi(p(\lambda))}{M\pi(p_m) + (\pi(p_m) - \pi(p(\lambda))} \\
&= \frac{N(M+1)(\pi(p_m)(s(p_m) - s(c))}{M\pi(p_m) + (\pi(p_m) - \pi(p(\lambda))}(\frac{s(p(\lambda)) - s(c)}{s(p_m) - s(c)} - \frac{\pi(p(\lambda))}{\pi(p_m)}) \\
&= \frac{N(M+1)}{M\pi(p_m) + (\pi(p_m) - \pi(p(\lambda))}\frac{\pi(p_m)}{\pi(p(\lambda))}(\frac{\Delta(p_m)}{\pi(p_m)} - \frac{\Delta(p(\lambda))}{\pi(p(\lambda))}) > 0.
\end{aligned} \tag{19}$$

If Assumption 1 is satisfied, equation (19) demonstrates that the central planner can increase patent protection in the North and reduce it in the South in order to improve the world welfare. This result can be related to the results in Gilbert and Shapiro (1990). In a two-country model, global welfare can be considered as social welfare in a closed economy. Therefore, the uniform universal standard can be thought of as granting the innovators monopoly power for a certain period in a closed economy. This kind of patent policy is not efficient which is indicated by Gilbert and Shapiro (1990). Increasing the length of IPRs protection in the North and decreasing the length of IPRs protection in the South can be interpreted as lengthening the patent. Allowing arbitrage, which is an alternative way of reducing the market power of the innovators, also can be considered as reducing the patent breadth. Therefore, the inefficiency problem can be improved in a global approach. Thus, we obtain the following

Proposition 5: *If for all* $\lambda \in [0, \bar{\lambda}]$, $\frac{\Delta(p(\lambda))}{\pi(p(\lambda))}$ *is increasing in* λ, *the uniform universal standard for IPRs protection will never achieve global social efficiency.*

5 Conclusions

We have developed a simple two-country model of IPRs protection, in which the assumption of market segmentation is relaxed. This paper examines how arbitrage affects the setting of patent policy in innovating and non-innovating countries. When patent lives are set non-cooperatively, it is surprising that the innovating country will not completely eliminate arbitrage, even though enforcement is costless. The reason is because granting the innovators monopoly power is not an efficient way to reward innovators. The existence of arbitrage can lower the market price and improve the innovating country welfare. The innovating country can strengthen the length of IPRs protection to achieve the maximal welfare. The results are analogous to the results in Gilbert and Shapiro (1990).

There exists a Nash equilibrium in this patent setting game. It is also shown that the innovating countries will offer a stronger IPRs protection than non-innovating countries when arbitrage is possible. Finally, we prove that a uniform universal standard for patent protection will never achieve global social efficiency.

6 Appendix

Proof of Lemma 1: Lemma 1 can be proven using Figure 4. First, we need to show that $A_n(\lambda) > A_s(\lambda)$. Let $B_n(\tau_s)$ be the best response function for the North, which is defined by equation (12) in Proposition 2. As in Proposition 2, $B_n(\tau_s)$ is a non-decreasing , continuous, and piecewise linear function on $[0, A_n(\lambda)]$.

$A_n(\lambda)$ is defined as the intersection of $B_n(\tau_s)$ and the 45 degree line, i.e. $B_n(\tau_s) = \tau_s$. By equation (8), we can get:

$$\Delta(A_n(\lambda))(\tau_n(1+M)\pi(p_m)) = (\tau_n s(p_m) + (1-\tau_n)s(c))\pi(A_n(\lambda))\gamma. \tag{20}$$

By equation (9), it yields:

$$(\Delta(p_m) + \pi(p_m))\tau_s(1+M_s)\pi(p_m)$$

$$= (\tau_s s(p_m) + (1-\tau_s)s(c))((1+M)\pi(p_m) - \pi(A_s(\lambda)))\gamma. \tag{21}$$

Let $F(\tau) = \frac{(\tau s(p_m)+(1-\tau)s(c))\gamma}{\tau(1+M)\pi(p_m)}$, and we can get:

$$F(A_n(\lambda)) = \frac{\Delta(p(A_n(\lambda)))}{\pi(p(A_n(\lambda)))} < \frac{1}{2}. \tag{22}$$

