

THREE ESSAYS ON INTERNATIONAL TRADE

By

Yih-ming Lin

Dissertation

Submitted to the Faculty of the  
Graduate School of Vanderbilt University  
in partial fulfillment of the requirements  
for the degree of

DOCTOR OF PHILOSOPHY

in

Economics

August, 2005

Nashville, Tennessee

Approved:

Professor Eric W. Bond

Professor Jennifer F. Reinganum

Professor Benjamin C. Zissimos

Professor Mikhael Shor

**To**

**My Parents and Hung-Pin**

## ACKNOWLEDGEMENTS

I am especially indebted to my advisor, Dr. Eric W. Bond. Without his direction, suggestions and comments, this dissertation would not have been possibly completed. I also would like to put in record my thanks and gratitude for my committee members: Dr. Jennifer F. Reinganum, Dr. Benjamin C. Zissimos, and Dr. Mikhael Shor for their generosity, time, and comments. They made me notice my mistakes and give me useful suggestions. I am thankful to my wife, Hung-pin, who took care of our family when I was away.

## TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS .....	iii
LIST OF TABLES .....	v
LIST OF FIGURES.....	vi
 Chapter	
I. INTRODUCTION.....	1
Essay 1: Bilateral trade flows and nontraded goods .....	1
Essay 2: Arbitrage and the harmonization of patent lives .....	1
Essay 3: Why does the U.S. prevent parallel import?.....	2
 II BILATERAL TRADE FLOWS AND NONTRADED GOODS .....	 3
Introduction .....	3
The model .....	6
Data .....	10
Empirical results .....	10
Conclusions .....	12
Appendix .....	13
References .....	15
 III ARBITRAGE AND THE HARMONIZATION OF PATENT LIVES .....	 17
Introduction .....	17
The model .....	19
The non-cooperative equilibrium.....	24
Efficient international agreement .....	32
Conclusions .....	34
Appendix .....	35
References .....	41
 IV WHY DOES THE U.S. PREVENT PARALLEL IMPORTS? .....	 43
Introduction .....	43
The single product case .....	46
The welfare effects of parallel trade incorporated with innovation.....	48
Conclusions .....	53
References .....	56

## LIST OF TABLES

Table	Page
II.1. 1-Si, Export/GDP Ratio (%), and Import/GDP Ratio (%) (1995).....	4
II.2. An Example with Three Countries .....	8
II.3. Estimation Results.....	12
IV.1. Summary of IPRs Exhaustion Regime .....	45

## LIST OF FIGURES

Figure	Page
III.1. Profit, $\pi(\lambda)$ , consumer's surplus, $s(\lambda)$ and social deadweight loss, $\Delta(\lambda)$ .....	23
III.2. Marginal benefit and marginal cost of increasing enforcement level, $\lambda$ .....	27
III.3. Right hand side (RHS) and left hand side (LHS) of equation (16) where $\lambda' > \lambda$ .....	29
III.4. The conditional best response function of the North and South .....	36
IV.1. Relative welfare, $\gamma=5$ .....	51
IV.2. Relative welfare, $\gamma=10/11$ .....	52
IV.3. The relation between elasticity of innovation, $\gamma$ and $x$ .....	54

# CHAPTER I

## INTRODUCTION

This dissertation contains three essays. In the first essay, I consider nontraded goods as an important impact on bilateral trade flows. A testable gravity equation is derived and a simple example is demonstrated. In my second essay, I study a North/South model of intellectual property rights protection in which markets are not completely segmented and investigate how arbitrage affects IPR protection. In the third essay, we propose a two-country model of parallel trade with innovation. This model tries to explain why some countries, such as the U.S., do not allow parallel trade and some, such as Japan, do. The abstracts of these three essays are as follows:

### **Essay 1: Bilateral Trade Flows and Nontraded Goods**

This essay develops a monopolistic competition model with nontraded goods, which provides an explanation for why the real volume of trade is much lower than predicted by Helpman and Krugman's (1985) model. Furthermore, it explains the phenomenon that the volume of trade among high-income countries is relatively larger than the volume of trade between high-income and low-income countries. We also derive a testable gravity equation from this model. A sample of 1995 including 118 countries is examined. Our results show that evidence from the data is consistent with the prediction of this model; further, the goodness-of-fit increases as nontraded goods are considered.

### **Essay 2: Arbitrage and the Harmonization of Patent Lives**

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-strategy 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 global Pareto efficiency when the markets are not perfectly segmented.

### **Essay 3: Why Does the U.S. Prevent Parallel Imports?**

In this essay, we propose a two-country model of parallel trade with innovation. This model tries to explain why some countries, such as the U.S., do not allow parallel imports and some, such as Japan, do. We find that the welfare effects of parallel trade are related to the elasticity of innovation. If the elasticity of innovation is high, the welfare of the importing country improves when the difference between the importing and exporting markets is small, but it becomes worse when the difference is large. Furthermore, global welfare decreases anyway when innovation is considered. For the case of low elasticity of innovation, it is possible for global welfare to be improved by allowing parallel trade.



## CHAPTER II

### BILATERAL TRADE FLOWS AND NONTRADED GOODS

#### Introduction

In the last few decades, the most important development in the theory of international trade is the monopolistic competition model. Helpman and Krugman (1985) proposed a model in which monopolistically competitive firms produce differentiated goods using an increasing returns to scale technology (IRS) and all individuals have identical homothetic preferences and a “love for variety”.<sup>1</sup> Their model provides an explanation for the phenomenon of large volumes of trade among similar countries with a factor-proportions view of intersectoral trade flows, which could not be explained by the traditional Heckscher-Ohlin (HO) theorem.

In Helpman and Krugman’s model, it is assumed that the economy has free trade, balanced trade, no transport cost, all tradeable goods, and identical production technology across countries. The monopolistic competition model yields the following equation to predict the volume of bilateral trade,

$$M_{ij} = \frac{Y_i Y_j}{Y_w} = s_j Y_i, \quad (1)$$

where  $M_{ij}$  is the imports of country  $i$  from country  $j$ ,  $Y_i$  ( $Y_j$ ) is the gross domestic product (GDP) of country  $i$  ( $j$ ),  $Y_w$  is the total world income, and  $s_j = \frac{Y_j}{Y_w}$  is the share of country  $j$  in total world income. Equation (1) means that the bilateral trade flows are positively related to the product of countries’ GDPs, which is the simplest form of gravity equation.

However, the volume of trade in the real world is much less than the amount predicted by equation (1). For example, the volume of trade in the world is about 5,214 billion U.S. dollars, which is much lower than the predicted number, 25,033 billion US dollars in 1995.<sup>2</sup> Furthermore, let us take a look at the country’s data. The Export/GDP ratio and Import/GDP ratio of a sample of countries are shown in Table 1.<sup>3</sup> The predicted Export/GDP ratio or Import/GDP ratio

---

<sup>1</sup>The term “love for variety ” is from Jensen (2000).

<sup>2</sup>The trade data is from the CD-ROM “World Trade Flows, 1980-1997, with Production and Tariff Data,” the predicted value is calculated based on the GDP data from World Development Indicators CD-ROM, 2000.

<sup>3</sup>This is the same sample employed in Jensen (2000). This sample includes ten rich countries, ten middle income countries, and ten poor countries.

Table 1:  $1 - s_i$ , Export/GDP Ratio(%), and Import/GDP Ratio(%)(1995)

Country	$1 - s_i$	$\frac{Export}{GDP}$	$\frac{Import}{GDP}$	Country	$1 - s_i$	$\frac{Export}{GDP}$	$\frac{Import}{GDP}$
Australia	0.987	20.02	20.13	Austria	0.992	38.69	39.10
Bangladesh	0.999	11.02	17.14	Brazil	0.975	7.72	9.49
Canada	0.980	38.23	35.26	Chile	0.998	30.55	28.73
Colombia	0.997	14.53	20.96	Denmark	0.994	35.36	31.36
Ecuador	0.999	29.69	28.63	Egypt	0.998	22.45	27.50
France	0.946	23.53	21.16	Germany	0.915	23.85	23.07
Ghana	1.000	24.50	32.82	Greece	0.996	15.84	24.98
India	0.987	10.90	14.07	Indonesia	0.993	26.31	27.65
Israel	0.997	31.32	47.59	South Korea	0.983	30.20	31.68
Malaysia	0.997	95.42	99.42	Netherlands	0.986	53.09	46.41
Pakistan	0.998	15.95	21.11	Philippines	0.997	36.36	44.18
Portugal	0.996	30.84	37.73	Spain	0.980	23.98	23.88
Sweden	0.992	40.93	34.50	Turkey	0.994	19.89	24.35
United Kingdom	0.961	28.61	29.29	United States	0.751	11.44	12.83
Venezuela	0.997	26.78	21.31	Zambia	1.000	37.63	43.40

Source: World Development Indicators CD-ROM,2000.

of country  $i$  is  $1 - s_i$ .<sup>4</sup> In this sample, it can be found that most of the country's Export/GDP ratios and Import/GDP ratios are between 10%-50%, which are much lower than  $(1 - s_i)$ , the expected ratio implied by the monopolistic competition model for all countries except Malaysia. Obviously, that is because the bilateral trade flows are overestimated by Krugman and Helpman's model.

Many possible factors could reduce the bilateral trade flows. For example, high transport cost could decrease the volume of trade. Most countries do not have completely free trade. Further, not all goods in the real world are tradeable such as services, education, and housing. All of these facts will reduce the volume of trade. However, most of the seminal studies do pay attention to the influence of transport cost. The gravity equation, in general, shows that the volume of bilateral trade is not only positively related to both incomes, but also negatively related to the distance between the two trade partners, in which the distance is used as a proxy variable for transport

<sup>4</sup> $\sum_{j \neq i} M_{ij} = \sum_{j \neq i} s_j Y_i = (1 - s_i) Y_i$ .

costs.

The study of the gravity equation is probably the most successful empirical work in international trade. The model was first introduced by Tinbergen (1962), in which the econometric model of trade flows was based only on intuitive justification. Anderson (1979) was the first to derive the gravity equation theoretically. Anderson (1979) assumed that preferences are Cobb-Douglas (or CES) and the goods are differentiated by countries of origin, which is called the Armington assumption. Bergstrand (1985) used CES preferences over Armington-differentiated goods to develop a general equilibrium model of world trade, thus yielding a reduced-form gravity equation for bilateral trade involving price indices.

The gravity equation model has been widely employed to infer substantial bilateral trade flow effects. Aitken (1973) showed that European trade is significantly influenced by membership in a trading community such as the European Economic Community (EEC) and the European Free Trade Association (EFTA), and the neighboring country dummy variable. Geraci and Prewo (1976) included average tariff rates, preferential trading group membership and common language in a gravity model and found that these effects have a significant impact on bilateral trade flows in a sample of OECD countries. Srivastava and Green (1986) incorporated some noneconomic variables in a gravity model and found that cultural similarity, political instability, economic union and former colonial status are all significant determinants of bilateral trade flows. Thursby and Thursby (1987) found strong support for the hypothesis that increased exchange rate variability affects bilateral trade flows. Recently, Bougheas, Demetriades and Morgenroth (1999) examined the role of infrastructure in a bilateral trade model with transport costs and the empirical results demonstrated that the level of infrastructure and the trade volume have a positive relationship.

