Why Does the US Prevent Parallel Imports?

Yih-Ming Lin
Department of Applied Economics, National Chiayi University, Chiayi, Taiwan
Email: yxL173@mail.nctu.edu.tw

Received November 1, 2013; revised December 1, 2013; accepted December 8, 2013

ABSTRACT
Most studies on parallel trade conclude that parallel imports, in general, benefit the importing country because it lowers the price of parallel imports and benefits to the consumers in the importing country. Richardson [1] explicitly indicates that there is no importing country not to permit parallel imports because they are discriminated against in its absence. However, an obvious counter example is observed in the US. In this paper, we propose a two-country model of parallel trade with innovation to explain why some countries, such as the US would like to prevent parallel imports. We show that the elasticity of innovation is crucial to the welfare of importing country and global welfare.

Keywords: Parallel Imports; Welfare; Innovation

1. Introduction
Most of the previous literatures, such as Malueg and Schwartz [2], Richardson [1], Maskus and Chen [3,4], and Chen and Maskus [5], on parallel trade conclude that parallel imports, in general, benefit the importing country because it lowers the price of parallel imported goods and benefits to the consumers. Among them, Richardson [1] explicitly indicates that there is no importing country not to permit parallel imports because they are discriminated against in its absence. However, an obvious counter example is observed in the US. For instance, the message “This International Edition is not for sale in the United States of America, its dependencies or Canada.” is common to be found in a copyright page or cover of the International Edition textbook, which indicates that the US does not allow new or used international edition textbooks to be circulated within the country. Therefore, this phenomenon raises the question: If parallel import is always in favor of the importing country, then why would the US government prevent parallel import?

Previous discussion on whether parallel trade should be prevented includes Maskus [6,7], Malueg and Schwartz [2], and so on. Maskus [6] summarizes the arguments which are in favor of preventing parallel trade. Among them, the most well-known explanation is that third-degree price discrimination may be beneficial to the global welfare, which will not necessarily happen but depend on whether the global output increases or not under each case (Varian [8]). Another argument opposing to parallel trade proposed by Malueg and Schwartz [2] is that some markets will become unserved due to a uniform price, the direct consequence of parallel trade. However, the above arguments seem hard to explain why a country is willing to block parallel import if its government concerns its own welfare, instead of the global welfare. In particular, the above discussions are concentrated on the benefits of the entire world, instead of that of an individual country.

In this paper, we build in a two-country model with innovation to analyze the welfare effects of parallel imports from the viewpoint of an importing country. The welfare effects of parallel trade on the importing country and global welfare are also investigated by a simulation approach. We investigate the welfare effects and how parallel trade relates to the incentive of product innovation by employing a differentiated-goods model, which extends the models of Li and Maskus [9], Li and Robles [10], and Li [11,12]. It distinguishes from the previous studies, while only a single product model is used to demonstrate how parallel trade affects the welfare in most of existing literatures. See Malueg and Schwartz [2], Maskus and Chen [3,4], and Chen and Maskus [5] for examples.

The remainder of this paper is organized as follows. In Section 2, we demonstrate a conventional two-country model.
model with linear demand. Following Section 2, we set up a two-country model of parallel trade incorporated with innovation to explain why a country would like to prevent parallel imports in Section 3. Section 4 gives conclusions.

2. 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 with two countries, North and South, denoted by \( n \) and \( s \), and a monopolistic manufacturer. The monopolist manufacturer produces and sells identical products in both countries. Further, we follow Malug and Schwartz [2] to assume that the demand function is linear such that the inverse demand function of the market \( i \) has the form

\[
p_i(q_i) = a_i(1-q_i), \quad i = n, s,
\]

where \( p_i \) and \( q_i \) represent the price and consumption in country \( i \), respectively. Both countries have horizontal intercepts at 1, but different vertical intercepts at \( a_i \). Without of loss of generality, we may 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 so that the producer can set different prices according to the demand elasticity in both countries in order to make its profit maximized, which is a standard case of third degree price discrimination. Moreover, suppose that the marginal production cost is constant and the same in the two countries. Without loss of generality, we may assume that marginal production cost is zero and there is no transaction cost for simplicity. Therefore, the monopoly price in the North, \( p^*_n = \frac{1+x}{2} \), is higher than the price in the South’s market, \( p^*_s = \frac{1-x}{2} \), which implies that the North is the importing country if parallel trade is exhibited.