Since $\frac{\Delta}{\pi} = \frac{1}{2}$ if the demand is linear and M is less than 1, and we can get:

$$F(A_s(\lambda)) = \frac{\Delta(p_m) + \pi(p_m)}{(1+M)\pi(p_m) - \pi(A_s(\lambda))} > \frac{\Delta(p_m) + \pi(p_m)}{(1+M)\pi(p_m)} > \frac{3}{4}. \tag{23}$$

Referring to the definitions of $A_n(\lambda)$ and $A_s(\lambda)$ in Propositions 2 and 3, it can be shown that $A_n(\lambda) > A_s(\lambda)$ if and only if $F(A_s(\lambda)) > F(A_n(\lambda))$. Therefore, we can conclude that the intersection of the North's reaction function and the 45 degree line, $A_n(p(\lambda))$, is to the right of the intersection of the South's reaction functions and the 45 degree line, $A_s(p(\lambda))$, under the assumptions of Lemma 1 and the constraint $\tau_n \geq \tau_s$.

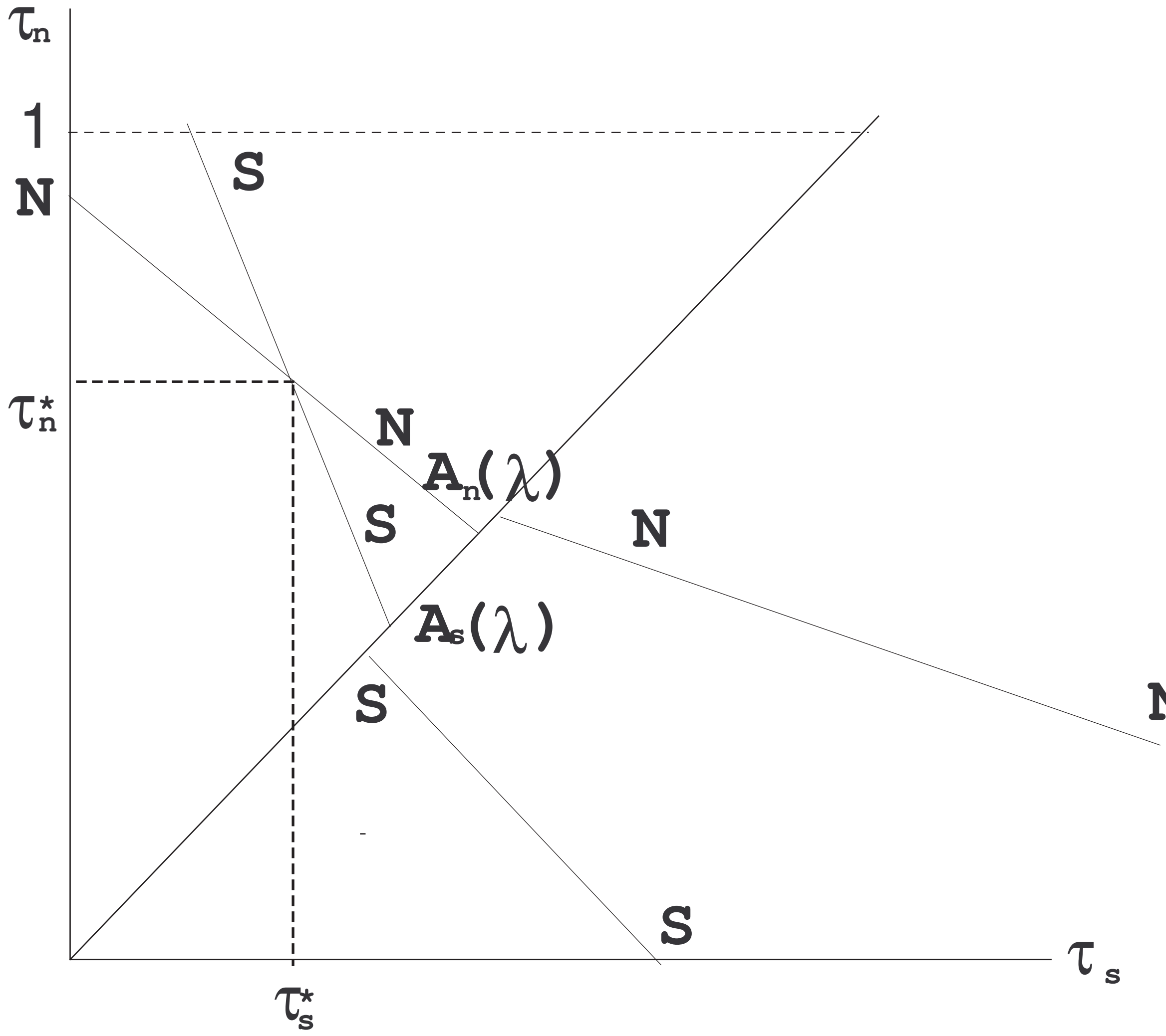

Figure 4: The conditional best response function of the North and South.

