Optimal Global Patent Design

Donald J. Wright[*]

March 2004

Abstract

The optimal patent breadth and length is derived for an innovating and a non-innovating country in the presence of imitation. It is found that the innovating country chooses stronger patent protection than the non-innovating country. These patents are compared to the optimal global patent design and it is found they are too weak from a global perspective. Finally, it is shown that the innovating country is unambiguously better off, while the non-innovating country may be worse off with the optimal global patent design.

JEL Classification Numbers: O34

Keywords: Patents

[*]Economics, Faculty of Economics and Business, University of Sydney, NSW, 2006, Australia, Ph: 61+2+93516609, Fax: 61+2+93514341, Email: don.wright@econ.usyd.edu.au.

1. Introduction

In the early 1980's the United States and many countries in Europe, unhappy with the lack of protection many developing countries afforded intellectual property, lobbied strongly for intellectual property rights to be placed on the agenda of the Uruguay Round of GATT. This lobbying was successful and eventually culminated in a TRIPS (Trade Related Intellectual Property Rights) agreement being signed in April 1994, effective as of 1st January 1995. This agreement set out minimum standards of protection for intellectual property, domestic enforcement procedures, and dispute settlement procedures. [1]

A national treatment clause forbade discrimination between members own nationals and nationals of other members and a most-favoured-nations clause forbade discrimination between nationals of other members. However, nothing required all member countries to have identical protection. With regard to patents, the TRIPS agreement required member countries to make patents available for any invention and that protection under the patent shall not end before the expiration of a period of 20 years. [2] Developed countries were required to comply with all components of the TRIPS agreement by 1st January 1996. Developing countries were given a general 5 year period for compliance and so were required to comply with all the components of the TRIPS agreement by 1st January 2000, however, they were required to comply with the national treatment and most-favoured-nations clauses by 1st January 1996. A further concession was granted to developing countries in that compliance could be delayed until 2005 for those inventions which were not provided with patents at the time of the agreement. Basically, the TRIPS agreement is seen as an extension of intellectual property rights protection to developing countries that previously offered little or no protection.

Surprisingly little theoretical work has been undertaken on the impact of extending patent protection to developing countries. In a model in which innovation

[1] An excellent discussion of these events is found in Lanjouw and Cockburn (2000)
[2] Details of the TRIPS agreement can be found at WTO (2000).

occurs in developed countries and imitation occurs in developing countries, Deardorff (1992) finds that the developed countries are unambiguously better off following the extension of patent protection to the developing world because (i) more R&D is done and more new products developed and (ii) they receive monopoly profits from the developing world. On the other hand, the developing countries can be worse off following the extension of patent protection because (i) consumer surplus is shifted to the industrialized countries in the form of monopoly profit and (ii) this might outweigh the welfare gain from there being more products available. Deardorff models the extension of patent protection to the developing world as a move from a situation in which no protection is given to one in which complete protection is given.

In a similar vein, Chin and Grossman (1990), examine the welfare consequences of moving to complete patent protection in a duopoly model in which there is one innovating firm in the North and one imitating firm in the South. Their findings are similar to those of Deardorff (1992). A weakness of these approaches is that in neither is the question of optimal patent design addressed. It may or may not be optimal for developed countries to give full patent protection and developing countries to give none. Also, from a global perspective full protection may or may not be optimal.

In this paper, optimal patent design is considered. It might be optimal for some countries to have a lot of protection and others to have little. Rather than considering extending patent protection, the optimal global patent is determined and compared to patents that arise where countries optimally determine their own patent design. In doing this, not only can the optimal patents designs of each country be compared, but both can be compared to the optimal global patent design. In addition, the welfare implications of moving to the optimal global patent design can be examined.

The two aspects of patent design that are considered in this paper are patent breadth and patent length. What is meant by patent length is clear, by patent breadth is meant the extent to which a patent allows competition in the product market. Narrowing patent breadth means that there is more competition in the product market. This paper extends the work of Klemperer (1990), Gilbert and Shapiro (1990) and Wright (1999) to a two country setting in which innovation is undertaken in one country while imitation occurs in both. Given that innovation typically occurs in developed countries and developing countries typically only imitate, the results of this paper can be reinterpreted as applying to developed and developing countries. Unlike Denicolo (1996), where potential innovators are involved in a patent race, in this paper there is only one innovator.

Initially, optimal patent design is considered from the point of view of a policy maker in each country. It is useful to divide the patent design problem into two stages. In the first stage, the policy makers in each country maximize expected welfare by choosing the reward the innovator receives from operating in their market. The greater is the aggregate reward, the more R&D the innovator does and the greater is the probability of a successful innovation. In the second stage, the policy maker chooses patent breadth and patent length to maximize welfare or what is the same thing, to minimize the deadweight loss associated with the monopoly power created by the patent. Where the two countries have symmetric demands and costs, it is shown in Propositions 1a and 1b, that the policy maker in the innovating country has a greater incentive to reward the innovator than the policy maker in the non-innovating country. This translates into stronger patent protection in the innovating country, patents are either longer or broader depending on whether broad or narrow patents of infinite length are optimal. These result are consistent with the observation that in the absence of a TRIPS agreement patent protection is stronger in the developed (innovating) world. However, it must be noted, where the two countries have asymmetric demands and costs, that if the innovating country has smaller demand and/or greater costs, then there is no presumption that the innovating country has stronger patent protection, Proposition 2a. On the other hand, if the innovating country has greater demand and/or lower costs then unambiguously it has stronger patent protection than the non-innovating country. These results are consistent with those found in a very different model developed by Grossman and Lai (2001).

Next, the problem faced by a global policy maker is solved. The optimal aggregate reward is allocated between the two countries to minimize deadweight loss. Where the countries are symmetric or differ in costs or demand intercepts, Propositions 3a and 4a show that if the optimal global patent design is broad, then although the optimal global aggregate reward is unique, patent lengths in each country are not. An implication of this last result is that patents of identical breadth and length, in each country, are globally optimal as long as they induce the optimal aggregate reward. The global optimality of identical patent protection in both countries is reinforced in Proposition 3b where it is shown that if the optimal global patent design is narrow and of infinite length, then the degree of narrowness in both countries is identical. These results suggest that the globally optimal patent design involves identical patents in both countries, however, if countries differ in the curvature of their demand functions, then Proposition 5 shows that identical patent designs in each country are no longer globally optimal.

Comparing the optimal global aggregate reward to the rewards that were optimal from each countries' perspective reveals that the optimal global aggregate reward is greater than the sum of the individual country optimal rewards, Propositions 6a and 6b. This suggests that the patents designed by the individual countries are too weak from the global perspective. Under symmetry, the implication is that the optimal global patent design is either longer or broader for the non-innovating country as compared to a situation in which the non-innovating country chooses its patent design. This provides some theoretical support for the 1994 TRIPS agreement which increased protection of intellectual property rights in the developing (non-innovating) world.

Finally, welfare of each country is compared under the optimal individual country patent design and the optimal global patent design. It is found, under symmetry, that the innovating country is better off and that the non-innovating country can be worse off. The non-innovating country gains from the increase in R&D, but can lose because of the increase in patent length or breadth. This is consistent with the findings of Deardorff (1992) and Grossman and Lai (2001), though in very different models. [3] One implication of the welfare analysis is that for the developing (non-innovating) countries to continue to embrace the 1994 TRIPS agreement some mechanism that redistributes the gains from the optimal global patent design needs to be implemented.

2. Patent Breadth and Model Structure

There are two countries, N, the innovating country and S, the non-innovating country. In each country there is a policy maker or there is a global policy maker that chooses patent breadth and length to maximize expected country or expected global welfare. What is meant by patent length is clear, by patent breadth is meant the extent to which a patent allows competition in the product market. Narrower patents are associated with more competition in the product market, broader patents with less.

Gilbert and Shapiro (1990) measure patent breadth by innovator flow profits, narrower patents are associated with smaller flow profits. In a spatial model, Klemperer (1990) measures patent breadth by the distance between the patented product and the products other firms can produce without infringing the patent, narrower patents are associated with a smaller distance. Gallini (1992) measures patent breadth by the size of imitation costs, narrower patents are associated with smaller imitation costs. Denicolo (1996) parameterizes the measure used by Gilbert and

[3] However, unlike Deardorff, in this paper, by definition, any strengthening of patent protection in the developing (non-innovating) world that arises from the move to the optimal global patent design can not reduce global welfare. This is a direct consequence of working with optimal patent designs.

Shapiro (1990) in a way such that innovator flow profit is a decreasing function of the parameter, narrower patents are associated with a larger parameter value. Finally, O'Donoghue, Scotchmer, and Thisse (1998) measure patent breadth by the quality difference between the patented product and the products other firms can produce without infringing the patent, narrower patents are associated with a smaller difference.