Jensen (2000) modifies the gravity model by using disaggregated production data instead of GDP as a measure of production for exporting countries. This is because nontraded goods are included in GDP. Therefore, GDP is not considered as a good indicator of potential exports. However, his gravity equation included a country's income share, in which the nontraded goods problem also exists. Helpman and Krugman (1985),<sup>5</sup> discussed the model with the introduction of nontraded goods. They characterize the integrated world equilibrium and the region of factor price equalization when nontraded goods exist. However, they do not analyze the relationship between the volume of trade and nontraded goods. Schmitt and Yu (2001) employ a monopolistic competitive model with traded and non-traded goods to show that an increase in the degree of economies of scale of production will increase the volume of intra-industry trade, which is the

---

<sup>5</sup>See Chapter 10.

finding of the empirical study of Harrigan (1994). To date, to the best of my knowledge, there is no paper that includes nontraded goods in the gravity model and examines the relationship between the consumption of nontraded goods and the volume of trade, even though nontraded goods have an important impact upon bilateral trade flows.

In this paper, the assumption that all goods are tradable is relaxed. I incorporate nontraded goods in a monopolistic competition model. This model tries to provide an explanation of why the real volume of trade is much lower than predicted. The intuition is quite simple. It is because the countries do not have so many goods that are tradable in the real world. An estimable gravity equation can be derived from this model. Moreover, since this model incorporates nonhomothetic preferences, it also can explain the phenomenon that the volumes of trade among the industrialized countries are relative large to the volumes of trade between developed and less-developed countries, which was proposed by Linder (1961) and Markusen (1986).<sup>6</sup>

The remainder of this paper is organized as follows: in Section 2, we set up the model and derive the gravity equation with nontraded goods. The individuals consume differentiated tradable goods as well as nontraded goods, and preferences are specified as an additively separable utility function. It shows that bilateral trade flow is not only related to the GDPs of importing and exporting countries, but also is a function of the per capita income of both countries. The bilateral trade flows are negatively related to the consumption of nontraded goods. Section 3 describes the data sets used in this study. The empirical results are given in Section 4. They strongly support the prediction of this model in a sample of 118 countries in 1995. Section 5 concludes.

## The Model

In this section, the model is set up. Consider an open economy which has balanced trade, no transport costs, and with identical production technologies. There are  $m$  countries and two types of goods, tradeable goods ( $x_k$ ) and nontraded goods ( $z$ ). The tradeable goods ( $x_k$ ) are differentiated manufactured goods which are produced with production technologies with increasing returns to scale. The nontraded good is a homogeneous commodity and is produced with a constant returns to scale (CRS) technology.

---

<sup>6</sup>Markusen (1986) proposed a nonhomothetic model to explain the difference between the volume of W-E trade and the volume of N-S trade. Unfortunately, his model does not offer a testable model to predict the volume of trade from his model. However, in the extension of his paper, Hunter and Markusen (1988) proposed a nonhomothetic model and estimated a linear expenditure system (LES) to show that the demand is nonhomothetic.

It is assumed that all individuals consume tradeable goods as well as nontraded goods. Consumers have the following identical nonhomothetic preferences, which are given by

$$U = (\sum_k x_k^\alpha)^{1/\alpha} + u(z), \quad 0 < \alpha < 1, \quad (2)$$

where  $u(\cdot)$  is a strictly concave function.

All individuals maximize their utility subject to their income. Since the subutility function for differentiated goods is homogeneous of degree one, we can use a two-stage budgeting procedure to solve this utility maximization problem.<sup>7</sup> The consumer's problem can be rewritten as

$$\begin{aligned} \text{Max} \quad & U = \mathcal{X} + u(z), \\ \text{s.t.} \quad & \mathcal{P}\mathcal{X} + p_z z = I. \end{aligned} \quad (3)$$

where  $I$  is individual income,  $\mathcal{X} = (\sum_k x_k^\alpha)^{1/\alpha}$  is the quantity index for differentiated goods,  $\mathcal{P} = (\sum_k p_k^{\alpha/(\alpha-1)})^{(\alpha-1)/\alpha}$  is the price index for  $\mathcal{X}$ , and  $p_z$  and  $p_k$  for good  $z$  and  $x_k$ , respectively.

If we consider  $\mathcal{X}$  as the numeraire, the utility can be considered as a quasi-linear utility function. According to the property of quasi-linear utility function, the consumption of  $z$  is constant which is determined by

$$u'(z) = \frac{p_z}{\mathcal{P}}, \quad (4)$$

for all consumers if the income is big enough. There is no income effect for  $z$ . Increasing individual income does not change the quantity of demand for good  $z$  at all, and all the extra income goes entirely to the consumption of differentiated goods. Let  $z = z^*$  satisfy equation (4) which means that  $z^*$  is the demand quantity of notraded good for every consumer in every country and thus is independent of individual income and the prices of the differentiated commodities. Furthermore, it is also assumed that the individual's income in every country is bigger than  $z^*$ . Due to this special property of the nonhomothetic preference, the production of nontraded good for country  $i$  is  $Z_i = n_i z^*$  and is produced first in country  $i$ , where  $n_i$  is the population of country  $i$ .

In addition, the differentiated manufactured goods are produced and freely traded. Just like the imperfect competition model proposed by Helpman and Krugman (1985) and Helpman (1987), a number of firms produce one differentiated commodity in a monopolistically competitive market. Also, all firms are equipped with identical IRS technology, and free entry leads to zero profit in equilibrium.

---

<sup>7</sup>The detail of derivation is in the Appendix.

Table 2: An Example with Three Countries

Country	Population	Per Capita Income	GDP	$\frac{Export^a}{GDP}$	$\frac{Export^b}{GDP}$
<i>A</i>	10	10	100	0.6	0.3937
<i>B</i>	10	10	100	0.6	0.3937
<i>C</i>	10	5	50	0.8	0.35

a. the expected ratio of  $\frac{Export}{GDP}$  in Helpman and Krugman's model.

b. the expected ratio of  $\frac{Export}{GDP}$  in this nontraded goods model.

In this model, the volume of trade is different from the result of Helpman and Krugman (1985) and Helpman (1987) but similar. Since this property of nonhomothetic preference is “love for variety” for differentiated goods, each country will demand all foreign varieties according to the country's share of world value of differentiated goods. Therefore, the value of differentiated goods that country  $i$  imports from  $j$ , denoted  $M_{ij}$ , is

$$M_{ij} = \frac{X_i X_j}{X_w}, \quad (5)$$

where  $X_j$  is the value of a differentiated good produced in country  $j$ , and  $X_w = \sum_j X_j$  is the world output of differentiated goods.

Let  $Y_i = X_i + Z_i$  denote the GDP of country  $i$ . Rearranging equation (5),  $M_{ij}$  can be rewritten as

$$M_{ij} = \frac{(Y_i - Z_i)(Y_j - Z_j)}{\sum_j (Y_j - Z_j)} = \frac{(Y_i - n_i z^*)(Y_j - n_j z^*)}{X_w}. \quad (6)$$

In the following, I take an example to show how this model can explain the question I proposed in Section 1. Consider a three-country economy, consisting of countries  $A$ ,  $B$ , and  $C$ . Suppose that there are 10 people in each country and each individual consumes 3 units of nontraded goods. Assume that the per capita income of  $A$  and  $B$  is 10 and the per capita income of  $C$  is 5. The summary of this example is shown in Table 2. In the imperfect competition model, the expected ratios of export over GDP for each country are 0.6, 0.6 and 0.8, respectively. However, the expected ratios are reduced if we consider the influence of nontraded goods. As can be seen in Table 2, the ratios are down to 0.3937, 0.3937 and 0.35, respectively. this can explain why the phenomenon in Table 1 happened.

Another advantage of this model is that the expected Export/GDP ratio of a small country is not necessary bigger than a large country. Since the Export/GDP ratio is  $1 - s_i$  in Helpman and Krugman's model, a country with a smaller GDP will have a larger Export/GDP ratio. However,

this is not consistent with the data. For example, since Colombia's GDP is apparently smaller than Germany's GDP, the Export/GDP and/or Import/GDP ratios of Colombia are smaller than Germany's ratios in in Table 1, which is not consistent with Helpman and Krugman's model. However, it can be explained by Table 2.

It can also be used to explain why the volume of trade among the developed countries is very large relative to the volume of N-S trade. Country  $A$  will import 30.625 ( $\frac{4900}{160}$ ) of goods from country  $B$ (the trade between developed countries) but only import 8.75 ( $\frac{1400}{160}$ ) from country  $C$ (the trade between the developed country and the developing country). In this case, the GDP of country  $B$  is twice as large as that of country  $C$ . However, the volume of trade between  $A$  and  $B$  is more than three times that of the trade volume between  $A$  and  $C$ .

Next, we are going to derive the gravity equation from this nontraded model. Taking natural logarithms of both sides of equation (6), it follows that

$$\begin{aligned} \ln(M_{ij}) &= -\ln(X_w) + \ln((Y_i - n_i z^*)(Y_j - n_j z^*)) \\ &= -\ln(X_w) + \ln((Y_i Y_j - Y_j n_i z^* - Y_i n_j z^* + n_i n_j (z^*)^2)), \end{aligned} \quad (7)$$

where  $\ln(X_w)$  is a constant.

In order to derive an estimable gravity equation, it is necessary to linearize the last term of equation (7). Applying the first order Taylor series approximation at  $z^* = 0$ , it yields the following estimable gravity equation

$$\begin{aligned} \ln(M_{ij}) &= -\ln(X_w) + \ln(Y_i Y_j) - \frac{n_i Y_j + n_j Y_i}{Y_i Y_j} z^* + \epsilon, \\ &= c + \ln(Y_i) + \ln(Y_j) - \left(\frac{1}{y_i} + \frac{1}{y_j}\right) z^* + \epsilon, \end{aligned} \quad (8)$$

where  $y_i$  is per capita income of country  $i$ ,  $\epsilon$  is the disturbance term. The above is our basic estimating equation.

Comparing the new gravity equation to the conventional gravity equation, there is a new item,  $-\left(\frac{1}{y_i} + \frac{1}{y_j}\right)z^*$ , in this model. It shows that the bilateral trade flow is not only related to the GDP of both countries, but also related to the demand for nontraded goods and the per capita incomes in both countries. According to the above equation, bilateral trade flow decreases if the consumption of nontraded goods increases or the per capita income decreases. It explains that the volume of trade between two rich countries is larger than that between two lower-income countries.

Since the pattern of bilateral trade is not included in the above description, in the following, the generalized model and prediction of the pattern of trade will be discussed. Suppose there are

three types of goods,  $X^1$ ,  $X^2$ , and  $z$ .  $X^1$  and  $X^2$  are tradable and differentiated goods produced by identical increasing return to scale technology. The definition of  $z$  is exactly as described. Furthermore, it is assumed  $X^1$  is relatively capital-intensive compared with  $X^2$ , and the modified utility function is specified as

$$U = (\Sigma x_{1k}^\alpha + \Sigma x_{2k}^\alpha)^{1/\alpha} + u(z), \quad 0 < \alpha < 1. \quad (9)$$

Given the above assumptions and the result of Helpman and Krugman (1985), we can obtain not only the result of equation (8), but also that a capital-abundant country imports labor-intensive goods from a labor-abundant country.