We will now show that there exists a unique solution satisfying equations (12) and (15), using the following five cases.

Case 1: If $A_s(\lambda) \geq 1$, which means $A_n(\lambda) > 1$. Since $B_n(\tau_s) = 1$ for all $\tau_s \in [0, 1]$ and $B_s(\tau_n) = 1$ for all $\tau_n \in [0, 1]$, the only solution is $(1, 1)$.

Let a "-" over a variable denote the maximum value of that variable and a "-" under a variable denote the minimum value of that variable.

Case 2: If $A_s(\lambda) < 1$, $B_s(1) > 0$, and $\bar{B}_n^{-1}(1) \geq B_s(1)$. Since $A_n(\lambda) > A_s(\lambda)$, $B_n^{-1}(\tau_n)$ is greater than $B_s(\tau_n)$ for all $\tau_n \in [A_n(\lambda), 1)$. Additionally, $B_n(\tau_s) = 1$ for all $\tau_s \in [0, \bar{B}_n^{-1}(1))$. Therefore, the unique solution is $(1, B_s(1))$.

Case 3: If $A_s(\lambda) < 1$, $B_s(1) > 0$ and $\bar{B}_n^{-1}(1) < B_s(1)$. Since $\bar{B}_n^{-1}(1) < B_s(1)$, there is no solution for all $\tau_s \in [0, B_s^{-1}(1)]$. $B_n(\tau_s)$ is a linear, decreasing function in τ_s. Therefore, $B_s^{-1}(.) - B_n(.)$ is positive when $\tau_s = B_s(1)$ because $B_s^{-1}(B_s(1)) = 1$ and $B_n(B_s(1)) < 1$. Furthermore, $B_s^{-1}(.) - B_n(.)$ is negative when $\tau_s = A_s(\lambda)$ because $B_s^{-1}(B_s(A_s(\lambda))) = A_s(\lambda) < A_n(\lambda) = B_n(A_n(\lambda)) < B_n(A_s(\lambda))$. By continuity, we can conclude that there exists a pair of (τ_n, τ_s) satisfying $1 > \tau_n > A_n(\lambda)$ and $0 < \tau_s < A_s(\lambda)$.

Case 4: If $A_s(\lambda) < 1$, $B_s(1) = 0$ and $B_n(0) \geq \underline{B}_s^{-1}(0)$. $B_n(\tau_s)$ is greater than $B_s^{-1}(\tau_s)$ for all $\tau_s \in (0, A_s(\lambda)]$. Also, $B_s(\tau_n) = 1$ for all $\tau_n \in [\underline{B_s^{-1}}(0), 1]$. Therefore, the unique solution is $(B_n(0), 0)$.

Case 5: If $A_s(\lambda) < 1$, $B_s(1) = 0$ and $B_n(0) < \underline{B}_s^{-1}(0)$. Since $B_n(0) < \underline{B}_s^{-1}(0)$, there is no solution for $\tau_s = 0$. Since $B_s(\tau_n)$ is a linear, decreasing function in τ_n, $B_s(\tau_n) > 0$ for all $\tau_n \in (\underline{B}_s^{-1}(0), A_s(\lambda)]$. Therefore, $B_n^{-1}(.) - B_s(.)$ is negative when $\tau_n = B_n(0)$. This is because $B_n^{-1}(B_n(0)) = 0$ and $B_s(B_n(0)) > B_s(\underline{B}_s^{-1}(0)) = 0$. Furthermore, $B_n^{-1}(.) - B_s(.)$ is positive when $\tau_n = A_n(\lambda)$ because $B_n^{-1}(B_n(A_n(\lambda))) = A_n(\lambda) > A_s(\lambda) = B_s(A_s(\lambda)) > B_s(A_n(\lambda))$. By continuity, we can conclude that there exists a pair of (τ_n, τ_s) satisfying $1 > \tau_n > A_n(\lambda)$ and $0 < \tau_s < A_s(\lambda)$. This completes the proof of Lemma 1.