As long as there exists a price difference, there are incentives for arbitrage. If parallel trade is allowed, the optimization problem of the monopolist is

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

where \( p \) is the uniform price, \( 1-p/(1+x) \) and \( 1-p/(1-x) \) are the quantities consumed in the North and the South, respectively. Taking derivative with respect to \( p \), we obtain

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

and the optimal price is \( p^* = \left( 1-x^2 \right)/2 \).

Note that \( p^* \) is greater than \( p^*_n \), but less than \( p^*_s \).

Next, we investigate the welfare effects of parallel trade. First, assume that the two markets are segmented, and let \( W_i \) represent the welfare of country \( i \) (\( i = n, s \)), which is equivalent to the sum of consumers’ and producer’s surplus of the country. Then \( W_n = 3(1+x)/8 \) and \( W_s = 3(1-x)/8 \) and the global welfare \( W_n + W_s = 3/4 \). Next, let’s consider the case where parallel trade is allowed. Solving the firm’s problem gives the optimal price \( p^* = (1-x^2)/2 \) and \( W_n = (1+x)^2(3-x)/8 \) and \( W_s = (1-x)^2(3+x)/8 \), respectively. The global welfare \( W_n + W_s = 3/4 + x^2/4 \) in this case. Consequently, we can conclude that the global welfare in the case of parallel trade is greater than in the case of segmented markets. In particular, global welfare is increasing in \( x \).

Furthermore, if the monopolistic manufacturer is from the North, the welfare of the South decreases once parallel trade is allowed which makes the price of goods higher in the South. Consequently, 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 can conclude that global welfare and the welfare of the North improve if parallel trade is allowed. Consequently, we may generalize the above linear demand model to different market size and can still conclude the same for importing country. In particular, the above conclusion for importing country holds under the model of Varian [8] as well as the multi-country model of Malug and Schwartz [2]. However, for simplicity, we shall still focus on the case of a two-country model with the same market size but generalize our analysis to a multiple-products setting.

3. The Welfare Effects of Parallel Trade Incorporated with Innovation

In this section, we develop a two-country model with innovation. Similar to the previous setting, we consider an economy with two countries, named South (\( s \)) and North (\( n \)), and a firm that develops new products in the North. But, here we assume that there are two sectors in a firm, including a homogenous good sector and a differentiated products sector. Following Deardorff [13] and Scotchmer [14], innovation is incorporated in the model through a two-stage model; i.e., the firm chooses the number of differentiated products to be produced in the first stage and then sells the products in both countries in the next stage.

Consumers in the same country are assumed to have identical preferences but consumers in different countries may have different preferences. Consumer \( i \) chooses \( z \) and \( x(j) \) to maximize his utility function \( u_i(z) \), where \( u'_i(z) > 0 \) and \( u''_i(z) < 0 \) and \( i = n, s \). The problem can be written as
\[ \text{Max} \int_0^N u_i(x_i(j))dj + z_j, \] (2)

Subject to \( \int_0^N p_i(j) x_i(j) dj + z_j = Y_i, \)
where \( j \) is the index of differentiated goods, \( N \) is a measure of 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 \( Y_i \) represents the income of an individual in country \( i \).

Solving the first order condition of the consumer’s problem yields \( x_i(j) = x_i(p_i(j)) \), where \( x_i = (u_i')^{-1} \), which suggests that the demand for the numeraire good is \( z_i = Y_i - \int_0^N p_i(j)x_i(j) dj + z_j \). Furthermore, 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.

The equilibrium of the two-stage model can be thought of as that representing the steady state of an infinite horizon general equilibrium model in which a firm innovates in every period and products have an exogenously given useful life, as has been shown by Grossman and Lai [15]. In order to make a comparison, we assume that the demand function for each differentiated product is the same as that in Equation (1) for a different country.

For the monopolistic manufacturer, we may, without loss of generality, assume that the production of one unit of all varieties of differentiated products has the same constant marginal cost \( c \), where \( c \) is set to be 0, similar to the assumption imposed in the previous section. Furthermore, we follow Grossman and Lai [15] and assume that the innovation cost function is specified as
\[ C(N) = N^\alpha, \] (3)
where \( \alpha > 1 \). Note that we rule out the case \( \alpha \leq 1 \), where both the marginal and average innovation cost are nonincreasing in \( N \). Equation (3) implies that the innovation cost is convex in \( N \). Next, we define
\[ \gamma = \frac{\alpha}{C'(N)/(NC''(N))} \]
to be the elasticity of innovation with respect to an increase in the profit from innovation. Our 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 [15].