## Data

There are three databases used in this study. The world bilateral trade flows data is originally obtained from the CD-ROM “World Trade Flows, 1980-1997, with Production and Tariff Data,” which is available from the Social Science Data Service, Institute of Governmental Affairs, University of California, Davis. The data used to test this model is from the year of 1995. There are 182 countries or regions in this database. In Feenstra (2000), it indicated that the main source for bilateral trade data is the United National Statistical Office. Those data were also published in the Yearbook of International Trade Statistics and fully reported in Commodity Trade Statistics by the United Nation.

Another data set used in the study is the GDP and per capita income data which is downloaded from Harvard University, CID(Center for International Development). The GDP and per capita income are PPP-adjusted. In Gallup and Sachs (1999), it indicated that most of the GDP and per capita income data are from the World Bank (1997, 1998). For countries which are missing in the World Bank, the data is obtained from CIA (1996,1997).

In order to compare this nontraded goods model and the conventional gravity model, a distance data set is also employed. The distance data set is downloaded from Purdue University.<sup>8</sup> This data set contains 137 countries and it provides the distance between capital cities in kilometers.

Combining the above three data sets, there are 118 countries employed in this paper. The names of those countries are listed in the Appendix. From the Appendix, it can be found that almost all of the countries or regions in the world are included.

---

<sup>8</sup><ftp://intrepid.mgmt.purdue.edu/pub/Trade.Data/dist.txt>.



## Empirical Results

In this section, the gravity equations derived from the nonhomothetic model are estimated by using the above data set. Based on equation (8), the gravity equation can be specified as the following two forms

$$\ln(M_{ij}) = \alpha + \beta_1 \ln(Y_i) + \beta_2 \ln(Y_j) - z^* \left( \frac{1}{y_i} + \frac{1}{y_j} \right) + \epsilon, \quad (10)$$

or

$$\ln(M_{ij}) = \alpha + \beta \ln(Y_i Y_j) - z^* \left( \frac{1}{y_i} + \frac{1}{y_j} \right) + \epsilon, \quad (11)$$

where  $\alpha$ ,  $\beta$  and  $z^*$  are the coefficients to be estimated. Theoretically,  $\hat{\alpha}$  is expected to be negative,  $\hat{z}^*$  should be positive, and  $\hat{\beta}_1$  and  $\hat{\beta}_2$  should be around 1.

In order to demonstrate the advantage of this new model, the conventional gravity equation is also to be estimated, which is

$$\ln(M_{ij}) = \alpha + \beta_1 \ln(Y_i) + \beta_2 \ln(Y_j) + \epsilon. \quad (12)$$

Ordinary Least Square(OLS) estimation is employed to estimate the above regressions. The estimation results are shown in Table 3. There are 13,806 observations in this sample. In Table 3, it shows all of the estimated coefficients in equation (10) and (12) are significant at 1% significant level. The most important is that all of the signs of the estimated coefficients coincide with our expectation. In particular,  $\hat{z}^*$  is around \$1,053 in this model, which means that all individuals consume at least about \$1,053 goods produced by their country every year. Comparing to Eq(12), it can be found that the estimated coefficients of GDPs in Eq(10) are significant closer to unity than in Eq(12) which means that this new model has a theoretical advantage. This nontraded goods gravity equation is more consistent with the data than the conventional gravity model. The  $R^2$  of Eq(10) is about 0.60, which is bigger than the  $R^2$  in Eq(12). That shows that the goodness-of-fit of this new model is better than the conventional model.

Next, I examine the performance of this model when a distance term exists and compare the traditional gravity equation with a distance term. The results are also given in Table 3 where Eq(10') and Eq(12') are Eq(10) and Eq(12) with distance term, respectively. The  $\delta$  is the coefficient of the distance term in each model. Just like the results above, the signs of estimated coefficients are as expected and significant. We also can find that the estimated coefficients of GDPs are also much closer to unity than the result of Eq(12').

Table 3: Estimation Results

Coefficient	Eq(10)	Eq(12)	Eq(10')	Eq(12')
$\alpha$	-21.78 (0.317)	-26.62 (0.235)	-9.63 (0.416)	-15.40 (0.361)
$\beta_1$	1.51 (0.017)	1.68 (0.015)	1.52 (0.016)	1.71 (0.014)
$\beta_2$	1.14 (0.017)	1.32 (0.015)	1.15 (0.016)	1.34 (0.014)
$z^*$	1053.34 (47.272)	- -	1163.45 (44.606)	- -
$\delta$	- -	- -	-1.41 (0.034)	-1.36 (0.034)
$R^2$	0.603	0.588	0.647	0.630

Notes: Standard errors are given in parentheses.

## Conclusions

In this paper, we tried to provide an alternative explanation of why the real volume of trade is much less than the volume predicted by Helpman and Krugman's model. We consider nontraded goods as an important impact on bilateral trade flows. Furthermore, this model also explains why the volume of trade between N-S is much less than the volume among developed countries.

We developed a model with nontraded goods and nonhomothetic preferences and derived an estimable gravity equation. The model tries to make a connection to bridge the gap between theory and empirical work on the role of nontraded goods. This model finds that bilateral trade flows are not only related to the GDP of two countries, but also related to the per capita income of the two countries. The individual consumption level of nontraded goods can be estimated, which is also related to bilateral trade flows. In the empirical work of this study, we employed a sample of 1995 data with 118 countries to test this model. The empirical results are consistent with the expectations from this model.

For the sake of simplicity, it is assumed that all individuals consume the same quantity of nontraded goods in this model. In further study, this restrictive assumption might be relaxed.

## Appendix

### Derivation of the quasi-linear like utility function:

Suppose that every consumer shares the same preference for the two types of goods:

$$U = (\sum_k x_k^\alpha)^{1/\alpha} + u(z), \quad 0 < \alpha < 1, \quad (13)$$

where  $u(\cdot)$  is a strictly concave function.

Given income  $I$  and a set of prices,  $p_z$  and  $p_k$  for good  $z$  and  $x_k$ , respectively, the consumer's problem is to maximize utility function (13) subject to the budget constraint,

$$\sum_k p_k x_k + p_z z = I. \quad (14)$$

This problem can be solved in two steps. First, define

$$(\sum_k x_k^\alpha)^{1/\alpha} = \mathcal{X}.$$

where  $\mathcal{X}$  can be interpreted as a composite index of consumption of differentiated goods. For any given value  $\mathcal{X}$ , each  $x_k$  needs to be chosen so as to minimize the cost of attaining  $\mathcal{X}$ . Therefore, we can solve the following minimization problem:

$$\text{Min } \sum_k p_k x_k \quad \text{s.t.} \quad (\sum_k x_k^\alpha)^{1/\alpha} = \mathcal{X}. \quad (15)$$

By the first order condition, we can solve that  $x_k = x_{k'} (p_k/p_{k'})^{1/(1-\alpha)}$ . Substituting it into the constraint, we obtain

$$x_{k'} = \frac{p_{k'}^{1/(\alpha-1)}}{(\sum_k p_k^{\alpha/(\alpha-1)})^{1/\alpha}} \mathcal{X}.$$

We can also derive an expression for the minimum cost of attaining  $\mathcal{X}$  which is:

$$\sum_k p_k x_k = (\sum_k p_k^{\alpha/(\alpha-1)})^{(\alpha-1)/\alpha} \mathcal{X}.$$

where the term multiplying  $\mathcal{X}$  can be interpreted as a price index of  $\mathcal{X}$ . Denote that

$$\mathcal{P} = (\sum_k p_k^{\alpha/(\alpha-1)})^{(\alpha-1)/\alpha}.$$

Therefore, the consumer's problem can be rewritten as

$$\begin{aligned} \text{Max } \quad & U = \mathcal{X} + u(z), \\ \text{s.t.} \quad & \mathcal{P}\mathcal{X} + p_z z = I. \end{aligned}$$

If we consider  $\mathcal{X}$  as the numeraire.

$$\begin{aligned} \text{Max} \quad & U = \mathcal{X} + u(z), \\ \text{s.t.} \quad & \mathcal{X} + \frac{p_z}{\mathcal{P}}z = \frac{I}{\mathcal{P}}. \end{aligned} \tag{16}$$

Equation (16) is a quasi-linear utility function. The consumption of  $z$  is determined by the first order condition which is

$$u'(z^*) = \frac{p_z}{\mathcal{P}}. \tag{17}$$

**Country Names:** The data contains 118 countries. They are as follows: Algeria, Angola, Argentina, Australia, Austria, Bangladesh, Belgium-Lux, Benin, Bolivia, Brazil, Bulgaria, Burkina Faso, Burundi, Cameroon, Canada, Central African Republic, Chad, Chile, China, Colombia, Congo, Costa Rica, Czechoslovakia, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Ethiopia, Finland, France, Gabon, Gambia, Germany, Ghana, Greece, Guatemala, Guinea, Guinea Bissau, Haiti, Honduras, Hong Kong, Hungary, India, Indonesia, Iran, Iraq, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kenya, Korea Republic (S), Kuwait, Laos, Liberia, Madagascar, Malawi, Malaysia, Mali, Mauritania, Mauritius, Mexico, Mongolia, Morocco, Mozambique, Myanmar, Nepal, Netherlands, New Zealand, Nicaragua, Niger, Nigeria, Norway, Oman, Pakistan, Panama, Papua. New. Guinea, Paraguay, Peru, Phillipines, Poland, Portugal, Romania, Rwanda, Saudi Arabia, Senegal, Sierra Leone, Singapore, Solomon Islands, Somalia, South Africa, Spain, Sri Lanka, Sudan, Sweden, Switzerland, Syrian. Arab. Republic, Taiwan, Thailand, Togo, Trinidad And Tobago, Tunisia, Turkey, Uganda, United Arab Emirates, United Kingdom, Uruguay, USA, USSR, Venezuela, Yemen, Yugoslavia, Zaire, Zambia, Zimbabwe.

## References

- Aitken, Norman D. (1973), "The Effect of the EEC and EFTA on European Trade: A Temporal Cross-Section Analysis," *American Economic Review*, 63, 881-892.
- Anderson, James E. (1979), "A Theoretical Foundation for the Gravity Equation," *American Economic Review*, 69, 106-116.
- Bergstrand, Jeffrey (1985), "The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence," *Review of Economics and Statistics*, 67, 474-481.
- Bougheas, Spiros, Demetriades, Panico O. and Edgar L.W. Morgenroth (1999), "Infrastructure, Transport Cost and Trade," *Journal of International Economic*, 47, 169-189.
- CIA (1996) *The World Factbook*. Washington, D.C. Central Intelligence Agency.
- CIA (1997) *The World Factbook*. <http://www.odci.gov/cia/publications/factbook/index.html>.
- Feenstra, Robert C. (2000), "World Trade Flows, 1980-1997, with Production and Tariff Data," Working Paper, University of California, Davis.
- Gallup, John L. and Jeffery D. Sachs (1999), "Geography and Economic Development," Working Paper, Harvard University.
- Geraci, Vincent J. and Wilfried Prewo (1977), "Bilateral Trade Flows and Transport Costs," *Review of Economics and Statistics*, 59, 67-74.
- Harrigan, James (1994), "Scale Economies and the Volume of Trade," *Review of Economics and Statistics*, 76, 321-328.
- Helpman, Elhanan (1987), "Imperfect Competition and International Trade: Evidence from Fourteen Industrial Countries," *Journal of the Japanese and International Economies*, 1, 62-81.
- Helpman, Elhanan and Paul Krugman (1985), *Increasing Returns, Imperfect Competition, and Foreign Trade*, Cambridge: MIT Press.
- Hunter, Linda and James R. Markusen (1988), "Per Capita Income as a Basis for Trade," in: Robert Feenstra, ed., *Empirical Methods for International Trade*, Cambridge: MIT Press.
- Hunter, Linda (1991), "The Contribution of Nonhomothetic Preferences to Trade," *Journal of International Economics*, 30, 345-358.