All these measures exhibit the characteristic that narrower patents are associated with more competition in the product market. This is also true of the measure used in this paper, namely, the number of imitators that the patent allows into the market, m_N and m_S, for countries N and S, respectively. With this measure, narrower patents are associated with more imitators in the market. This approach to patent breadth is equivalent to a patent system in which the government of country $i = N, S$ awards a patent to the innovator and m_i additional licenses to competing firms. These m_i licensees are then placed on the same footing as the patentee and the industry within each country is characterized by imperfect competition between $m_i + 1$ identical firms. The imperfect competition means that the patentee and each imitator realize a positive flow of economic profit equal to Π_i which is a decreasing function of m_i. This hypothetical patent system creates a useful proxy for patent breadth - the index m_i, with larger $m_i's$ corresponding to narrower patents and smaller $m_i's$ corresponding to broader patents - and permits a relatively straightforward mathematical approach, though somewhat unrealistic, to the patent breadth issue.

The North is distinguished from the South by being the country in which the innovator is located. Imitators are located in the North and the South. Although the North and the South differ in other dimensions as well, for the purposes of this paper, in which the tightening of intellectual property rights in the South following the 1994 TRIPS Agreement is the main motivating factor, the location of the innovator is central. [4] Nevertheless, the model can be thought of more broadly as applying to

[4]Other differences between the North and the South are discussed in the paper, but do not play

any two countries which are distinguished by whether or not they have an innovating firm located within them.

Patents with breadth m_N and m_S and length T_N and T_S, ensure that a successful product innovator, located in country N, faces limited competition from imitators, in each country, for the length of the patent. After the patent expires, the innovation is freely available to any firm so the industry in each country becomes perfectly competitive. The patents generate a flow of profits, a reward, for the innovator. Given this reward, the single innovator chooses R&D to maximize expected profit. [5] It is assumed that the innovator sells the new product in both countries while imitators only sell in the country in which they are located.

Any policy maker's problem can be divided into two stages. In the first stage, the policy maker maximizes expected welfare by choosing the innovator's reward. In the second stage, the policy maker chooses patent breadth and length to maximize welfare or minimize the deadweight loss associated with the creation of the reward. In determining the optimal reward and the optimal patent breadth and length, the policy maker trades off the traditional deadweight loss associated with monopoly power against the benefits of a greater probability of a successful innovation caused by the positive effect a greater innovator reward has on R&D. As is usual, this two stage problem is solved backwards.

Finally, it is assumed that any policy maker commits to the optimal patent design obtained from the solution of the two stage problem. In the absence of such commitment, after the innovator has undertaken R&D, it would be optimal for any policy maker to reduce patent protection to zero. The innovator, realizing this, would undertake no R&D. Assuming commitment overcomes this problem and can be rationalized by appealing to the reputation effects that would be present in a model in which more than one innovation was considered.

a central role in comparing the independent country equilibrium to the global equilibrium.

[5] Unlike Denicolo (1996), where it is assumed that there are many potential innovators involved in a patent race, here there is only one potential innovator. Issues of the timing of innovation, Denicolo (1999) and Duffy (2002), and the sequential nature of innovation, Scotchmer and Green (1995) are not considered.

3. Country Policy Makers

In this section, each country has a policy maker that maximizes its expected welfare.

3.1. Stage Two - Optimal Patent Breadth and Length

Given some reward to be earned in each country, each country's policy maker chooses patent breadth and length to maximize welfare.

Country N

Welfare in country N is the sum of consumer surplus, country N industry profit, and the profit earned by the innovator in country S. Let consumer surplus as a function of the number of firms in country N be $C_N(m_N + 1)$, where m_N is the number of imitators in country N and let $\Pi_N(m_N+1)$ denote innovator and imitator profit as a function of the number of firms in country N. [6] For mathematical convenience it is assumed that m_N is a continuous variable. It is also assumed that consumer surplus is an increasing function of the number of firms in the industry and that firm profit is a decreasing function of the number of firms in the industry, that is, $C'_N(m_N + 1) > 0$ and $\Pi'_N(m_N + 1) < 0$ for all $(m_N + 1)$. Let C_N^c denote competitive consumer surplus, that is, consumer surplus after the patent expires and the industry is competitive and let r be the discount rate.

The policy maker's problem is to maximize welfare by choosing patent breadth and length subject to the innovator receiving a reward V_N from selling in country N and V_S from selling in country S. It is assumed that the innovation occurs at time $t = 0$ so the policy maker's problem is,

$$\begin{aligned} \max_{m_N,T_N} \quad W_N \quad &\equiv \int_0^{T_N} [C_N(m_N + 1) + (m_N + 1) \cdot \Pi_N(m_N + 1)] \exp^{-rt} dt \\ &+ \int_{T_N}^{\infty} C_N^c \exp^{-rt} dt + V_S \end{aligned} \tag{1}$$

subject to

[6] For a given number of imitators, innovator and imitator profit are assumed equal. Assumptions that guarantee this are given in the following section.

$$\int_0^{T_N} \Pi_N(m_N+1)\exp^{-rt}dt = V_N. \tag{2}$$

This constraint ensures a reward of V_N is earned by the innovator in country N. Define $\tilde{m}_N(V_N)$ by

$$\int_0^{\infty} \Pi_N(\tilde{m}_N)\exp^{-rt}dt \equiv V_N, \tag{3}$$

it is the maximum number of imitators that can enter the industry and still allow the innovator to earn reward V_N. In the Appendix, it is shown that $\tilde{m}'_N(V_N) < 0$. Now $0 \leq T_N \leq \infty$ so the solution to the policy maker's problem must involve $m_N \leq \tilde{m}_N$. Substituting (2) into (1) and rearranging yields

$$\max_{m_N} \quad W_N \equiv \frac{C_N^c}{r} + V_S + V_N - \alpha_N \cdot V_N \tag{4}$$

subject to

$$m_N \leq \tilde{m}_N(V_N), \tag{5}$$

where

$$\alpha_N = \frac{C_N^c - C_N(m_N+1)}{\Pi_N(m_N+1)} - m_N. \tag{6}$$

The ratio of deadweight loss to innovator profit is given by $\alpha_N - 1$, so maximizing (4) is equivalent to minimizing this ratio. Once patent breadth, m_N, is chosen to maximize welfare, patent length, T_N, is chosen to ensure constraint (2) is satisfied. Note that if constraint (5) binds, then $T_N = \infty$.

As long as the industry is not perfectly competitive, it seems reasonable to assume that the ratio of deadweight loss to innovator profit is greater than zero for all m_N so $\alpha_N - 1 > 0 \quad \forall \quad m_N$, that is,

$$\alpha_N > 1 \quad \forall \quad m_N. \tag{7}$$

Let the number of imitators that solves problem (4) - (5) be denoted $\hat{m}_N$ and let $\hat{\alpha}_N$ be (6) after $\hat{m}_N$ is substituted into it. $\hat{m}_N$ actually minimizes α_N and $\hat{\alpha}_N$ is the minimized value of α_N. If $\hat{m}_N = 0$, then the optimal patent is as broad as possible and all imitators are excluded from the industry. On the other hand, if $\hat{m}_N = \tilde{m}_N(V_N)$, then the optimal patent is of infinite length and as narrow as possible, that is, as many imitators are allowed into the industry as is consistent with the innovator receiving reward V_N.

Maximized welfare is

$$\hat{W}_N = \frac{C_N^c}{r} + V_S + V_N - \hat{\alpha}_N(V_N) \cdot V_N. \tag{8}$$

If the constraint on the number of imitators, (5), does not bind, then $\hat{\alpha}_N$ is a constant. On the other hand, if the constraint on the number of imitators binds, then the optimal patent is as narrow as possible and $\hat{\alpha}_N$ is a function of V_N, that is, $\hat{\alpha}_N(V_N)$. At this stage it is useful to sign $\hat{\alpha}'_N(V_N)$. As constraint (5) binds, the policy maker would like more imitators in the market but is constrained by the innovator's reward. Therefore, the case being considered is where α_N is a decreasing function of m_N. An increase in V_N tightens constraint (5), reduces the number of imitators, and so increases the value of α_N. Therefore, $\hat{\alpha}'_N(V_N) > 0$ as shown in the Appendix.

Country S

Welfare in country S is the sum of consumer surplus and aggregate profit earned by country S imitators. Innovator profit earned in S is part of country N welfare. The policy maker's problem is similar to that in country N, namely,

$$\max_{m_S, T_S} \quad W_S \equiv \int_0^{T_S} [C_S(m_S+1) + m_S \cdot \Pi_S(m_S+1)] \exp^{-rt} dt + \int_{T_S}^{\infty} C_S^c \exp^{-rt} dt \tag{9}$$

subject to

$$\int_0^{T_S} \Pi_S(m_S+1) \exp^{-rt} dt = V_S. \tag{10}$$

Rearranging yields

$$\max_{m_S} \quad W_S \equiv \frac{C_S^c}{r} - \alpha_S \cdot V_S, \tag{11}$$

subject to

$$m_S \leq \tilde{m}_S(V_S), \tag{12}$$

where $\tilde{m}_S$ is defined by

$$\int_0^{\infty} \Pi_S(\tilde{m}_S) \exp^{-rt} dt \equiv V_S \tag{13}$$

and

$$\alpha_S = \frac{C_S^c - C_S(m_S + 1)}{\Pi_S(m_S + 1)} - m_S. \tag{14}$$

As for country N, for country S, where constraint (12) binds, $\tilde{m}'_S(V_S) < 0$ and

$$\alpha_S > 1 \quad \forall \quad m_S. \tag{15}$$

Let the number of imitators that solves problem (11) - (12) be denoted $\hat{m}_S$ and let $\hat{\alpha}_S$ be (14) after $\hat{m}_S$ is substituted into it. Maximized welfare is

$$\hat{W}_S = \frac{C_S^c}{r} - \hat{\alpha}_S(V_S) \cdot V_S. \tag{16}$$

If the constraint on the number of imitators does not bind, then $\hat{\alpha}_S$ is a constant. On the other hand, if the constraint on the number of imitators binds, then the optimal patent is as narrow as possible and $\hat{\alpha}_S$ is a function of V_S, that is, $\hat{\alpha}_S(V_S)$. Using an identical argument to that used for country N, it is clear that if constraint (12) binds, then $\hat{\alpha}'_S(V_S) > 0$.