Proof of Lemma 2:

To prove Lemma 2, we first show that the intersection point of the North's conditional best response function and the 45 degree line is to the right of the intersection point of the South's conditional best response function and the 45 degree line in the case of $\tau_n < \tau_s$. We then demonstrate the slope of the North's best response function of the North is flatter than the South's.

Since patent protection is longer in the South, arbitrage will potentially occur there. This changes the profit function, the welfare functions, and the social surplus. We will redefine most of the equations in sections 2 and 3.

First, the profit function Π is:

$$\Pi(\tau_s, \tau_n, \lambda_s, N) = \tau_n(1 + M)\pi(p_m) + M(\tau_s - \tau_n)\pi(p_s(\lambda_s)). \tag{24}$$

Total differentiating $C(N^*) = \Pi(\tau_s, \tau_n, \lambda_s, N^*)$ yields:

$$\frac{dN^*}{N^*} = [(\pi + M(\pi - \pi(p_s)))d\tau_n + M\pi(p_s)d\tau_s + (\tau_s - \tau_n)M\pi'(p(\lambda_s))p'(\lambda_s)d\lambda_s]\frac{\gamma}{\Pi},$$

Similar to equations (2) and (3), the welfare functions can be rewritten as:

$$\begin{aligned}\tilde{W}_n(\tau_n, \tau_s, \lambda_s, N^*(\tau_s, \tau_n, \lambda_s)) &= (\tau_n(1 + M)\pi + M(\tau_s - \tau_n)\pi(p_s(\lambda_s)))N^* \\ &+ N^*(\tau_n s(p_m) + (1 - \tau_n)s(c)) - C(N^*), \end{aligned} \tag{25}$$

and

$$\tilde{W}_s(\tau_n, \tau_s, \lambda_s, N^*(\tau_s, \tau_n, \lambda_s)) = N^* M(\tau_n s(p_m) + (\tau_s - \tau_n) s(p_s(\lambda_s)) + (1 - \tau_s) s(c)). \quad (26)$$

Differentiating equation (25) with respect to τ_n, we obtain the following first order condition:

$$\begin{aligned} \frac{\partial \tilde{W}_n}{\partial \tau_n} &= N^*(-\Delta + M(\pi - \pi(p_s(\lambda_s)))) \\ &+ (\tau_n s(p_m) + (1 - \tau_n) s(c))(\pi + M(\pi - \pi(p_s(\lambda_s)))) \frac{N^* \gamma}{\Pi} = 0. \end{aligned} \quad (27)$$

Again, taking the derivative of equation (26) with respect to λ_s, τ_s, we obtain the following first order conditions:

$$\begin{aligned} \frac{\partial \tilde{W}_s}{\partial \tau_s} &= N(s(p_s(\lambda_s)) - s(c)) \\ &+ [(\tau_n s(p_m) + (\tau_s - \tau_n) s(p_s(\lambda_s)) + (1 - \tau_s) s(c)) M \pi(p_s(\lambda_s))] \frac{N^* \gamma}{\Pi} = 0, \end{aligned} \quad (28)$$

and

$$\begin{aligned} \frac{\partial \tilde{W}_s}{\partial \lambda_s} &= N^* s'(p_s(\lambda_s))(\tau_s - \tau_n) + [(\tau_n s(p_m) + (\tau_s - \tau_n) s(p_s(\lambda_s)) \\ &+ (1 - \tau_s) s(c))] \frac{N^* \gamma (\tau_s - \tau_n) M \pi'(p_s(\lambda_s)) p_s'(\lambda_s)}{\Pi} = 0. \end{aligned} \quad (29)$$

Equation (29) shows that the South also will not eliminate all arbitrage that occurs there. The reason is the same as the North's, which we indicated in Proposition 1.