- Jensen, Paul E. (2000), “ Analysis of Bilateral Trade Patterns with Panel Data,” *Review of International Economics*, 8, 86-99.
- Linder, Staffan B. (1961), *An Essay on Trade and Transformation*, New York: Wiley and Sons.
- Markusen, James R. (1986), “Explaining the Volume of Trade: An Eclectic Approach,” *American Economic Review*, 76, 1002-1011.
- Schmitt, Nicolas and Zhihao Yu (2001), “Economies of Scale and the Volume of Intra-Industry Trade,” Manuscript.
- Srivastava, Rajendra K. and Robert T. Green (1986), “Determinants of Bilateral Trade Flows,” *Journal of Business*, 59, 623-640.
- Thursby, Jerry G. and Marie C. Thursby (1987), “Bilateral Trade Flows, the Linder Hypothesis, and Exchange Risk,” *Review of Economics and Statistics*, 69, 488-495.
- Tinbergen, Jan (1962), *Shaping the World Economy: Suggestions for an International Economic Policy*, New York: Twentieth Century Fund.
- World Bank. (1997) *World Development Indicators 1997 CD-ROM*. Washington, D.C.: International Bank for Reconstruction and Development.
- World Bank. (1998) *World Development Indicators 1998 CD-ROM*. Washington, D.C.: International Bank for Reconstruction and Development.

## CHAPTER III

### ARBITRAGE AND THE HARMONIZATION OF PATENT LIVES

#### 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 policymaking, 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 spurred 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.

## 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.<sup>1</sup> 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

---

<sup>1</sup>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.

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:

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

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

where  $T$  is the length of the production period,  $\omega$  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  $\omega$  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  $x = (u')^{-1}$ .<sup>2</sup> The demand for the 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.

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

---

<sup>2</sup>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 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).

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.<sup>3</sup> Furthermore, the profit function is assumed to be strictly concave on  $[c, \bar{p}]$ , where  $\pi(c) = \pi(\bar{p}) = 0$ .

Let  $\lambda$  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  $\lambda$  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  $\lambda$  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  $i$ .<sup>4</sup> The conditions under which this will hold in the Nash equilibrium will be discussed below. Thus,

---

<sup>3</sup>This assumption follows Grossman and Lai (2004).

<sup>4</sup>We will discuss the case of  $t_n < t_s$  in the Appendix.

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:

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

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  $\pi(p_m)$  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  $\lambda$  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  $p'(\cdot)$  is positive, we can see that a decrease in  $\lambda$  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  $\tau_i$ , there exists a one-to-one relationship between  $t_i$  and  $\tau_i$ . For the sake of simplicity, we designate  $\tau_i$  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 ( $\tau_i$ ) and  $\lambda$ . 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:

$$W_n(\tau_s, \tau_n, \lambda, N) = (\tau_s(1 + M)\pi(p_m) + (\tau_n - \tau_s)\pi(p(\lambda)))N \\ + (\tau_s s(p_m) + (1 - \tau_n)s(c) + (\tau_n - \tau_s)s(p(\lambda)))N - C(N) - G(\lambda). \quad (2)$$

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  $\pi$  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 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

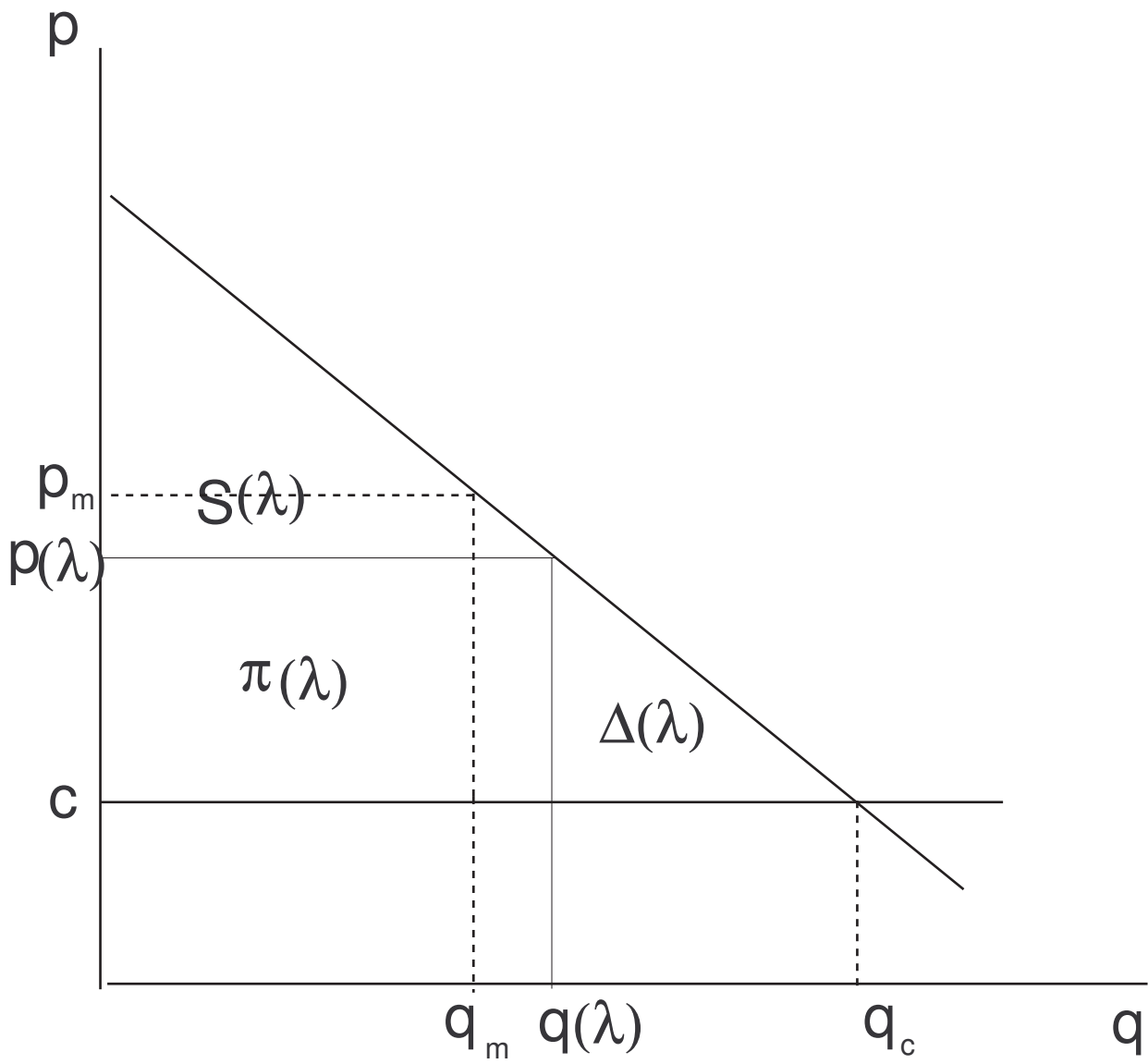


Figure 1: Profit,  $\pi(\lambda)$ , consumer's surplus,  $s(\lambda)$  and social deadweight loss,  $\Delta(\lambda)$ .

$s(p(\lambda))$  when the patent expires in the South but remains in effect in the North, with an enforcement level,  $\lambda$ .

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  $C(\cdot)$  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.  $G'(\cdot)$  is supposed to be positive which implies that the enforcement cost is increasing in enforcement level,  $\lambda$ .

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

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

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.

### 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,  $\tau_i$ , and the level of enforcement,  $\lambda$ .<sup>5</sup> 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,  $N^* = N^*(\tau_s, \tau_n, \lambda)$ .

---

<sup>5</sup>It is not reasonable if the North chooses  $\lambda$  first and then two countries choose  $\tau_n$  and  $\tau_s$ . If the two countries choose  $\tau_n$  and  $\tau_s$  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.

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

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

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 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

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

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

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

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

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

and

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

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

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

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

---

<sup>6</sup>It is surprising that we will find that the North will not choose to completely eliminate arbitrage even though enforcement is costless.

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:

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

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  $\lambda$  goes to  $\bar{\lambda}$ , which means the marginal benefit of enforcement goes to 0 as  $\lambda$  goes to  $\bar{\lambda}$ . Moreover,  $\Delta'(p(\lambda))$  is positive and not close to zero as the  $\lambda$  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  $\lambda$  will increase social welfare in the North for any  $(\tau_s, \tau_n)$ , which guarantees that the North's optimal enforcement policy will be less 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  $\lambda$ , we can therefore conclude that there exists an interior optimal  $\lambda^*(\tau_s, \tau_n) \in (0, \bar{\lambda})$  for any  $(\tau_s, \tau_n)$  such that the first order condition for  $\lambda$  is satisfied.

Figure 2 illustrates the existence of an optimal level of enforcement to prevent arbitrage that is smaller than  $\bar{\lambda}$ ,<sup>7</sup> 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.

---

<sup>7</sup>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  $\lambda$ .



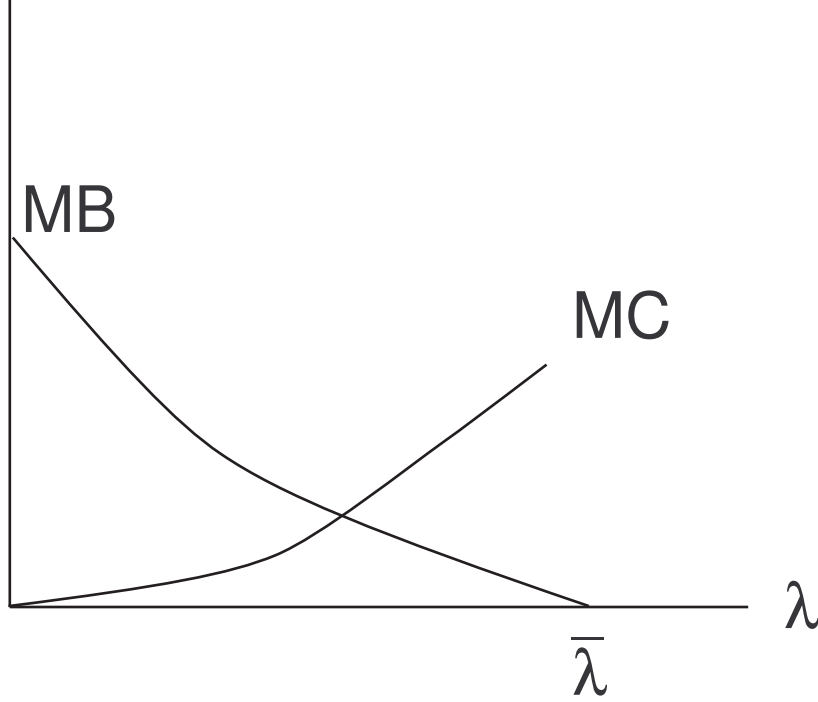


Figure 2: Marginal benefit and marginal cost of increasing enforcement level,  $\lambda$ .