3.2. Stage One - Optimal Rewards

In this stage, the policy makers choose the rewards to maximize expected welfare taking into account the effect their choice of rewards has on the innovator's R&D.

3.2.1. The Innovator

The innovator takes the rewards, V_N and V_S, as given and chooses R&D to maximize expected profit. This R&D occurs at time $t = 0$ and immediately results in an innovation which immediately receives patent protection. [7] Let $\rho(R)$ denote the probability, as a function of R&D, that the innovator is successful in inventing a new product. It is assumed that $\rho'(R) > 0$ and $\rho''(R) < 0$ so that more R&D increases the probability of a successful innovation though it increases at a decreasing rate. The innovator's problem is

$$\max_R \ \rho(R) \cdot (V_N + V_S) - \omega R, \tag{17}$$

[7] As in Nordhaus (1969), the model is static in the sense that the size of the reward has no effect on the timing of the innovation. Two recent papers that allow the size of the reward to affect the timing of the innovation are Duffy (2002) and Denicolo (1999).

where ω is the unit cost of R&D. The first order condition for a maximum is

$$\rho'(R) \cdot (V_N + V_S) - \omega = 0. \tag{18}$$

In the appendix, it is shown that $R'(V_N + V_S) > 0$ and that $R''(V_N + V_S) < 0$. [8] These results have intuitive appeal, a bigger reward, earned in either country, increases the amount of R&D undertaken, though at a decreasing rate.

3.2.2. The Policy Maker in Country N

Given the reward obtained in country S, the country N policy maker chooses V_N to maximize expected welfare taking into account the effect its choice of the reward has on innovator R&D. Its problem is

$$\max_{V_N} \quad E\hat{W}_N \equiv \rho(R(V_N+V_S)) \cdot [C_N^c/r + (1-\hat{\alpha}_N(V_N)) \cdot V_N + V_S] - \omega R(V_N+V_S) \tag{19}$$

with first order condition

$$\begin{aligned} \frac{dE\hat{W}_N}{dV_N} &= (1 - \hat{\alpha}_N(V_N) - \hat{\alpha}'_N(V_N) \cdot V_N) \cdot \rho(R(V_N + V_S)) - \omega R'(V_N + V_S) \\ &+ [C_N^c/r + (1 - \hat{\alpha}_N(V_N)) \cdot V_N + V_S] \cdot \rho'(R(V_N + V_S)) \cdot R'(V_N + V_S) \\ &= 0 \end{aligned} \tag{20}$$

for an interior solution. The trade off the policy maker faces when increasing V_N is between decreasing expected welfare via increasing deadweight loss and increasing R&D costs, the first and second terms on the left hand side of (20) and increasing expected welfare via increasing the probability of a successful innovation, the third term on the left hand side of (20). For the problem to make economic sense the term $(1 - \hat{\alpha}_N(V_N) - \hat{\alpha}'_N(V_N) \cdot V_N)$ must be negative overall otherwise an infinite reward would be optimal.

The second order condition for a maximum is assumed to be satisfied and the solution is assumed to be unique. The condition under which the second order condition is satisfied is given in the Appendix.

[8] In signing $R''(V_N + V_S)$ it is assumed that $\rho'''(\cdot) = 0$.

3.2.3. The Policy Maker in Country S

The problem faced by the policy maker in country S is very similar to that faced by the policy maker in country N and is given by

$$\max_{V_S} \quad E\hat{W}_S \equiv \rho(R(V_N + V_S)) \cdot [C_S^c/r - \hat{\alpha}_S(V_S) \cdot V_S] \tag{21}$$

with first order condition

$$\begin{aligned} \frac{dE\hat{W}_S}{dV_S} &= (-\hat{\alpha}_S(V_S) - \hat{\alpha}'_S(V_S) \cdot V_S) \cdot \rho(R(V_N + V_S)) \\ &+ [C_S^c/r - \hat{\alpha}_S(V_S) \cdot V_S] \cdot \rho'(R(V_N + V_S)) \cdot R'(V_N + V_S) = 0 \end{aligned} \tag{22}$$

for an interior solution. The trade off faced by the policy maker in country S is similar to that faced by the policy maker in country N. The difference arises because it is assumed that there is a single innovator located in country N, that is, only one country N firm has the ability to innovate in the industry under consideration. [9] Finally, it is assumed that the second order condition for a maximum is satisfied, and that the solution is unique.

3.2.4. Nash Equilibrium

The Nash equilibrium rewards are obtained by simultaneously solving (20) and (22). Since (i) the sets of rewards, V_N and V_S, are compact and convex, (ii) $E\hat{W}_N$ and $E\hat{W}_S$ are both continuous in both V_N and V_S, (iii) $E\hat{W}_N$ is concave in V_N, and (iv) $E\hat{W}_S$ is concave in V_S, a pure strategy Nash equilibrium exists. [10] It is further assumed that the Nash equilibrium is unique and stable. In the Appendix, it is shown that the equilibrium locus (reaction function) of the policy maker in country S is negatively sloped while the equilibrium locus of the policy maker in country N can be positively or negatively sloped. These two possibilities are shown in Figures 1a and 1b. In Figure 1a the equilibrium locus of the policy maker in country N is drawn flatter than that of the country S policy maker to ensure stability. Let

[9] If the innovation market in country N was competitive and free entry ensured zero expected profits for innovators, then each country's policy maker would face an identical problem.

[10] Dasgupta and Maskin (1986, p4)

the Nash equilibrium rewards be denoted by $\hat{V}_N$ and $\hat{V}_S$ for countries N and S, respectively.

To obtain some insight into the characteristics of the Nash equilibrium the model is made less general by assuming a particular market structure. It is assumed that one patent holder and m_N entrants in country N and m_S entrants in country S produce perfect non-patent infringing substitutes at a constant marginal cost of c_N and c_S in countries N and S respectively. [11] It is assumed that entry is simultaneous. The equilibrium concept adopted is Cournot equilibrium. In this equilibrium, in each country, entrants and patent holders earn the same flow profit. Importantly, in addition, it is assumed that demand in each country takes the following form,

$$p_i = a_i - bQ_i^{d_i} \quad i = N, S \tag{23}$$

where p_i is the per-unit price if Q_i units of the good are sold and a_i, b and d_i are positive parameters. This demand specification includes, convex ($d_i < 1$), linear ($d_i = 1$), and concave ($d_i > 1$) demand curves.

Given this structure, Wright (1999) demonstrates that

$$\alpha_i(m_i) = \frac{\left(\frac{1+d_i+m_i}{1+d_i}\right)\left(\frac{(1+m_i)^{1-d_i}}{1+d_i+m_i}\right)^{\frac{1}{d_i}}}{-}\left(\frac{1+m_i}{1+d_i}\right) \cdot (2 + d_i + m_i) + 1 \qquad i = N, S. \tag{24}$$

Note that α_i is independent of a_i, b and c_i. Wright (1999) also shows that $\alpha_i(m_i)$ is a monotonically increasing (decreasing) function if $d_i > 1$ ($d_i < 1$) and is a constant equal to 1.5 if $d_i = 1$. It turns out that $1 < \alpha_i(m_i) < 2$. Since the policy maker acts to minimize α_i in the second stage, $\hat{m}_N = 0 = \hat{m}_S$ if $d_i > 1$ and $\hat{m}_N = \tilde{m}_N(V_N)$ and $\hat{m}_S = \tilde{m}_S(V_S)$ if $d_i < 1$. The above is summarized in the following .

Lemma 1: *If demand is concave ($d_i > 1$), then the optimal patent design specifies a patent of maximum breadth. If demand is convex ($d_i < 1$), then the optimal patent design specifies a patent of minimum breadth, but infinite length.*

[11] This set up was used in Gallini (1992) and Wright (1999). Another interpretation is that the innovator patents a process for producing a homogeneous product at marginal cost c_N and c_S in countries N and S, respectively, and this product can only be produced using this process. Imitators imitate the process.