Next, we examine the conditional best response function of the North when τ_n is equal to τ_s. According to the first conditions of equations (27) and (28), the conditional best response functions should satisfy the following conditions if the interior solution is available:

$$(\Delta - M(\pi - \pi(p_s(\lambda_s))))(\tau_n(1+M)\pi)$$

$$= (\tau_n s(p_m) + (1-\tau_n)s(c))(\pi + M(\pi - \pi(p_s(\lambda_s))))\gamma,$$

and,

$$(s(c) - s(p_s(\lambda_s)))(\tau_s(1+M)\pi) = (\tau_s s(p_m) + (1-\tau_s)s(c))M\pi(p_s(\lambda_s))\gamma$$

Letting $F(\tau) = \frac{(\tau s(p_m)+(1-\tau)s(c))\gamma}{\tau(1+M)\pi}$, we get:

$$F(\tau_n) = \frac{(\Delta - M(\pi - \pi(p_s(\lambda_s))))}{\pi + M(\pi - \pi(p_s(\lambda_s)))} < \frac{1}{2} < 1 < \frac{1}{M} < \frac{s(c) - s(p_s(\lambda_s))}{M\pi(p_s(\lambda_s))} = F(\tau_s), \qquad (30)$$

Since $F(.)$ is a decreasing function in τ, equation (30) indicates that that the intersection point of the North's conditional best response function and the 45 degree line is to the right of the intersection point of the South's conditional best response function and the 45 degree line.

Next, we show that the slope of the North's conditional best response function is flatter than the South's. Differentiating equation (28) with respect to τ_n and τ_s yields the slope of the conditional best response function of the South:

$$\frac{d\tau_n}{d\tau_s}|_S = -\frac{(s(c) - s(p_s(\lambda_s)))M\pi(p_s(\lambda_s))(1+\gamma)}{(s(c) - s(p_s(\lambda_s)))[\pi + M(\pi - \pi(p_s(\lambda_s)))] + (s(p_s(\lambda_s)) - s(p_m))M\pi(p_s(\lambda_s))\gamma}.$$

Differentiating equation (27) with respect to τ_n and τ_s yields the slope of the North's *"best response function"*:

$$\frac{d\tau_n}{d\tau_s}|_N = -\frac{M\pi(p_s(\lambda_s))(\Delta - M(\pi - \pi(p_s(\lambda_s))))}{[\pi + M(\pi - \pi(p_s(\lambda_s)))][\gamma(\pi + \Delta) + \Delta - M(\pi - \pi(p_s(\lambda_s)))]}.$$

Let $Q = \pi - \pi(p_s(\lambda_s))$. If $\Delta - MQ < 0$, there is no best response function for the North since $\frac{\partial \tilde{W}_n}{\partial \tau_n} > 0$. If $\Delta - MQ > 0$, we need to show

$$\frac{(1+\gamma)}{[\pi + MQ] + (s(p_s(\lambda_s)) - s(p_m))M\frac{\pi(p_s(\lambda_s))}{s(c)-s(p_s(\lambda_s))}\gamma} > \frac{(\Delta - MQ)}{[\pi + MQ][\gamma(\pi+\Delta) + \Delta - MQ)]},$$

Since $\frac{(1+\gamma)}{[\pi+MQ]+(s(p_s(\lambda_s))-s(p_m))M\frac{\pi(p_s(\lambda_s))}{s(c)-s(p_s(\lambda_s))}\gamma} > \frac{(1+\gamma)}{[\pi+MQ]+(s(p_s(\lambda_s))-s(p_m))M\gamma} > \frac{(1+\gamma)}{[\pi+MQ]+(\pi+\Delta)M\gamma}$

, it suffices to show

$$\frac{(1+\gamma)}{[\pi + MQ] + (\pi+\Delta)M\gamma} > \frac{(\Delta - MQ)}{[\pi + MQ][\gamma(\pi+\Delta) + \Delta - MQ]}$$

Since $\gamma\pi\pi > \gamma M\pi\Delta$ and $\gamma\Delta\pi > \gamma M\Delta^2$, we yield

$$\begin{aligned}(1+\gamma)[\pi + MQ][\gamma(\pi+\Delta) + \Delta - MQ] &= (\Delta - MQ)[\pi + MQ] + \gamma(\Delta - MQ)[\pi + MQ] \\ &\quad + \gamma\pi\pi + \gamma\pi MQ + \gamma\Delta\pi + \gamma\Delta MQ + \gamma^2(\pi+\Delta)[\pi + MQ] \\ > (\Delta - MQ)[\pi + MQ] + (\Delta - MQ)(\pi+\Delta)M\gamma &= ([\pi + MQ] + (\pi+\Delta)M\gamma)(\Delta - MQ)\end{aligned}$$

This completes the proof of Lemma 2.