Next, we investigate the relationship between  $\tau_s$  and  $\tau_n$  when the price  $p(\lambda)$  is between the monopoly price  $p_m$  and the competitive price  $c$ . Assume that  $\gamma$  is constant.<sup>8</sup> The optimal 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

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

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:*

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

<sup>8</sup>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  $\alpha$  are constant. We can obtain that  $\gamma$  is equal to  $\frac{1}{\alpha-1}$  which is a constant.

$A_n(\lambda)$  is the intersection of the North's "conditional best response function"<sup>9</sup> 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  $\tau_n$  is a linear function of  $\tau_s$ . 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:

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

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  $\lambda$ .

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  $\tau_n$  and  $\lambda$ . If  $\lambda$  increases, then the RHS of equation (13) will shift down. Moreover, the LHS of equation (13) will shift up if  $\lambda$  goes up. Thus, by Figure 3, it shows that for any given  $\tau_s$ , the patent length of the North is decreasing in  $\lambda$  if Assumption 1 is satisfied.

Propositions 1 and 2 imply that, for any given  $\tau_s$ , 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,  $\tau_s$ , 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.<sup>10</sup> Thus, the North's welfare can be improved.

---

<sup>9</sup>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  $\lambda$  is also an endogenous variable in this model.

<sup>10</sup>In Gilbert and Shapiro (1990), the patent breadth has many different definitions. They indicate 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.

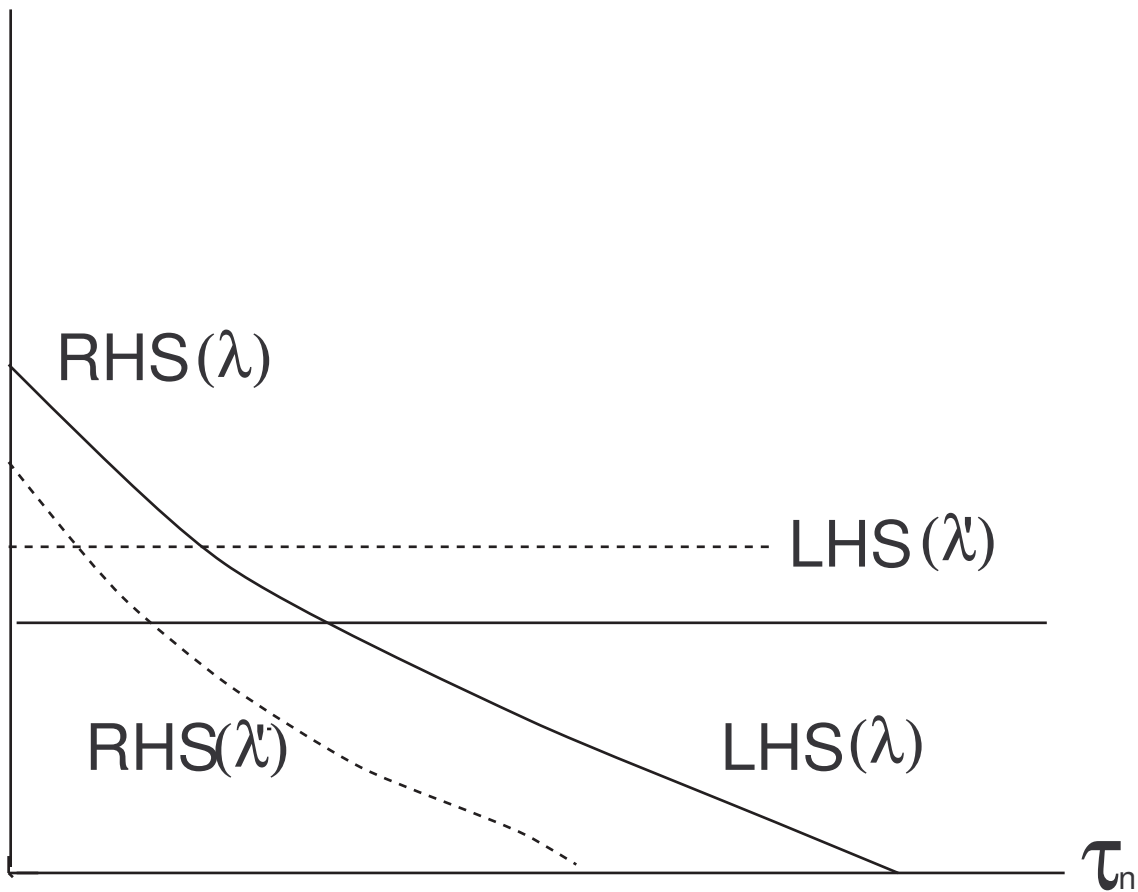


Figure 3: Right hand side (RHS) and left hand side (LHS) of equation (16) where  $\lambda' > \lambda$ .

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

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

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*

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

$A_s(\lambda)$  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  $\gamma$  is constant, there exists a unique pair  $(\tau_n, \tau_s)$  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  $\tau_n$  and  $\tau_s$  are continuous functions of  $\lambda$ , respectively. The continuity of  $\tau_n$  and  $\tau_s$  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  $\gamma$  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  $\gamma$  is constant, then*

$$(i) \frac{\partial \tilde{W}_n}{\partial \tau_n} > 0 \text{ for } \tau_s \in [0, A_s(\lambda)],$$

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

*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  $\gamma$  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  $\lambda$ . Next, we need to show that there exists a  $\lambda^*$  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  $\lambda^*$  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

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

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

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  $(\tau_s^e, \tau_n^e)$  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  $\tau_s^*$  is greater than  $\tau_s^e$  and  $\tau_n^* \leq 1$ . This completes the proof of the Corollary.

### 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  $\lambda$  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  $\lambda$  and  $N$ , increasing  $\tau_n$  and reducing  $\tau_s$  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

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

In order to maximize total welfare, the central planner can choose three variables,  $\lambda$ ,  $\tau_n$ , and  $\tau_s$ . Taking the derivative with respect to  $\lambda$ , we obtain the following

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

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  $\tau_n$  and reducing  $\tau_s$  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:

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

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

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

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  $\lambda$ , the uniform universal standard for IPRs protection will never achieve global social efficiency.*



## 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.

## 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:

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

By equation (9), it yields:

$$\begin{aligned} & (\Delta(p_m) + \pi(p_m))\tau_s(1+M_s)\pi(p_m) \\ &= (\tau_s s(p_m) + (1-\tau_s)s(c))((1+M)\pi(p_m) - \pi(A_s(\lambda)))\gamma. \end{aligned} \quad (21)$$

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

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

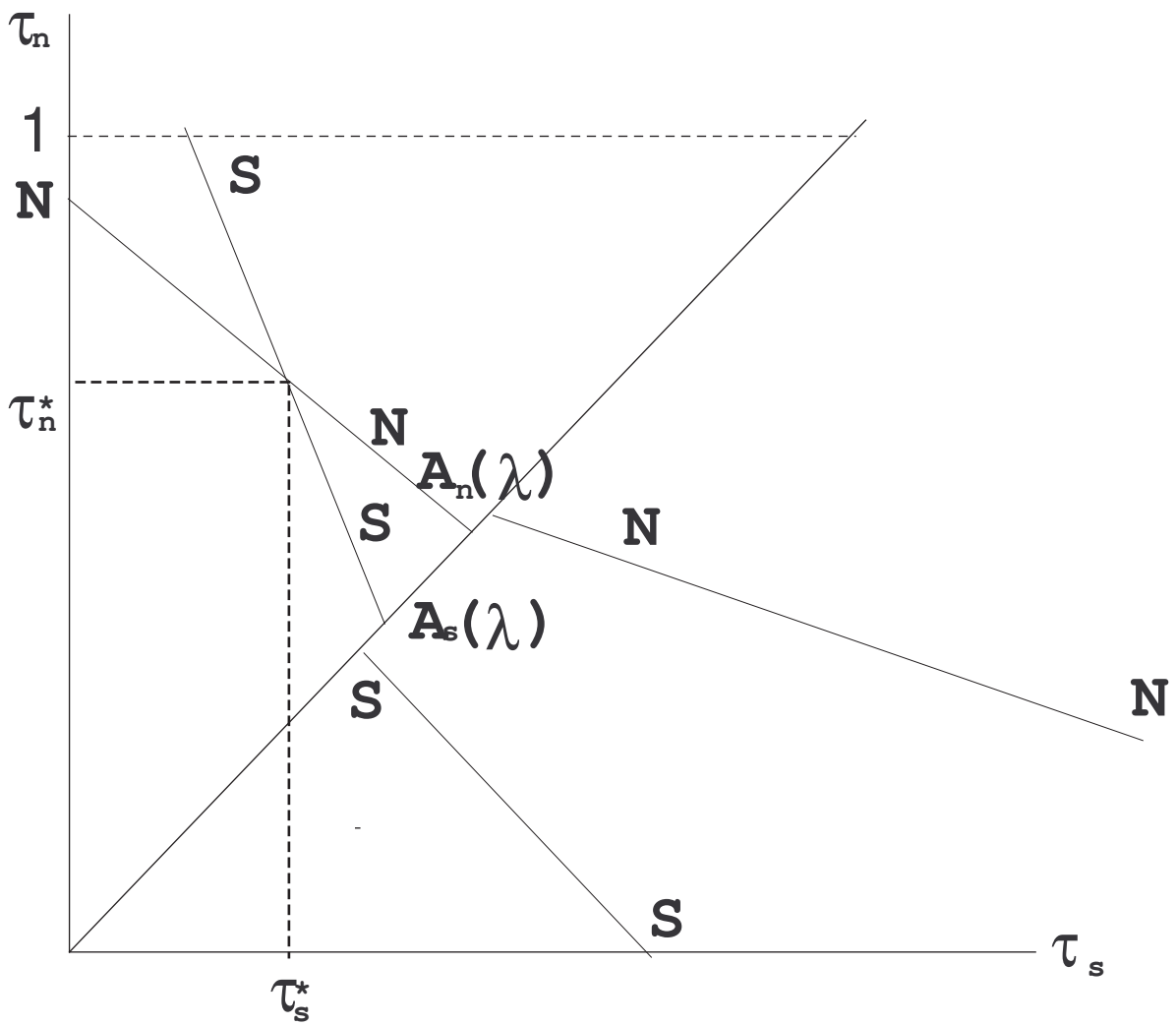


Figure 4: The conditional best response function of the North and South.

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

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

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  $\tau_n \geq \tau_s$ .

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  $B_n^{-1}(1) \geq B_s(1)$ . 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  $\tau_s \in [0, B_n^{-1}(1))$ . Therefore, the unique solution is  $(1, B_s(1))$ .