The intuition is clear since $\alpha_i - 1$ is the ratio of deadweight loss to innovator profit. With linear demand, $d_i = 1$, simple geometry reveals that this ratio always equals $\frac{1}{2}$. An increase in the number of imitators decreases deadweight loss and innovator profit in the same proportion. In this case any patent breadth and length consistent with constraints (2) and (10) are optimal. However, if demand is concave, $d_i > 1$, an increase in the number of imitators decreases deadweight loss by a smaller proportion than it reduces innovator profit because industry output rises a little, but price falls a lot as does innovator output. As a result, the ratio of deadweight loss to innovator profit is increasing in the number of imitators and so is minimised by having zero imitators, that is, an extremely broad patent. A similar argument establishes that narrow patents are optimal if $d_i < 1$. [12]

The following propositions give some insight into the characteristics of the Nash equilibrium. Propositions 1a and 1b show that patent protection is stronger in the innovating country, country N, than in the non-innovating country, country S.

Proposition 1a: Symmetric Costs and Demand Intercepts - Broad Patent *If $a_N - c_N = a_S - c_S$ and $d_N = d_S > 1$ so that $C_N^c = C_S^c$, $\hat{m}_N = \hat{m}_S = 0$, and $\hat{\alpha}_N = \hat{\alpha}_S$, then in a Nash Equilibrium $\hat{V}_N > \hat{V}_S \geq 0$ and the optimal patent has greater length in country N than in country S, that is, $\hat{T}_N > \hat{T}_S \geq 0$.*

Proof: (i) $V_N(V_S)$ downward sloping. Suppose condition (22) is satisfied, and that $V_N = V_S$. Let the solution of these two conditions be given by $V_S^a = V_N^a$. At this reward pair, $\frac{dE\hat{W}_N}{dV_N} = \rho((V_N + V_S)) > 0$. By the second order condition for a maximum, the best response to V_S^a is $V_N > V_N^a$. This best response is given by V_N^b in Figure 1a. As $V_N(V_S)$ is downward sloping and flatter than $V_S^0(V_N)$ by stability, the Nash equilibrium is at $\hat{V}_N > \hat{V}_S$ in Figure 1a. By conditions (2) and (10), $\hat{T}_N > \hat{T}_S$. Nothing rules out $\hat{V}_S = 0$, this would be the outcome if the vertical intercept of $V_S(V_N)$ lies below the vertical intercept of $V_N(V_S)$.

(ii) $V_N(V_S)$ upward sloping. The proof is identical to (i) above except Figure 1b

[12] More detailed intuition for Lemma 1 can be found in Wright (1999).

is used. **(Q.E.D.)**

This proposition deals with the case where the optimal patent is as broad as possible. It has intuitive appeal, an increase in V_N is not as costly to country N as an increase in V_S is to country S as welfare of country N includes innovator profit earned in country S. In addition, an increases in V_N yields more expected welfare to country N than an increase in V_S yields to country S as the increased probability of a successful innovation applies to innovator profit earned in country S. This makes increases in V_N more beneficial to country N than increases in V_S are to country S. This is reflected in the equilibrium rewards. Via constraints (2) and (10) the greater equilibrium reward in country N translates into a greater equilibrium patent length in country N. Note that it might be optimal for country S to give no patent protection, $\hat{V}_S = 0$ and $\hat{T}_S = 0$.

Proposition 1b: Symmetric Costs and Demand Intercepts - Narrow Patent
If $a_N - c_N = a_S - c_S$ *and* $d_N = d_S < 1$ *so that* $C_N^c = C_S^c$, $\hat{m}_N = \tilde{m}_N(V_N)$, *and* $\hat{m}_S = \tilde{m}_S(V_S)$, *then in a Nash Equilibrium* $\hat{V}_N > \hat{V}_S \geq 0$ *and the optimal patent has greater breadth in country* N *than in country* S, *that is,* $\hat{m}_N < \hat{m}_S$. *In addition,* $\hat{\alpha}_N > \hat{\alpha}_S$.

Proof: Proposition 1b is qualitatively the same as Proposition 1a and its proof is identical in structure to that of Proposition 1a. The last inequality follows because $\alpha_i(m_i)$ is a decreasing function of m_i. **(Q.E.D.**

The intuition for Proposition 1b is identical to that of Proposition 1a, note however, where patent protection is given, that patent lengths are infinite. The following proposition, Proposition 2a, and the discussion that follows shows that there are conditions under which patent protection is not necessarily stronger in the innovating country than in the non-innovating country.

Proposition 2a: Asymmetric Costs and Demand Intercepts - Broad Patent
If $a_N - c_N < a_S - c_S$ *and* $d_N = d_S > 1$ *so that* $C_N^c < C_S^c$, $\hat{m}_N = \hat{m}_S = 0$ *and* $\hat{\alpha}_N = \hat{\alpha}_S$, *then in a Nash Equilibrium* $\hat{V}_N \geq \hat{V}_S \geq 0$ *or* $0 \leq \hat{V}_N < \hat{V}_S$ *depending*

on the size of the difference between $a_N - c_N$ and $a_S - c_S$. Define $V_S^{crit} > \hat{V}_N$ by $\frac{V_S^{crit}}{\Pi_S(1)} = \frac{\hat{V}_N}{\Pi_N(1)}$. If $\hat{V}_S < (=, >) V_S^{crit}$, then $\hat{T}_S < (=, >) \hat{T}_N$.

Proof: Note that $\alpha_i(m_i)$ does not depend on marginal cost or the demand intercept. It will suffice to consider the case where initially $a_N - c_N = a_S - c_S$ and $a_S - c_S$ is increased so that $a_N - c_N < a_S - c_S$. Examination of (20) reveals that the increase in $a_S - c_S$ does not effect the equilibrium locus of country N. On the other hand, examination of (22) reveals that the increase in $a_S - c_S$ effects the equilibrium locus of country S through an increase in $\frac{C_S^c}{r}$. For a given V_N, the increase in $a_S - c_S$ makes $\frac{dE\hat{W}_S}{dV_S}$ more positive and so by the second order condition for a maximum shifts the country S equilibrium locus to the right. Such a shift is shown in Figures 1a and 1b, where the country S equilibrium locus shifts from V_S^0 to V_S^1. The bigger the increase in $a_S - c_S$, the bigger the shift to the right. As drawn, the shifts in Figures 1a and 1b result in a new equilibrium in which $\hat{V}_N = \hat{V}_S$, however, depending on the size of the rightward shift (the size of the increase in $a_S - c_S$) any relationship between $\hat{V}_N$ and $\hat{V}_S$ is possible.

Now consider patent length. In the Appendix, it is shown that T_S and T_N are the same monotonically increasing function of $\frac{V_S}{\Pi_S(1)}$ and $\frac{V_N}{\Pi_N(1)}$, respectively. Now $\Pi_S(1) > \Pi_N(1)$ because $a_S - c_S > a_N - c_N$, so $V_S^{crit} > \hat{V}_N$. By monotonicity, if $\hat{V}_S < (=, >) V_S^{crit}$, then $\hat{T}_S < (=, >) \hat{T}_N$.

(Q.E.D.)

In Proposition 1a, under symmetric costs and demands, it was shown that the policy maker in country N had a larger incentive than the policy maker in country S to increase its reward to the innovator because it received profit generated by the innovator in country S. Proposition 2a demonstrates that this is not generally true. With $a_N - c_N < a_S - c_S$, competitive consumer surplus is greater in country S than country N and so on this count the welfare payoff to an increased reward is greater in country S than country N. If the difference between $a_N - c_N$ and $a_S - c_S$ is large enough, the competitive consumer surplus increase in country S more than offsets the effect of the remitted profit to country N so that overall the policy maker in country S has a greater incentive to increase the innovator's reward than the policy maker in country N. For example, if country S has a significant cost advantage over country N, $c_S << c_N$, then the innovation is worth more in country S than N, therefore, the country S policy maker has a greater incentive to encourage R&D through patent protection. If $a_N - c_N > a_S - c_S$, then the remitted profit effect and the consumer surplus effect reinforce each other and $\hat{V}_N > \hat{V}_S$. Therefore, if the market in country N is larger than in country S, $a_N > a_S$, and/or marginal cost is less in country N than in country S, then country N has even stronger patent protection relative to country S than in the case of symmetry.

In a recent paper, Grossman and Lai (2001), develop a differentiated products general equilibrium model with deterministic R&D and examine optimal patent length. They found that, in equilibrium, the optimal patent length in the North is greater than in the South as long as the North is at least as large as the South and inventive capacity in the North is much greater than in the South. These results are entirely consistent with Propositions 1a-2a above as the North is assumed to be the only country with inventive capacity. The intuition for the Grossman and Lai result is similar to that of this paper even though the two models are quite different.

Symmetric Costs and Asymmetric Demand Parameters, d_i: First of all, consider the case where demand is such that a broad patent is optimal in both countries, $d_i > 1 \ for \ i = N, S$. Now assume $d_N > d_S$ so $\hat{\alpha}_N < \hat{\alpha}_S$ and $\frac{C_N^c}{r} < \frac{C_S^c}{r}$. Using a similar argument to that used in the proof of Proposition 1a leads to an ambiguous relationship between $\hat{V}_N$ and $\hat{V}_S$. Although the cost of V_N in terms of deadweight loss is reduced there is also a decrease in $\frac{C_N^c}{r}$ so that overall $\frac{dE\hat{W}_N}{dV_N}$ can not be signed at $V_S^a = V_N^a$. As a consequence, the relationship between $\hat{T}_N$ and $\hat{T}_S$ is ambiguous. Similar arguments can be applied to the case where narrow patents are optimal so that the relationships between $\hat{V}_N$ and $\hat{V}_S$ and $\hat{T}_N$ and $\hat{T}_S$ are ambiguous.