Proof of Lemma 3:

(i) is proven by a two-step approach. First, we show that the intersection of the Northern reaction function and the 45 degree line is to the right of $A_s(\lambda)$. We, then, show that $\frac{\partial \tilde{W}_n}{\partial \tau_n} > 0$ for $\tau_s \in [0, A_s(\lambda)]$. Equations (23) and (30) proves this. Equation (27) shows that for any (τ_n, τ_s) below the conditional best response function of North, $\frac{\partial \tilde{W}_n}{\partial \tau_n}$ is positive. This completes the proof of (i).

The proof of (ii) is analogous to (i). Equations (22) and (30) show that the intersection of the Southern conditional best response function and the 45 degree line is to the left of $A_n(\lambda)$. Equation (28) shows that for any point above the conditional best response function of South, $\frac{\partial W_s}{\partial \tau_s}$ is negative. This completes the proof of (ii).

References

[1] Braga, C.P. (1990), "The Developing Country Case for and against Intellectual Property Protection," in Siebeck, W.E.(Eds), *Strengthening Protection of Intellectual Property in Developing Countries: A Survey of the Literature.* World Bank Discussion Paper, no. 112.

[2] Chin, Judith C. and Gene M. Grossman (1990), "Intellectual Property Rights and North-South Trade," in Jones R., Krueger, A.O.(Eds), *The Political Economy of International Trade,* Cambridge, MA: Basil Blackwell, 90-107.

[3] Deardorff, Alan V. (1992), "Welfare Effects of Global Patent Protection," *Economica,* 59, 35-51.

[4] Diwan, Ishak and Dani Rodrik (1991), "Patents, Appropriate Technology, and North-South Trade," *Journal of International Economics,* 23, 79-90.

[5] The Economist (1996), "Stolen Melodies," May 11; 339, 64.

[6] Gallini, Nancy T., and Suzanne Scotchmer (2002), "Intellectual Property: When Is It the Best Incentive System?" *Innovation Policy and the Economy,* Vol 2, Adam Jaffe, Joshua Lerner and Scott Stern, eds, MIT Press, 51-78.

[7] Gilbert, Richard and Carl Shapiro (1990), "Optimal Patent Length and Breadth," *RAND Journal of Economics,* 21, 106-112.

[8] Grossman, Gene M. and Elhanan Helpmam (1994), "Protection for Sale," *American Economic Review,* 84, 833-850.

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

[10] Krugman, Paul R. (1992), "Regionalism versus Multilateralism," in J. De Melo and A. Panagarya, eds. *New Dimensions in Regional Integration.* Cambridge University Press, 58-79.

[11] Lai, Edwin L.C. and Larry D. Qiu (2003), "The North's Intellectual Property Rights Standard for the South?" *Journal of International Economics,* 59, 183-209.

[12] Maskus Keith E., (2000), *Intellectual Property Rights in the Global Economy*, Washington, DC: Institute for International Economics.

[13] Mitra, Devashish (1999), "Endogenous Lobby Formation and Endogenous Protection: A Long-Run Model of Trade Policy Determination," *American Economic Review,* 89, 1116-1134.

[14] Nordhaus, William D. (1969), *Invention, Growth, and Welfare: A Theoretical Treatment of Technological Change,* Cambridge, Mass., M.I.T. Press.

[15] Scherer, Frederic M. (1980), *Industrial Market Structure and Economic Performance,* Chicago: Rand McNally College Pub. Co.

[16] Scotchmer, Suzanne (2004), "The Political Economy of Intellectual Property Treaties," *Journal of Law, Economics, and Organizations,* 20, 415-437.

[17] Spors, Kelly K. (2004), "Canadian Drugs Are Getting Cheaper; Online Pharmacies Catering to American Cut Costs by Finding New Surppliers," *Wall Street Journal,* July 22.

[18] State Legislatures (2004), "Canadian Drugs: Panacea or Problem?" June; 30, 6.

[19] Weil, Elizabeth (2004), "Grumpy Old Drug Smugglers," *New York Times Magazine,* May 30, 42-45.