**Case 3:** If  $A_s(\lambda) < 1$ ,  $B_s(1) > 0$  and  $B_n^{-1}(1) < B_s(1)$ . Since  $B_n^{-1}(1) < B_s(1)$ , there is no solution for all  $\tau_s \in [0, B_n^{-1}(1)]$ .  $B_n(\tau_s)$  is a linear, decreasing function in  $\tau_s$ . Therefore,  $B_s^{-1}(\cdot) - B_n(\cdot)$  is positive when  $\tau_s = B_s(1)$  because  $B_s^{-1}(B_s(1)) = 1$  and  $B_n(B_s(1)) < 1$ . Furthermore,  $B_s^{-1}(\cdot) - B_n(\cdot)$  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  $(\tau_n, \tau_s)$  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  $\tau_n$ ,  $B_s(\tau_n) > 0$  for all  $\tau_n \in (\underline{B}_s^{-1}(0), A_s(\lambda)]$ . Therefore,  $B_n^{-1}(\cdot) - B_s(\cdot)$  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,  $B_n^{-1}(\cdot) - B_s(\cdot)$  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  $(\tau_n, \tau_s)$  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  $\Pi$  is:

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

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:

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

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  $\tau_n$ , 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  $\lambda_s, \tau_s$ , 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  $\tau_n$  is equal to  $\tau_s$ . 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:

$$\begin{aligned} &(\Delta - M(\pi - \pi(p_s(\lambda_s))))(\tau_n(1 + M)\pi) \\ &= (\tau_n s(p_m) + (1 - \tau_n) s(c))(\pi + M(\pi - \pi(p_s(\lambda_s))))\gamma, \end{aligned}$$

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:

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

Since  $F(\cdot)$  is a decreasing function in  $\tau$ , 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  $\tau_n$  and  $\tau_s$  yields the slope of the conditional best response function of the South:

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

Differentiating equation (27) with respect to  $\tau_n$  and  $\tau_s$  yields the slope of the North's "best response function":

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

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  $(\tau_n, \tau_s)$  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

- 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.
- 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.
- Deardorff, Alan V. (1992), "Welfare Effects of Global Patent Protection," *Economica*, 59, 35-51.
- Diwan, Ishak and Dani Rodrik (1991), "Patents, Appropriate Technology, and North-South Trade," *Journal of International Economics*, 23, 79-90.
- The Economist (1996), "Stolen Melodies," May 11; 339, 64.
- 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.
- Gilbert, Richard and Carl Shapiro (1990), "Optimal Patent Length and Breadth," *RAND Journal of Economics*, 21, 106-112.
- Grossman, Gene M. and Elhanan Helpman (1994), "Protection for Sale," *American Economic Review*, 84, 833-850.
- Grossman, Gene M. and Edwin L.C. Lai (2004), "International Protection of Intellectual Property," *American Economic Review*, 94, 1635-1653.
- 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.
- 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.
- Maskus Keith E., (2000), *Intellectual Property Rights in the Global Economy*, Washington, DC: Institute for International Economics.

- Mitra, Devashish (1999), "Endogenous Lobby Formation and Endogenous Protection: A Long-Run Model of Trade Policy Determination," *American Economic Review*, 89, 1116-1134.
- Nordhaus, William D. (1969), *Invention, Growth, and Welfare: A Theoretical Treatment of Technological Change*, Cambridge, Mass., M.I.T. Press.
- Scherer, Frederic M. (1980), *Industrial Market Structure and Economic Performance*, Chicago: Rand McNally College Pub. Co.
- Scotchmer, Suzanne (2004), "The Political Economy of Intellectual Property Treaties," *Journal of Law, Economics, and Organizations*, 20, 415-437.
- Spors, Kelly K. (2004), "Canadian Drugs Are Getting Cheaper; Online Pharmacies Catering to American Cut Costs by Finding New Suppliers," *Wall Street Journal*, July 22.
- State Legislatures (2004), "Canadian Drugs: Panacea or Problem?" June; 30, 6.
- Weil, Elizabeth (2004), "Grumpy Old Drug Smugglers," *New York Times Magazine*, May 30, 42-45.



## CHAPTER IV

### WHY DOES THE U.S. PREVENT PARALLEL IMPORTS?

#### Introduction

Parallel imports, also called “gray-market imports,” are authorized for their first sale, not counterfeited or pirated, but imported by unauthorized resellers. Those commodities are identical to legitimate products except that they may be packaged differently and may not carry the original manufacturer’s warranty. It occurs when one manufacturer owns the patents or national trademarks in several countries. In order to make its maximal profits, the manufacturers would rationally grant low prices in low-income countries or in the markets with elastic demand and set high prices in high-income countries or in less-elastic-demand markets.

One of the most important issues of parallel trade is the welfare effects. Varian (1985) and Malueg and Schwartz (1994) indicate that parallel trade would result in uniform pricing internationally and would reduce global welfare because parallel trade could make some markets unserved. Maskus and Chen (2002, 2004) develop a model of parallel trade and vertical price control in which a manufacturer sells abroad through a retailer and tries to achieve the profit-maximizing price by setting its wholesale price and licence fee. It concluded that banning parallel trade is always beneficial to the monopolist producer, but could raise or reduce global surplus. Chen and Maskus (2005) generalize the results of Maskus and Chen (2004) and find that open parallel trading regimes may improve welfare within regional trade agreement. Richardson (2002) indicates that, in a Nash equilibrium, no importing countries would not permit parallel importing because they are discriminated against in its absence.

The existing literature on parallel trade always conclude that the welfare of importing country is seemingly improved because the price of parallel imports decreases. However, there exists the phenomenon that some developed countries, such as the U.S., do not like to permit parallel imports. For example, we can easily find the following messages in international edition textbook: “*This International Edition is not for sale in the United States of America, its dependencies or Canada.*”<sup>1</sup> This message or similar statements usually appear on the copy right page of the International

---

<sup>1</sup>See Nicholson (1985).

Edition textbook.<sup>2</sup> This indicates that the U.S. does not allow the international edition textbooks to circulate in the U.S. and Canada; this is true even for used books.<sup>3</sup>

Policy toward parallel trade varies worldwide. In general, policies can be divided into three categories, national exhaustion, regional exhaustion, and international exhaustion. National exhaustion means that intellectual property rights (IPRs) end upon original sale within a nation but IPRs owners may prevent parallel trade with other countries. International exhaustion means that IPRs are exhausted upon original sale and parallel trade is permitted. Regional exhaustion means that IPRs are completed within a group of countries, allowing parallel trade among them, but are not exhausted outside the region.

Some countries' policies are summarized in Table 1. It shows that Japan permits parallel imports unless they are barred by contract or its original sale is regulated by price-control. Thus, we conclude that Japan is regularly open for parallel trade. Moreover, the European Union policy is regional exhaustion. It also shows that the U.S. policy is generally national exhaustion. The U.S. policy permits trademark owners to block parallel imports with some exceptions, such as when both the foreign and the U.S. companies belong to the same entity or when the foreign and the U.S. companies are in a parent-subsidiary relationship, which means the U.S. significantly restricts parallel trade. Therefore, parallel trade in the U.S. is more restrictive than in other developed countries.

In this paper, we would like to ask: What is the reason for the difference between U.S. and Japanese policy? If parallel importing is always in favor of the importing country, it is difficult to explain why the U.S. would not like to allow parallel importing. Most of the existing literature on parallel trade (i.e. Malueg and Schwartz (1994), Maskus and Chen (2002, 2004), Chen and Maskus (2005)) usually employed a single product model to demonstrate how parallel trade impacts welfare. To date, to the best of my knowledge, no paper discusses how parallel trade affects the incentives of innovation. Furthermore, most papers on parallel trade use the global welfare approach to explain why some countries prevent parallel trade, which I think that is unreasonable. They do not explain by the viewpoint of importing country.

The purpose of this paper is to propose a simple two-country model with innovation to analyze the welfare effects of parallel imports. We try to use this model to explain why some of the

---

<sup>2</sup>In McGraw-Hill international edition textbook, the following statement, “*This book cannot be re-exported from the country to which it is consigned by McGraw-Hill.*,” appears.

<sup>3</sup>For example, we can not sell the international edition used books on [www.amazon.com](http://www.amazon.com), [ebay.half.com](http://ebay.half.com) or other similar websites.

Table 1: Summary of IPR Exhaustion Regimes.

Country	Trademarks	Patents	Copyrights
European Union	Community exhaustion	Community exhaustion	Community exhaustion
USA	National exhaustion limited by consumer control and no consumer confusion	National exhaustion	National exhaustion
Japan	International exhaustion unless agreed by contract or original sale is price-controlled	Same as trademarks	Same as trademarks
Australia	International exhaustion	National exhaustion unless sold by patent owner without clear restrictions	National exhaustion exception for compact discs and books

Sources: Maskus (2000a)

developed countries, such as the U.S., will not allow parallel imports. We utilize a simulation approach to investigate the welfare effects of parallel trade on the importing country and global welfare.

The remainder of this paper is organized as follows: In section 2, we demonstrate that the welfare of the importing country improves in a conventional two-country model with linear demand. In section 3, we set up a two-country model of parallel trade, in which innovation is incorporated. We employ a simulation approach to explain why the U.S. does not allow parallel imports. Section 4 gives conclusions.

### The Single Product Case

In this section, we describe a simple, conventional, two-country model of parallel trade with a single product. Consider an economy in which there are two countries, North and South, denoted  $n$  and  $s$ , and a monopolistic manufacturer. The monopolist manufacturer produces and sells identical products in both countries.

The inverse demand function in market  $i$  is

$$p_i(q_i) = a_i(1 - q_i), \quad i = n, s. \quad (1)$$

where  $p_i$  and  $q_i$  is the price and consumption in country  $i$ , respectively.<sup>4</sup> This kind of linear demand function has equal horizontal intercept, at 1, but different vertical intercepts,  $a_i$ . Furthermore, we assume that  $a_n = 1 + x$  and  $a_s = 1 - x$ , where  $x < 1$ .

The monopolist determines the price in both markets to maximize its profits. Suppose that the markets are segmented. The producer can price discriminate in the two countries so as to make its maximal profit. This is third degree price discrimination. This means that the producer can set different prices in different markets according the demand elasticity of market. Suppose that the marginal production cost is constant and the same in the two countries. For simplicity, I assume that marginal production cost is zero, without loss of generality. Therefore, the monopoly price in the North,  $p_n^m = \frac{1+x}{2}$ , would be higher than the price in the South's market,  $p_s^m = \frac{1-x}{2}$ , which implies that the North is the importing country if parallel trade exists.

Since there exists a price difference, there exist incentives for arbitrage. Parallel imports are brought into the North from the South. For simplicity, we assume that there is no transaction

---

<sup>4</sup>The specification of this linear demand function follows Malueg and Schwartz (1994).

cost. Since arbitrage is unlimited, the market of the parallel imports in the North is competitive.<sup>5</sup> Therefore, the optimization problem of the monopolist is to choose the price  $p$  to maximize

$$\pi(p) = p\left[\left(1 - \frac{p}{1+x}\right) + \left(1 - \frac{p}{1-x}\right)\right], \quad (2)$$

where  $p$  is the uniform price,  $\left(1 - \frac{p}{1+x}\right)$  and  $\left(1 - \frac{p}{1-x}\right)$  are the consumptions in the North and the South, respectively.

The monopolist chooses the uniform price  $p$  to maximize its profit. Taking derivative with respect to  $p$ , we obtain the following:

$$\frac{\partial \pi(p)}{\partial p} = 2 - \frac{2p}{1+x} - \frac{2p}{1-x}.$$

The optimal price can be solved by the first order condition. We obtain

$$p^* = \frac{1-x^2}{2}. \quad (3)$$

Note that  $p^*$  is greater than  $p_s^m$ , but less than  $p_n^m$ .