Summary: Under symmetry, country N has a greater incentive to reward the innovator than country S because the innovator is located in country N. If broad patents are optimal, then this greater reward is achieved through longer patents in N than S. On the other hand, if narrow patents of infinite length are optimal, then this greater reward is achieved by having less imitators in N than S, that is, by having a broader patent in country N than S.

Under asymmetry, the greater incentive country N has to reward the innovator because the innovator is located in country N can be offset by the impact differences in cost and demand parameters have on competitive consumer surplus in each country. As a result, the relationship between the Nash equilibrium rewards in each country is in general ambiguous. This ambiguity in rewards translates into an ambiguity concerning patent lengths, where broad patents are optimal, and patent breadths, where narrow patents of infinite length are optimal.

4. Global Policy Maker

In this section, it is assumed that a global policy maker chooses V_N and V_S to maximize expected global welfare, where global welfare is the sum of welfare in countries N and S. The global policy maker's problem is

$$\begin{aligned} \max_{V_N, V_S} \quad E\hat{W}_G \ &\equiv\ [\frac{C_N^c}{r} + (1 - \hat{\alpha}_N(V_N)) \cdot V_N + \frac{C_S^c}{r} + (1 - \hat{\alpha}_S(V_S)) \cdot V_S] \cdot \rho(R(V_N + V_S)) \\ &-\ \omega R(V_N + V_S) \end{aligned} \tag{25}$$

with first order conditions

$$\begin{aligned} \frac{dE\hat{W}_G}{dV_N} \ &=\ (1 - \hat{\alpha}_N(V_N) - \hat{\alpha}'_N(V_N) \cdot V_N) \cdot \rho(R(V_N + V_S)) - \omega R'(V_N + V_S) \\ &+\ [\frac{C_N^c}{r} + (1 - \hat{\alpha}_N(V_N)) \cdot V_N \\ &+\ \frac{C_S^c}{r} + (1 - \hat{\alpha}_S(V_S)) \cdot V_S] \cdot \rho'(R(V_N + V_S))R'(V_N + V_S) = 0 \end{aligned} \tag{26}$$

and

$$\begin{aligned}\frac{dE\hat{W}_G}{dV_S} &= (1-\hat{\alpha}_S(V_S)-\hat{\alpha}_S'(V_S)\cdot V_S)\cdot\rho(R(V_N+V_S))-\omega R'(V_N+V_S)\\ &+ [\frac{C_N^c}{r}+(1-\hat{\alpha}_N(V_N))\cdot V_N\\ &+ \frac{C_S^c}{r}+(1-\hat{\alpha}_S(V_S))\cdot V_S]\cdot\rho'(R(V_N+V_S))R'(V_N+V_S)=0 \qquad (27)\end{aligned}$$

for an interior solution. The second order conditions for a maximum are assumed to be satisfied and the solution is assumed to be unique. The equilibrium rewards are obtained by simultaneously solving (26) and (27).

The following propositions give some insight into the characteristics of this equilibrium. Propositions 3a, 3b, and 4a establish that the optimal global patent design may involve patents with identical breadth and length in both the developed and developing countries.

Proposition 3a: Symmetric Costs and Demand Intercepts - Broad Patent *If $a_N - c_N = a_S - c_S$ and $d_N = d_S > 1$ so that $C_N^c = C_S^c$, $\hat{m}_N = \hat{m}_S = 0$, and $\hat{\alpha}_N = \hat{\alpha}_S = \hat{\alpha}$, then the optimal global aggregate reward $\bar{V} = \bar{V}_N + \bar{V}_S$ is unique, but its division between countries N and S is not. Any division of the aggregate reward between countries N and S maximizes global welfare. If $\bar{V}_N > (=,<)\bar{V}_S$ in this division, then $\bar{T}_N > (=,<)\bar{T}_S$.*

Proof: If $\hat{\alpha}_N = \hat{\alpha}_S = \hat{\alpha}$, then conditions (26) and (27) are identical and can only be solved for $\bar{V} = \bar{V}_N + \bar{V}_S$. The result on patent lengths just follows from constraints (2) and (10).

The intuition is clear. With identical demand and costs in the two countries and so identical patent breadths, the deadweight loss per unit of profit is the same in both countries. Therefore, from the global policy maker's perspective it does not matter, in terms of deadweight loss, in which country the innovator's reward is earned. All that matters is that the aggregate reward optimally trades off decreases in expected welfare caused by increases in deadweight loss against increases in expected welfare caused by increases in the probability of a successful innovation. As the division of the optimal aggregate reward, $\bar{V}$, between the two countries is not unique neither are the optimal patent lengths. Once again this result is consistent with Grossman and Lai (2001 p.32) where it was found that "the same global reward for innovation can be achieved with different combinations of the two patent lengths."

Proposition 3b: Symmetric Costs and Demand Intercepts - Narrow Patent
If $a_N - c_N = a_S - c_S$ and $d_N = d_S < 1$ so that $C_N^c = C_S^c$, $\hat{m}_N = \tilde{m}_N(V_N)$, and $\hat{m}_S = \tilde{m}_S(V_S)$, then the optimal global rewards are such that $\bar{V}_N = \bar{V}_S$. As a result, $\hat{m}_N = \hat{m}_S$, and $\hat{\alpha}_N = \hat{\alpha}_S$.

Proof: Substitute $V_S = V_N$ into (26) and solve for $\bar{V}_N$. Since (26) and (27) are identical at $V_S = V_N$, this $\bar{V}_N = \bar{V}_S$ also satisfies (27). As the solution to (26) and (27) is unique, the solution is $\bar{V}_N = \bar{V}_S$. The remaining results in the proposition follow from constraints (2) and (10).

Once again, the intuition is clear. With identical demands and marginal costs, $\hat{\alpha}_N(V_N)$ and $\hat{\alpha}_S(V_S)$ are the same increasing function of the rewards. Therefore, deadweight loss per-unit of profit is the same increasing function of the reward in each country. For a given aggregate reward $V_N + V_S$ and so a given incentive for R&D, deadweight loss is minimized by having $V_N = V_S$. If $V_N > V_S$, it is more costly in terms of deadweight loss to raise the reward in country N than S. Therefore, shifting some of the reward from country N to country S reduces deadweight loss. This was not true in Proposition 3a because there $\hat{\alpha}_N = \hat{\alpha}_S$ and did not depend on V_N or V_S. Since $\bar{V}_N = \bar{V}_S$, symmetry ensures patents of equal narrowness in both the developed and developing country, that is, $\hat{m}_N = \hat{m}_S$.

Proposition 4a: Asymmetric Costs and Demand Intercepts - Broad Patent
If $a_N - c_N < a_S - c_S$ and $d_N = d_S > 1$ so that $C_N^c < C_S^c$, $\hat{m}_N = \hat{m}_S = 0$ and $\hat{\alpha}_N = \hat{\alpha}_S$, then the optimal global aggregate reward $\bar{V} = \bar{V}_N + \bar{V}_S$ is unique, but its division between countries N and S is not. Any division of the aggregate reward between countries N and S maximizes global welfare. If $\bar{V}_N > (=,<)\bar{V}_S$ in this division, then $\bar{T}_N > (=,<)\bar{T}_S$.

Proof: Since $\hat{\alpha}_N$ and $\hat{\alpha}_S$ are independent of $a_N - c_N$ and $a_S - c_S$ respectively, the proof is identical to the proof of Proposition 3a.

The intuition for this proposition is identical to that of Proposition 3a. Proposition 5 below demonstrates that there are conditions under which identical patents designs in countries N and S are not optimal. First of all, consider the case where demand is such that a broad patent is optimal in both countries, $d_i > 1 \; for \; i = N, S$, but $\hat{\alpha}_N \neq \hat{\alpha}_S$. In this case, the global policy maker would raise as much of the optimal aggregate reward as possible in the country with the lower $\hat{\alpha}$ because the cost of the reward in terms of deadweight loss is lower there. It would achieve this by making patent length in the country with the lower $\hat{\alpha}$ as large as needed. If a patent of infinite length in the country with the lower $\hat{\alpha}$ is not sufficient to raise the optimal aggregate reward, then the difference is raised in the other country.

Second, consider the case where demand is such that a narrow patent is optimal in both countries, $d_i < 1 \; for \; i = N, S$, but $\alpha_N(\tilde{m}_N) > (<) \alpha_S(\tilde{m}_S) \; \forall \; \tilde{m}_N = \tilde{m}_S$. In this case, an interior solution would have the optimal aggregate reward allocated between the two countries so that $\hat{\alpha}_N = \alpha_N(\tilde{m}_N) = \alpha_S(\tilde{m}_S) = \hat{\alpha}_S$. If the number of imitators in each country was such that $\alpha_N > \alpha_S$, then the aggregate reward could be raised with lower cost if more of the reward was allocated to country S and less to country N. This involves reducing the number of imitators in country S and increasing the number of imitators in country N. Since $\alpha_N(\tilde{m}_N)$ and $\alpha_S(\tilde{m}_S)$ are decreasing functions of the number of imitators, the reallocation increases α_S and decreases α_N. This reallocation continues until $\alpha_S = \alpha_N$. With the optimal allocation of the reward, the country with the lower $\alpha(\tilde{m})$ function will have less imitators, that is, it will have a broader patent. The relationship between $\hat{V}_N$ and $\hat{V}_S$ is ambiguous. The discussion above is summarized in the following proposition.

Proposition 5: *If $d_i > 1, \; i = N, S$, so that broad patents are globally optimal, and $\hat{\alpha}_N \neq \hat{\alpha}_S$, then optimal patent length is greater in the country with the lower welfare cost per-unit of profit. If $d_i < 1, \; i = N, S$, so that narrow patents of infinite length* *are globally optimal, and* $\alpha_N(\tilde{m}_N) \neq \alpha_S(\tilde{m}_S)\ \forall\ \tilde{m}_N = \tilde{m}_S$*, then optimal patent breadth is greater in the country with the lower welfare cost per-unit of profit.*

The problem faced by the global policy maker can be viewed as a two-stage problem. In the first stage, the policy maker determines the aggregate global reward and in the second stage allocates this reward between the two countries to minimise deadweight loss by choosing patent design in each country. The second stage problem can be reinterpreted in the familiar multi-product Ramsey pricing framework to provide further intuition. The global policy maker chooses patent design in each country (rather than a regulator choosing the price of two products) to minimise deadweight loss, subject to the constraint that the innovator receives the globally optimal aggregate reward (rather than the regulated firm breaking even). If narrow patents are optimal, the global policy maker, following the Ramsey approach, chooses m_N and m_S, to equalise the ratio of deadweight loss to innovator profit in each country. On the other hand, if broad patents are optimal, the global policy maker is indifferent to the allocation of rewards between the two countries where $\hat{\alpha}_N = \hat{\alpha}_S$ or, where $\hat{\alpha}_N \neq \hat{\alpha}_S$ the global policy maker collects as much of the reward as possible from the country with the lower $\hat{\alpha}_i$.

5. Comparison of Country Equilibrium and Global Equilibrium

It this section, the results for the situation where each country chooses its reward independently will be compared to the situation in which a global policy maker chooses the rewards in each country. Propositions 6a and 6b and the discussion that follows them shows that the optimal global aggregate reward is greater than the sum of the individual country optimal rewards. As a result, the patents designed by the individual countries are too weak from the global point of view.

Proposition 6a: Symmetric Costs and Demand Intercepts - Broad Patent

(i) If $a_N - c_N = a_S - c_S$ *and* $d_N = d_S > 1$ *so that* $C^c_N = C^c_S$*,* $\hat{m}_N = \hat{m}_S = 0$*, and*

$\hat{\alpha}_N = \hat{\alpha}_S = \hat{\alpha}$, then the aggregate reward chosen by the global policy maker is greater than the sum of the rewards chosen by each country's policy maker in the Nash equilibrium, that is, $\bar{V} > \hat{V}_N + \hat{V}_S$. (ii) If $\bar{V}_N = \bar{V}_S$ and so $\bar{T}_N = \bar{T}_S$, then $\bar{V}_S > \hat{V}_S$ and so $\bar{T}_S > \hat{T}_S$.

Proof: (i) Adding (20) and (22), evaluated at $\hat{V}_N$, $\hat{V}_S$, and noting that $\hat{\alpha}'(V_N) = \hat{\alpha}'(V_S) = 0$ since $\hat{\alpha}_N = \hat{\alpha}_S = \hat{\alpha}$ is a constant yields

$$\begin{aligned} & (1-\hat{\alpha}) \cdot \rho(R(\hat{V}_N + \hat{V}_S)) - \omega R'(V_N + V_S) \\ + \quad & [\frac{C_N^c}{r} + \frac{C_S^c}{r} + (1-\hat{\alpha}) \cdot (\hat{V}_N + \hat{V}_S)]\rho'(R(\hat{V}_N + \hat{V}_S)) \cdot R'(\hat{V}_N + \hat{V}_S) \\ \equiv \quad & \hat{\alpha} \cdot \rho(R(\hat{V}_N + \hat{V}_S)) \end{aligned} \tag{28}$$

Given $\hat{\alpha}_N = \hat{\alpha}_S = \hat{\alpha}$ is a constant, the left hand sides of (26) and (27) are identical and are written as

$$\begin{aligned} \frac{dE\hat{W}_G}{d(V_N + V_S)} \quad = \quad & (1-\hat{\alpha}) \cdot \rho(R(V_N + V_S)) - \omega R'(V_N + V_S) \qquad (29) \\ + \quad & [\frac{C_N^c}{r} + \frac{C_S^c}{r} + (1-\hat{\alpha}) \cdot (V_N + V_S)]\rho'(R(V_N + V_S)) \cdot R'(V_N + V_S) \end{aligned}$$

Evaluating (29) at $\hat{V}_N, \hat{V}_S$ and using (28) yields

$$\frac{dE\hat{W}_G}{d(V_N + V_S)} = \hat{\alpha} \cdot \rho(R(\hat{V}_N + \hat{V}_S)) > 0. \tag{30}$$

So at $(\hat{V}_N + \hat{V}_S)$ expected global welfare can be increased by increasing the aggregate reward. Therefore, $\bar{V} > \hat{V}_N + \hat{V}_S$. (ii) From Proposition 1a, $\hat{V}_N > \hat{V}_S$. Given $\bar{V}_N = \bar{V}_S$ and the first part of this Proposition, then $\bar{V}_S > \hat{V}_S$ and so from constraint (10) $\bar{T}_S > \hat{T}_S$.

When making its reward decision, the policy maker in country N does not take into account the fact that an increase in V_N increases expected welfare in country S. Similarly, the policy maker in country S does not take into account the fact that an increase in V_S increases expected welfare in country N. The global policy maker internalizes these externalities and so gives the innovator a greater aggregate reward than would be the case if rewards were chosen at the country level. Where the optimal global patent design involves patents of equal length, this greater aggregate reward translates into longer patents in the non-innovating country compared to under the individual country optimal policy. Unfortunately, in this case, nothing can be said about the relationship between optimal global rewards and optimal individual country rewards, and so patent lengths, in the innovating country.

Proposition 6b: Symmetric Costs and Demand Intercepts - Narrow Patent *(i) If $a_N - c_N = a_S - c_S$ and $d_N = d_S < 1$ so that $C_N^c = C_S^c$, $\hat{m}_N = \tilde{m}_N(V_N)$, and $\hat{m}_S = \tilde{m}_S(V_S)$ then the aggregate reward chosen by the global policy maker is greater than the sum of the rewards chosen by each country's policy maker in the Nash equilibrium, that is, $\bar{V} > \hat{V}_N + \hat{V}_S$. (ii) Given Proposition 3b, that is, $\bar{V}_N = \bar{V}_S$, then $\bar{V}_S > \hat{V}_S$ and so $\tilde{m}_S(\bar{V}_S) < \tilde{m}_S(\hat{V}_S)$.*

Proof: This proof is identical in construction to the proof of Proposition 6a. (i) Adding (20) and (22) evaluated at $\hat{V}_N$, $\hat{V}_S$ and substituting into (26) and (27) yields

$$\frac{dE\hat{W}_G}{dV_N} = (\hat{\alpha}_S + \hat{\alpha}'_S(V_S) \cdot V_S) \cdot \rho(R(\hat{V}_N + \hat{V}_S)) > 0 \tag{31}$$

and

$$\frac{dE\hat{W}_G}{dV_S} = (\hat{\alpha}_N + \hat{\alpha}'_N(V_N) \cdot V_N) \cdot \rho(R(\hat{V}_N + \hat{V}_S)) > 0. \tag{32}$$

So at $(\hat{V}_N + \hat{V}_S)$ expected global welfare can be increased by increasing one of V_N or V_S. Therefore, $\bar{V} = \bar{V}_N + \bar{V}_S > \hat{V}_N + \hat{V}_S$. (ii) As $\tilde{m}_S(V_S)$ is a decreasing function of V_S, $\tilde{m}_S(\bar{V}_S) < \tilde{m}_S(\hat{V}_S)$, since $\bar{V}_S > \hat{V}_S$.

The intuition for this result is identical to that for Proposition 6a except patent breadth rather than patent length is varied to achieve the aggregate reward. As a result, in the developing country, patent breadth is broader under the optimal global policy than it is under the individual country optimal policy.

The intuition for Proposition 6a carries over to the cases involving asymmetric costs and demands as well so that the optimal global policy always involves a larger aggregate reward than the sum of the individual country rewards. Through (18) and the second order condition for a maximum, this larger aggregate reward translates to more R&D under the global policy and so a higher probability that an innovation will result.