Next, we investigate the welfare effects of parallel trade. First, we want to show that parallel trade will improve global welfare. When the two markets are segmented, we can get that the sum of consumers' surplus and producer's surplus in the North is  $\frac{3(1+x)}{8}$  and  $\frac{3(1-x)}{8}$  for the South. Thus, we find that the global welfare is  $\frac{3}{4}$ .

If parallel trade exists, the optimal price,  $p^*$ , is  $\frac{1-x^2}{2}$ . Thus, the sum of consumers' surplus and producer surplus in the North is  $(1+x)^2(3-x)/8$  and  $(1-x)^2(3+x)/8$  in the South. Therefore, global welfare is  $\frac{3}{4} + \frac{x^2}{4}$ . Thus, we can conclude that the global welfare in the case of parallel trade is greater than in the case of segmented markets. Furthermore, global welfare is increasing in  $x$ .

If the monopolist is from the North, the welfare of the South decreases when parallel trade happens since the price of the product in the South goes up. Thus, the welfare of the North improves. Moreover, if the monopolist is from the South, it is easy to show that the welfare of the North improves because the price in the North decreases. Therefore, we obtain the following conclusions:

**Proposition 1:** (a) *Global welfare is improved by allowing parallel trade.* (b) *The welfare of the North increases if parallel trade allowed.*

---

<sup>5</sup>There exists an alternative approach of analysis in distribution of parallel imports. Maskus and Chen (2002, 2004), and Chen and Maskus (2005) propose a model of duopoly, in which parallel importer and the producer compete in the importing market. However, we think that similar results for the issue of this paper apply for their model. Furthermore, unlimited arbitrage and competitive market make this analysis simpler.

Proposition 1 indicates that the welfare of importing country improves when the demand functions are linear in the case of a single product model. Actually, Tirole (1988) indicates that when the demand functions are linear, the welfare of third price discrimination is lower than uniform pricing. Thus, we can generalize the above linear demand model to different market size and the results remain the same. Furthermore, we also can conclude that the importing country increases their welfare by parallel importing in Varian (1985) and Malueg and Schwartz (1994), in which a multi-country model is employed. For simplicity, we just focus on the case of a two-country model with the same market size.

In the next section, we investigate the welfare effects of parallel trade in the case of multiple products and try to explain why some developed countries do not allow parallel imports.

### **The Welfare Effects of Parallel Trade Incorporated with Innovation**

In this section, we develop a two-country model of parallel trade, in which innovation is incorporated. Consider an economy with two countries, named South ( $s$ ) and North ( $n$ ), and a firm that develops new products in the North. The model follows Deardorff (1992) and Scotchmer (2004) by considering a two-stage model of innovation. It is assumed that there are two sectors, a homogenous good sector and a differentiated products sector. In the first stage, 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 their country but the consumers have different preferences in the South and the North. Each consumer chooses  $z$  and  $x(j)$  to maximize his utility which is given by:

$$\begin{aligned} & \int_0^N u_i(x_i(j))dj + z_i, \\ \text{subject to } & \int_0^N p_i(j)x_i(j)dj + z_i = Y_i, \end{aligned} \tag{4}$$

where  $j$  is the index of differentiated goods,  $N$  is the measure of new differentiated products in the North,  $x_i(j)$  is the consumption of a differentiated good  $j$  in country  $i$ ,  $z_i$  is the consumption of the homogenous good, and  $Y_i$  is individual income in country  $i$ . We assume that  $u' > 0$  and  $u'' < 0$ . From the first order conditions, it yields  $x_i(j) = x_i(p(j))$  where  $x_i = (u'_i)^{-1}$ . The demand for the numeraire good is given by:

$$z_i = Y_i - p_i(j)x_i(p_i(j)).$$

The indirect utility function can be written as:

$$U = \int_0^N s_i(p_i(j))dj + Y_i,$$

where  $s_i(p_i(j)) = u_i(x_i(p_i(j))) - p_i(j)x_i(p_i(j))$  is the consumer surplus associated with a representative differentiated product,  $Y_i$  is consumer's income. The two-stage 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). For the sake of comparison, we assume that the demand function for each differentiated product is just like equation (1) in a different country.

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. Just like section 2, we assume that  $c$  is equal to zero.

In the following, we employ a simulation approach to investigate the welfare effects of parallel trade for the North and the global welfare effects.

The innovation cost function is specified as follows:

$$C(N) = N^\alpha, \tag{5}$$

where  $\alpha$  is greater than one.<sup>6</sup> According to equation (5), we know that the innovation cost is convex in  $N$ . We also denote that

$$\gamma = C'(N)/(NC''(N))$$

is the elasticity of innovation with respect to an increase in the profit from innovation, which follows Grossman and Lai (2004). This specification of innovation cost function has a constant elasticity of innovation, which has the same properties as the innovation cost function in Grossman and Lai (2004).

The problem of this firm is to choose the number of the new differentiated products,  $N$ , and the price in both markets to maximize its profit. Since the  $p_s$  and  $p_n$  are determined as in Section

---

<sup>6</sup>Otherwise, the marginal innovation cost and the average innovation cost are going down, which means that the number of innovation becomes infinite.

2, the optimal  $N^*$  is determined by marginal profit and the marginal innovation cost of innovating a new differentiated good. Thus,

$$C'(N^*) = \pi_n(p_n) + \pi_s(p_s).$$

where the right hand side is the marginal profit generated by an additional product.

The welfare of the North is composed of the profit of the firm in the North and the South, the consumers' surplus in the North, and the innovation cost.

$$W_n(N) = N(\pi_n(p_n) + \pi_s(p_s) + s_n(p_n)) - C(N).$$

Moreover, since the firm is in the North, the welfare of the South is only the consumers' surplus generated in the South.

$$W_s(N) = N s_s(p_s).$$

Furthermore, global welfare is defined as the sum of the welfare of the North and the South, i.e.  $W = W_n + W_s$ .

We examine the welfare effects of parallel trade by numerical simulation for  $x \in [0, 0.5]$  for different innovation cost functions.<sup>7</sup> Since we obtain similar results for different cases of the innovation cost function, we just illustrate two cases of  $\gamma$  in Figure 1 and Figure 2. In Figure 1 and Figure 2, the x-axis is defined as  $x$  and the y-axis is defined as the relative welfare of the North and relative global welfare, .

Two cases of elasticity of innovation,  $\gamma = 5$  and  $\gamma = \frac{10}{11}$  are illustrated in Figure 1 and Figure 2, respectively. The former represents the cases of high elasticity of innovation and the latter represents the cases of low elasticity of innovation. “ $W_n$  ratio” is defined as the North's welfare when parallel trade happens divided by the welfare of the North when parallel trade does not exist. Thus, “ $W_n$  ratio” is greater than 1 means that the North's welfare improves when parallel trade exists. Analogously, the  $GW$  ratio is the global welfare ratio, which is greater than 1 means the global welfare improves if parallel trade happens.

Figure 1 shows that two North's welfare is concave in  $x$ . It is increasing in  $x$  first, and then decreasing in  $x$ . We can find that there exists a  $x = 0.117$ , which makes the North obtain the maximal welfare improvement ( $W_n$  ratio) from parallel imports. For  $x \in [0.117, 0.5]$ , the North welfare is decreasing in  $x$  and there exists  $x = .231$  such that the North welfare of parallel trade is equal the welfare if parallel trade does not exist. This means that if  $x$  is big enough, parallel

---

<sup>7</sup>Since there will not exist parallel trade for the case of  $x > 0.5$ , we just focus on the case of  $x \in [0, 0.5]$ .



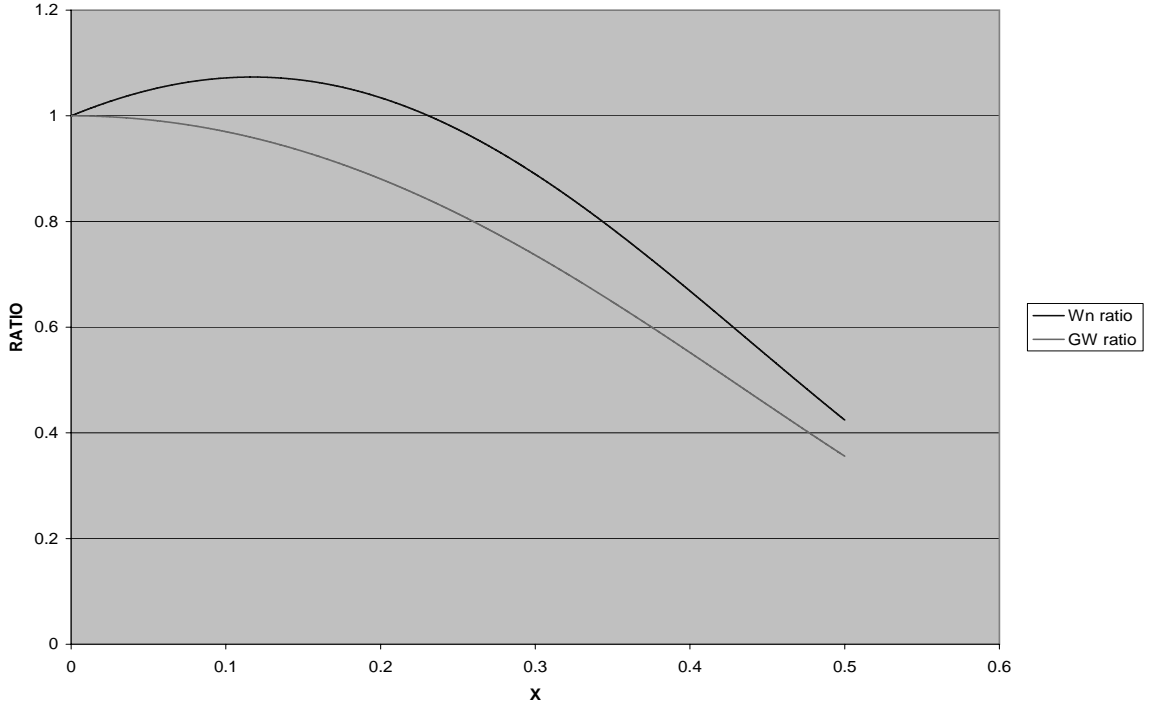


Figure 1: Relative welfare,  $\gamma = 5$ .

importation is not beneficial to the importing country. Furthermore, it also shows that the global welfare ratio is less than or equal to 1, and it is decreasing in  $x$  for all  $x$  when  $\gamma = 5$ . This means that global welfare decreases because of parallel importing when the elasticity of innovation is relative high.

Figure 2 shows the simulation results of the relative welfare of the North and the global welfare for the case of  $\gamma = \frac{10}{11}$ . It indicates that the welfare of North and global welfare are analogous to the case of  $\gamma = 5$ . The main difference is that global welfare could improve because of parallel trade. We also can find that the improvement of Global welfare is very small. Furthermore, it also shows that the North welfare of parallel imports is always greater than the welfare of banning parallel trade.