Welfare: By construction, aggregate welfare under the optimal global policy is greater than under the optimal individual country policies. [13] However, this does not ensure that both countries are better off under the optimal global policy. Both countries gain from the extra R&D increasing the probability of a successful innovation, but as welfare in each country also depends on patent length and patent breadth in each country, the changes in patent breadth and length that arise from the optimal global policy must be considered.

It is shown below under symmetry and where the optimal global patent involves identical patent designs in each country, that country N is unambiguously better off with optimally global patents, while country S might be worse off.

Symmetric Costs and Demand Intercepts - Narrow Patents: In the proof of Proposition 6b it was established that $\bar{V}_S > \hat{V}_S$, that is, more reward is raised in country S under the optimal global policy than under the individual country policy. It was also established that the optimal global policy has a broader patent in country S than the optimal individual country policy, that is, $\tilde{m}_S(\bar{V}_S) < \tilde{m}_S(\hat{V}_S)$. Since $\alpha_i(m_i)$ is a decreasing function of m_i, deadweight loss in country S is greater under the optimal global policy than under the optimal individual country policy. Therefore, it is possible that country S might be worse off under the optimal global policy compared to the optimal individual country policy. If country S is worse off under the optimal global policy, then country N must be better off because global welfare is greater under the optimal global policy. If country S is better off under the optimal global policy, then country N must also be better off because it receives the same R&D benefit as country S, but at a lower cost, $\bar{V}_S - \hat{V}_S > \bar{V}_N - \hat{V}_N$. Therefore, country N is unambiguously better off under the optimal global policy.

[13] This contrasts with Deardorff (1992), where it was found the world could lose from the extension of patent protection to all countries. This difference arises because in this paper optimal patent designs are used while Deardorff uses patents that are extremely broad and of infinite length, that is, patents that give complete protection.

Symmetric Costs and Demand Intercepts - Broad Patents: Proposition 3a established that the optimal global aggregate reward was unique, but its division between countries N and S was not. As a result, divisions exist such that either country can be worse off under the optimal global policy as compared to the optimal individual country policy. However, no divisions exists under which both countries are worse off. As any division is optimal, equal division is optimal. In this case, $\bar{V}_N = \bar{V}_S$ and $\bar{T}_N = \bar{T}_S$. Where the optimal global patent design involves broad patents of equal length in both countries, Proposition 6a demonstrated that $\bar{V}_S > \hat{V}_S$ and $\bar{T}_S > \hat{T}_S$. Using similar arguments to those in the previous sub-section reveals that country N is unambiguously better off with the optimal global patent design whereas country S might be worse off.

These findings are consistent with Deardorff (1992) except here patents design is optimal both before and after the global patent policy is introduced. In the light of the preceding arguments, it is not surprising that the developed (innovating) countries were strong advocates of a TRIPS agreement that strengthened patent protection in the developing (non-innovating)world so that both the developed and developing world's patent protection became similar. [14]

This welfare analysis suggests that if developing (non-innovating) countries are going to continue to embrace the 1994 TRIPS agreement, then the issue of the redistribution of the global gains from strengthening patent protection needs to be addressed. From an efficiency perspective, it is not appropriate to manipulate patent breadth and length so that no country is made worse under a global patent regime, rather an explicit redistribution mechanism, that is independent of patent design, is needed.

[14] Where country costs and demands are asymmetric, no unambiguous statements about the welfare implications of optimally global patents can be made. However, if under asymmetry, it is assumed that $\hat{V}_N > \hat{V}_S$, then the arguments above can be used to show that country N is unambiguously better off with the optimal global patent design whereas country S might be worse off.

6. Conclusion

In this paper, optimal patent design has been examined in a global context. First, it was shown that innovating countries provide greater patent protection, in the form of longer or broader patents, than non-innovating countries. To the extent that innovation typically occurs in developed nations, this effect may help explain why prior to the conclusion of the TRIPs agreement, developed countries typically chose more extensive patent protection than developing countries. Secondly, it was shown that the optimal global patent design might involve identical patent breadth and length in all countries, though not necessarily. Thirdly, it was shown that the patents designed by individual countries provide too little protection from a global point of view and finally it was shown that although innovating (developed) countries are unambiguously better off with the optimal global patent design, non-innovating (developing) countries may be worse off. This last result suggests that the success of the present and any future TRIPS agreement might ultimately depend on how redistribution of the gains from the optimal global patent regime is achieved.

Some comments on the specific nature of the model used in this paper seem appropriate. First, the results are derived in a model in which (i) patent design does not affect the timing of innovation, (ii) only one firm has the ability to innovate in the industry under consideration, and (iii) innovation is not sequential. Although parts of the literature on optimal patent design relax each of these assumptions, Denicolo (1996, 1999), Duffy (2002), and Green and Scotchmer (1995), they are standard in a large part of the literature, Gallini (1992), Gilbert and Shapiro (1990), Klemperer (1990), and Nordhaus (1969). The results in this paper are best viewed as extending this latter literature.

Secondly, the demand structure assumed has the property that the deadweight loss per unit of profit is a monotonic function of one of its parameters and independent of the others. This greatly simplifies the determination of whether broad or narrow patents of infinite length are optimal. Using a different assumption about demand would complicate the analysis, but not change the general thrust of the analysis.

Thirdly, imitators were assumed to sell only in the country in which they are located. If imitators from one country were allowed to sell in the market of the other, then the profits earned in the other country would effect the Nash equilibrium rewards of each country. Once again, this complicates the analysis, without offering much additional insight. Similarly, if the number of imitators was determined by the interaction of imitation costs and free entry conditions, as in Gallini (1992) or Wright (1999), rather than being part of patent design, then the optimal patent design would depend on imitation costs. However, all that would change is that where narrow patents were optimal, less narrow patents would now be optimal. [15] Nothing qualitatively would change.

Finally, it should be noted that patent design in this paper is product specific. Different products have different demand structures, R&D costs, production costs, and probability functions for a successful innovation, and so have different optimal patent designs. As much of this information is private information, future research will have to be aimed at designing an optimal menu of patent designs from which innovators choose.

[15] A similar result is found in Takalo (1998).

7. References

Chin, J. C. and Grossman, G. M. (1990). 'Intellectual Property Rights and North-South Trade', In R. W. Jones and A.Ø. Krueger (eds.), *The Political Economy of International Trade: Essay in Honour of Robert E. Baldwin*, Basil Blackwell, Cambridge, MA.

Dasgupta, P. and Maskin, E. (1986). 'The Existence of Equilibrium in Discontinuous Economic Games, 1: Theory', *Review of Economic Studies*, **53**, 1-26.

Deardorff, A. V. (1992). 'Welfare Effects of Global Patent Protection', *Economica*, **59**, 35–51.

Denicolo, V. (1996). 'Patent Races and Optimal Patent Breadth and Length', *Journal of Industrial Economics*, **44**, 249-65.

Denicolo, V. (1999). 'The Optimal Life of a Patent when the Timing of Innovation is Stochastic', *International Journal of Industrial Organization*, **17**, 827-846.

Duffy, J. F. (2002). 'A Minimum Optimal Patent Length', College of William and Mary - Marshall Wythe School of Law Working Paper.

Green, J. R. and Scotchmer, S. (1995). 'On the Division of profit in Sequential Innovation', *Rand Journal of Economics*, **26**, 20–33.

Gallini, N. T. (1992). 'Patent Policy and Costly Imitation', *Rand Journal of Economics*, **23**, 52–63.

Gilbert, R. and Shapiro, C. (1990). 'Optimal Patent Length and Breadth', *Rand Journal of Economics*, **21**, 106–12.

Grossman, G. M. and Lai, E. (2001). 'International Protection of Intellectual Property', (http://www.princeton.edu/ grossman/Ipr082902.pdf)

Klemperer, P. (1990). 'How Broad Should the Scope of Patent Protection be?', *Rand Journal of Economics*, **21**, 113–30.

Lanjouw, J. O. and Cockburn, I. (2000). 'Do Patents Matter? Empirical Evidence After the Gatt' NBER Working Paper 7495.

Nordhaus, W. D. (1969). *Invention, Growth and Welfare: A Theoretical Treatment* *of Technological Change*, Cambridge MA: MIT Press.

O'Donoghue, T., Scotchmer, S., and Thisse, J-F. (1998). 'Patent Breadth, Patent Life, and the Pace of Technological Progress', *Journal of Economics and Management Strategy*, **17**, 1-32.

Takalo, T. (1998) 'Innovation and Imitation Under Imperfect Patent Protection', *Journal of Economics*, **67**, 229-41.

Wright, D. J. (1999). 'Optimal Patent Breadth and Length with Costly Imitation', *International Journal of Industrial Organization*, **17**, 419–36.

WTO, 2000, *Overview: the TRIPS Agreement*, World Trade Organization, www.wto.org.

8. Appendix

Cournot Equilibrium Outputs

Firm profit in country $i = N, S$ is given by

$$\Pi_i = (a - b(Q_{-i} + q_i)^{d_i} - c_i) \cdot q_i, \tag{33}$$

where Q_{-i} is the sum of all other firms outputs. Differentiating with respect to q_i, setting this equal to zero and using symmetry, $Q_{-i} = m_i q_i$, yields

$$q_i^{d_i} = \frac{(a - c)}{b((m_i + 1)^{d_i} + d_i \cdot (m_i + 1)^{d_i - 1})}. \tag{34}$$

Rearranging yields

$$q_i = \left(\frac{(a - c) \cdot (m_i + 1)^{1 - d_i}}{b \cdot (1 + d_i + m_i)}\right)^{\frac{1}{d_i}}. \tag{35}$$

Narrow Patent

Applying the Implicit Function Theorem to (3) of the text yields

$$\tilde{m}'_N(V_N) = \frac{r}{\frac{d\Pi_N}{d\tilde{m}_N}(\tilde{m}_N(V_N))} < 0 \tag{36}$$

since

$$\frac{d\Pi_N}{d\tilde{m}_N}(\tilde{m}_N(V_N)) < 0. \tag{37}$$

An identical result holds for $\tilde{m}'_S(V_S)$.

By the Chain Rule

$$\hat{\alpha}'_N(V_N) = \frac{\partial \alpha_N}{\partial m_N} \frac{d\tilde{m}_N}{dV_N} > 0 \tag{38}$$

since $\frac{\partial \alpha_N}{\partial m_N} < 0$ for narrow patents and $\frac{d\tilde{m}_N}{dV_N} < 0$ from above. An identical result holds for $\hat{\alpha}'_S(V_S)$.

The Innovator

Applying the Implicit Function Theorem to first order condition (18) of the text yields

$$R'(V_N + V_S) = -\frac{\rho'(\cdot)}{\rho''(\cdot)(V_N + V_S)} > 0 \tag{39}$$

since $\rho'(\cdot) > 0$ and $\rho''(\cdot) < 0$. Differentiating again and assuming $\rho'''(\cdot) = 0$ gives

$$R''(V_N + V_S) = \frac{-R'(V_N + V_S) \cdot (V_N + V_S) + \frac{\rho'(\cdot)}{\rho''(\cdot)}}{(V_N + V_S)^2} < 0 \tag{40}$$

The Policy Maker in Country N

The second order condition for a maximum is

$$\begin{aligned}\frac{d^2 E\hat{W}_N}{d^2 V_N} &= 2(1 - \hat{\alpha}_N(V_N) - \hat{\alpha}'_N(V_N) \cdot V_N)\rho'(\cdot)R'(\cdot) \\ &- (2\hat{\alpha}'_N(V_N) + \hat{\alpha}''_N(V_N) \cdot V_N) \cdot \rho(\cdot) - \omega \cdot R''(V_N + V_S) \\ &+ [C^c_N/r + (1 - \hat{\alpha}_N(V_N)) \cdot V_N + V_S] \cdot R'(\cdot)^2 \cdot \rho''(\cdot) \\ &+ [C^c_N/r + (1 - \hat{\alpha}_N(V_N)) \cdot V_N + V_S] \cdot R''(\cdot) \cdot \rho'(\cdot) < 0 \end{aligned} \tag{41}$$

Applying the Implicit Function Theorem to (20) of the text gives the slope of the country N equilibrium locus as

$$\begin{aligned}\frac{dV_N}{dV_S} &= -\Bigg[\rho'(\cdot)R'(\cdot) + (1 - \hat{\alpha}_N(V_N) - \hat{\alpha}'_N(V_N) \cdot V_N)\rho'(\cdot)R'(\cdot) \\ &+ [C^c_N/r + (1 - \hat{\alpha}_N(V_N)) \cdot V_N + V_S] \cdot R'(\cdot)^2 \cdot \rho''(\cdot) \\ &+ [C^c_N/r + (1 - \hat{\alpha}_N(V_N)) \cdot V_N + V_S] \cdot R''(\cdot) \cdot \rho'(\cdot) - \omega \cdot R''(V_N + V_S)\Bigg] \\ &\quad \frac{d^2 E\hat{W}_N}{d^2 V_N}\end{aligned} \tag{42}$$

The sign of this is ambiguous because the first and last terms in the [ ] bracket are positive.

The Policy Maker in Country S

The second order condition for a maximum is

$$\begin{aligned}\frac{d^2 E\hat{W}_S}{d^2 V_S} &= 2(-\hat{\alpha}_S(V_S) - \hat{\alpha}'_S(V_S) \cdot V_S)\rho'(\cdot)R'(\cdot) \\ &- (2\hat{\alpha}'_S(V_S) + \hat{\alpha}''_S(V_S) \cdot V_S) \cdot \rho(\cdot) \\ &+ [C^c_S/r - \hat{\alpha}_S(V_S) \cdot V_S] \cdot R'(\cdot)^2 \cdot \rho''(\cdot) \\ &+ [C^c_S/r - \hat{\alpha}_S(V_S) \cdot V_S] \cdot R''(\cdot) \cdot \rho'(\cdot) < 0\end{aligned} \tag{43}$$

Applying the Implicit Function Theorem to (22) of the text gives the slope of the country S equilibrium locus as

$$\begin{aligned}\frac{dV_S}{dV_N} &= -\Big[(-\hat{\alpha}_S(V_S) - \hat{\alpha}'_S(V_S) \cdot V_S)\rho'(\cdot)R'(\cdot) \\ &+ [C_S^c/r - \hat{\alpha}_S(V_S) \cdot V_S] \cdot R'(\cdot)^2 \cdot \rho''(\cdot) \\ &+ [C_S^c/r - \hat{\alpha}_S(V_S) \cdot V_S] \cdot R''(\cdot) \cdot \rho'(\cdot)\Big] / \frac{d^2 E\hat{W}_S}{d^2 V_S} \\ &< 0 \end{aligned} \tag{44}$$

Asymmetric Costs - Broad Patent

(i) $\frac{dT_N}{d(\frac{V_N}{PN})} > 0$ and $\frac{d^2T_N}{d(\frac{V_N}{\Pi_N})^2} < 0$.

Applying the Implicit Function Theorem to (2) yields

$$\frac{dT_N}{d(\frac{V_N}{\Pi_N})} = \frac{1}{e^{-rT_N(\cdot)}} > 0 \tag{45}$$

and

$$\frac{d^2T_N}{d(\frac{V_N}{\Pi_N})^2} = \frac{1}{(e^{-rT_N(\cdot)})^2} \cdot re^{-rT_N(\cdot)} \cdot \frac{dT_N}{d(\frac{V_N}{\Pi_N})}. \tag{46}$$

An identical result holds for T_S.

Figure 1a

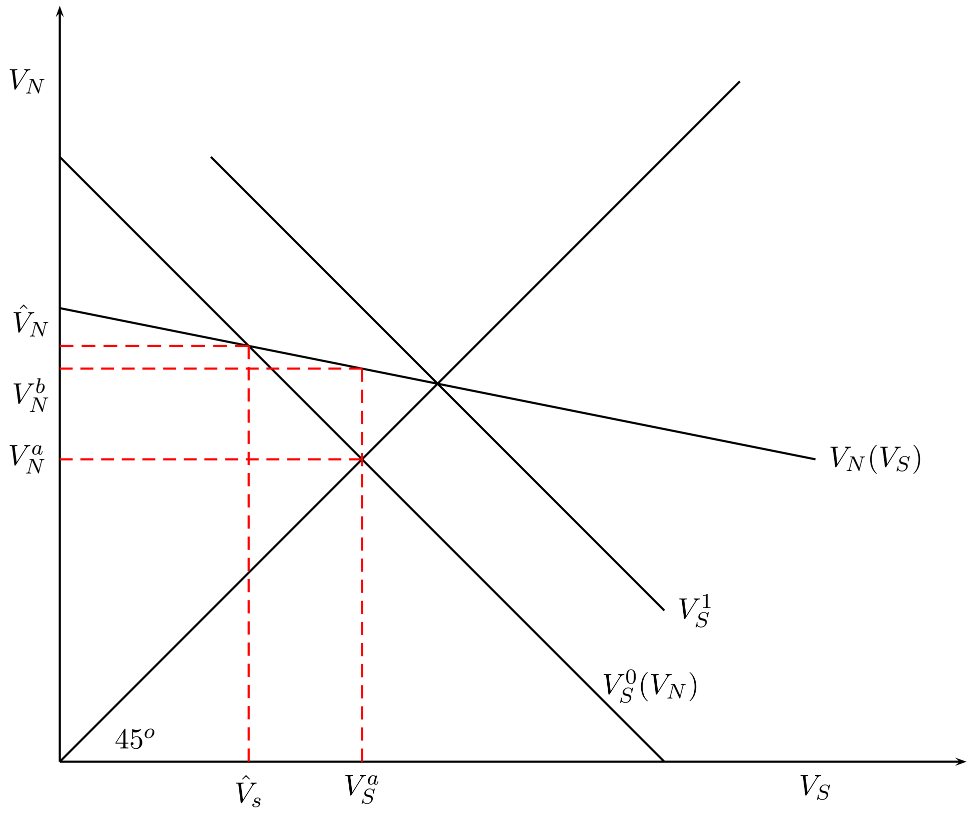

Figure 1b