We summarize our simulation results by Figure 3 for all different elasticity of innovation. The y-axis is defined as the elasticity of innovation,  $\gamma$  and the x-axis is  $x$ , the difference between the slope of demand function in different market. Figure 3 contains three regions:  $A$ ,  $B$ , and  $C$ . The

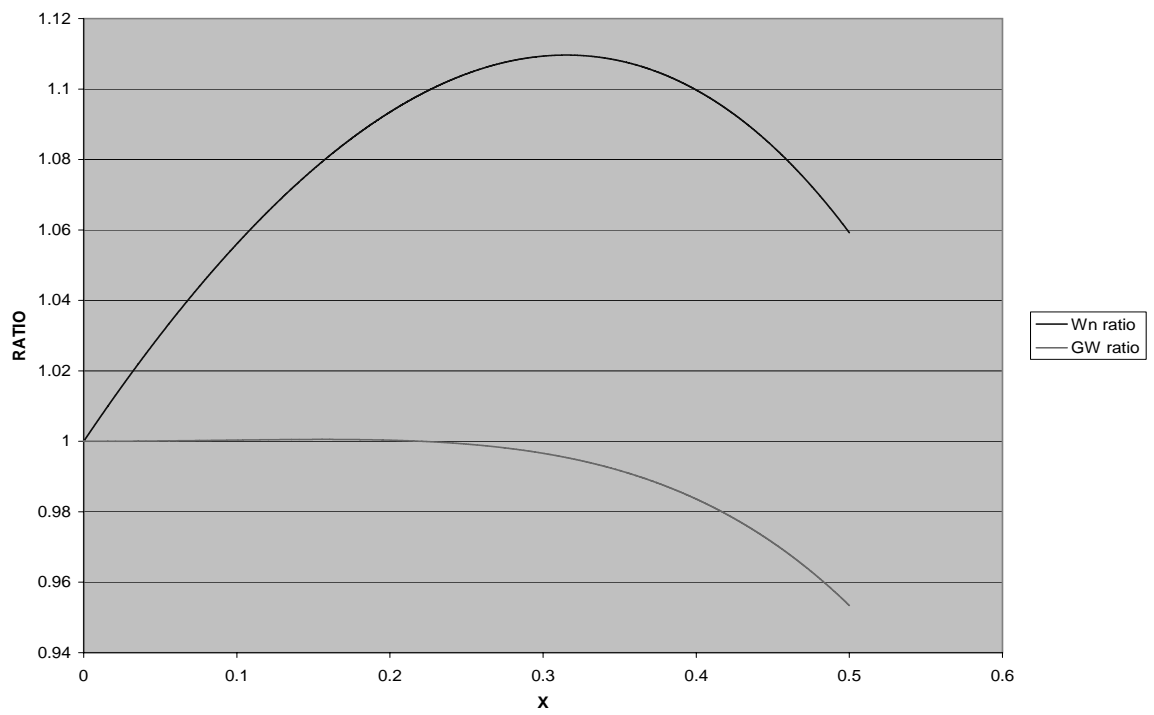


Figure 2: Relative welfare,  $\gamma = \frac{10}{11}$ .

case of  $(x, \gamma)$  in  $A$  means that parallel trade will reduce the North's welfare as well as global welfare. Figure 3 shows when both  $\gamma$  and  $x$  are high, parallel trade hurts the North's and global welfare.

Region  $C$  means that both global welfare and the North's welfare improve. Figure 3 shows that when elasticity of innovation is low, parallel trade will increase the North's and global welfare. Region  $B$  means the North's welfare increases but global welfare decreases if parallel trade exists.

Therefore, we summarize the results as the following proposition:

**Proposition 2:** *(a) When elasticity of innovation is high, the innovating country welfare increases in  $x$  first, and it decreases in  $x$  when  $x$  is getting large. Global welfare decreases in  $x$ . (b) When elasticity of innovation is low, the innovating country welfare and global welfare increase in  $x$  first, and they decrease in  $x$  when  $x$  is getting large. (c) When elasticity of innovation and  $x$  are high, parallel trade decreases the innovating country welfare as well as global welfare. (d) When elasticity of innovation is low, parallel trade will improve the innovating country welfare as well as global welfare. However, the increase in global welfare is limited.*

Proposition 2 indicates that parallel imports can always make the North's welfare improve when the difference between the North and the South is small. On the other hand, if the two markets differ so much, parallel imports will make the North worse. Thus, Proposition 2 can give us a reason to explain why some countries permit parallel importing, but some countries do not.

The other finding is that global welfare decreases for the case of high elasticity of innovation for any  $x$ . This finding is quite different from the results in section 2, in which global welfare always increases for any case of linear demand function in a single product model. The reason is because the number of innovations decreases due to the reduced profit. For the case of low elasticity of innovation, global welfare could improve when the difference between the two markets is not so big. The reason is because parallel trade reduces the profit of monopolist. Thus, parallel trade reduce the incentives for innovation. Therefore, the number of innovations decreases. Since parallel trade in a single product case will improve the welfare of the importing country and global welfare which is indicated in section 2, the welfare effects of parallel trade become ambiguous.

## Conclusions

We introduced a model of parallel trade in which innovation is incorporated and tried to explain why there exists huge difference of parallel trade policy among countries. The reason is because

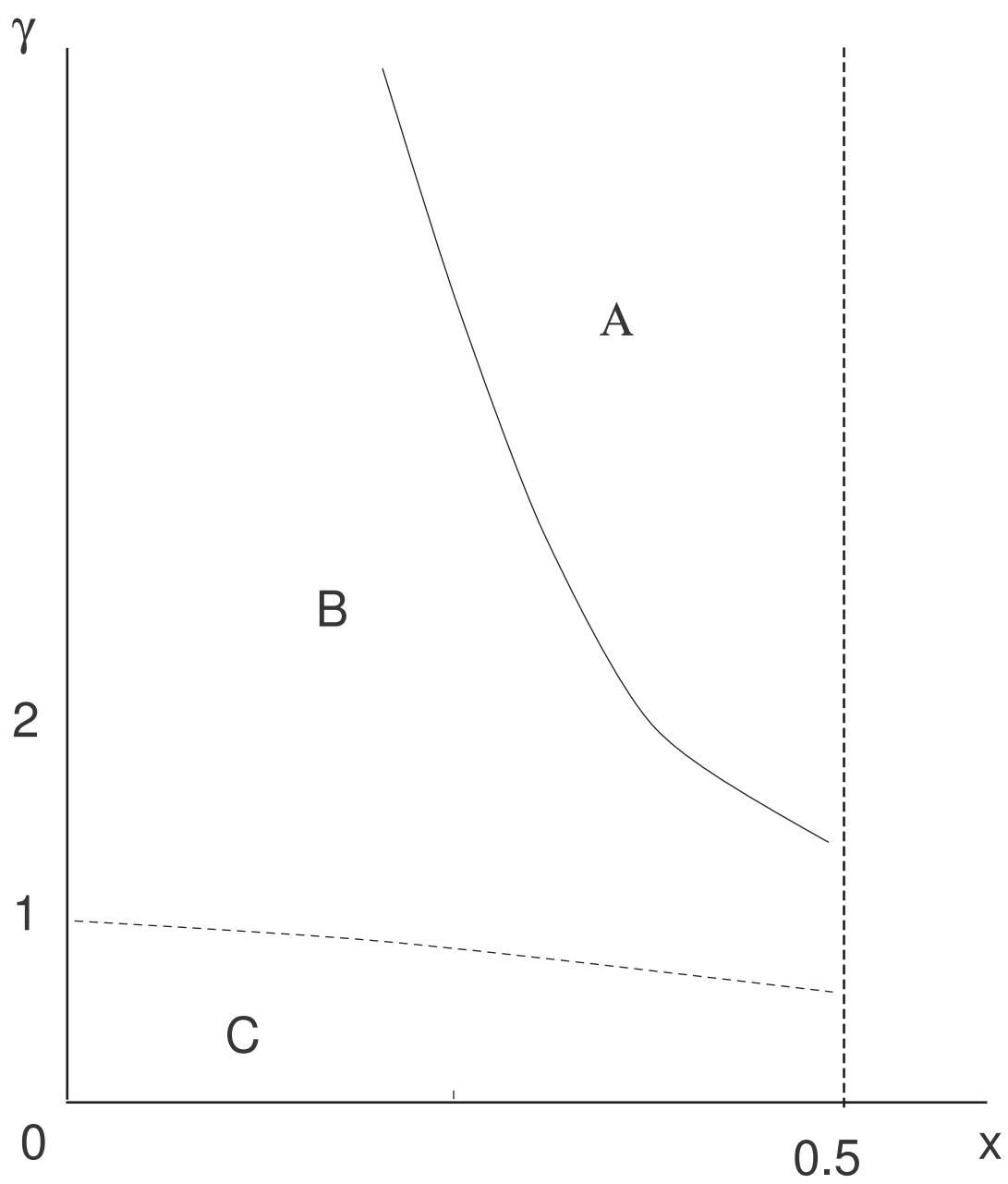


Figure 3: The relation between elasticity of innovation,  $\gamma$ , and  $x$ .

if parallel imports are allowed, it will reduce the incentives for innovation of the innovators, then reduce the number of new innovation. The results show that the welfare effect is not only related to the difference of the two markets, but is also related to the elasticity of innovation. Parallel trade improves the welfare of importing country. However, parallel trade also reduces the incentives to innovate. Thus, the welfare effect of parallel imports is ambiguous. It explains that some countries parallel import and some do not. Furthermore, we also show that when elasticity of innovation and  $x$  are high, parallel trade decreases the innovating country welfare as well as global welfare. When elasticity of innovation is low, parallel trade will improve the innovating country welfare as well as global welfare. However, the increase in global welfare is limited.

## References

- Chen, Yongmin and Keith E. Maskus (2005), "Vertical Pricing and Parallel Imports," *Journal of International Trade and Economic Development*, 14, 1-18.
- Deardorff, Alan V. (1992), "Welfare Effects of Global Patent Protection," *Economica*, 59, 35-51.
- Grossman, Gene M. and Edwin L.C. Lai (2004), "International Protection of Intellectual Property," *American Economic Review*, 94, 1635-1653.
- Malueg, David A. and Marius Schwartz (1994), "Parallel Imports, Demand Dispersion, and International Price Discrimination," *Journal of International Economics*, 37, 167-195.
- Maskus, Keith E. (2000a), *Intellectual Property Rights in the Global Economy*, Washington, DC: Institute for International Economics.
- Maskus, Keith E. (2000b), "Parallel Imports," *The World Economy*, 23, 1269-1284.
- Maskus, Keith E. and Yongmin Chen (2002), "Parallel Imports in a Model of Vertical Distribution: Theory and Evidence and Policy," *Pacific Economic Review*, 7, 319-334.
- Maskus, Keith E. and Yongmin Chen (2004), "Vertical Price Control and Parallel Imports: Theory and Evidence," *Review of International Economic*, 12, 551-570.
- Nicholson, Walter (1985), *Microeconomic Theory*, Third Edition, The Dryden Press.
- Richardson, Martin (2002), "An Elementary Proposition Concerning Parallel Imports," *Journal of International Economics*, 56, 233-245.
- Scotchmer, Suzanne (2004), "The Political Economy of Intellectual Property Treaties," *Journal of Law, Economics, and Organizations*, 20, 415-437.
- Tirole, Jean (1988), *The Theory of Industrial Organization*, Cambridge, MA, The MIT press.
- Varian, Hal R. (1985), "Price Discrimination and Social Welfare," *American Economic Review*, 75, 870-875.