The firm’s problem is to choose the number of the new differentiated products, \( N \), and the price in both markets to maximize its profit. Since the \( p_i \) and \( p_s \) are determined as in Section 2, the optimal \( N^* \) is determined by marginal profit and the marginal innovation cost of innovating a new differentiated good. Thus,
\[ C'(N^*) = \pi_u(p_u) + \pi_s(p_s), \]
where the right hand side is the marginal profit generated by an additional differentiated product.

The welfare of the North is a composition of the profit of the firm in the North and the South, the consumers’ surplus in the North, and the innovation cost.
\[ W_s(N) = N(\pi_u(p_u) + \pi_s(p_s) + s_u(p_u)) - C(N). \]

Moreover, since the firm is in the North, the welfare of the South includes only the consumers’ surplus generated in the South.
\[ W_s(N) = N s_s(p_s). \]

Furthermore, the global welfare is the sum of two countries welfare.

Next, we utilize a simulation approach to investigate the welfare effects of parallel trade for the North and the global welfare effects. The welfare effects of parallel trade is examined by numerical simulation for \( x \in [0, 0.5] \) under different innovation cost functions. Here, we concentrate on the cases only when \( x \in [0, 0.5] \), because no parallel trade will happen when \( x > 0.5 \). Similar results for different cases of the innovation cost function are obtained, therefore, we use two representative cases of \( \gamma \) in Figures 1 and 2 to illustrate the results. In Figures 1 and 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 \( \gamma = 10/11 \) are illustrated in Figures 1 and 2, respectively. The former represents the cases of high elasticity of innovation and the latter represents the cases of low elasticity of innovation. We define the “\( W_s \) ratio” as the North welfare with the occurrence of parallel trade divided by that of the North without the occurrence of parallel trade. Then “\( W_s \) ratio” \( > 1 \) means that the North welfare is improved when parallel trade exists. Analogously, we can also define the global welfare ratio, denoted by “\( GW \) ratio” and the global welfare with an occurrence parallel trade improves when “\( GW \) ratio” \( > 1 \).

Figure 1 shows that the two ratios are concave in \( x \). They are 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 \( W_s \) ratio, i.e., the maximal welfare improvement for the North from parallel imports. For \( x \in [0.117, 0.5] \), \( W_s \) ratio is decreasing in \( x \). In addition, when \( x = 0.231 \), we have \( W_s \) ratio = 1, i.e. the North welfare of parallel trade is equal to that without parallel trade. It suggests that if \( x \) is big enough, parallel importation is not beneficial to the importing country. Furthermore, it also shows that \( GW \) ratio is no greater than 1, and is decreasing in \( x \) for all \( x \) when \( \gamma = 5 \). It suggests that when the elasticity of innovation is relatively high then the global welfare will decrease due to parallel import-
Next, the simulation results of the relative welfare of the North and the global welfare for the case of $\gamma = 10/11$ is shown in Figure 2. 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 may be improved because of parallel trade which, however, may not be significant. Furthermore, it also shows that the North welfare of parallel imports is always greater than that under the banning of parallel trade.

To sum up the results, Figure 3 is constructed under various elasticities of innovation. The $y$-axis is defined as the elasticity of innovation, $\gamma$, and the $x$-axis is the difference between the slope of demand function in different market, $x$. Figure 3 contains three regions: A, B, and C. The points $(x, \gamma)$ falling into the area A represent the cases where parallel trade will reduce both of $W_n$ and global welfare. In particular, when both $\gamma$ and $x$ are high, parallel trade lowers the welfare of the North and the whole world. Moreover, the points in region C of Figure 3 represents the cases that both of $W_n$ and global welfare are better off. When elasticity of innovation is low, parallel trade will increase the North’s and global welfare. Finally, region B means the North’s welfare increases but global welfare decreases if parallel trade exists. Therefore, we conclude our results in Proposition 1.
Figure 3. The relation between elasticity of innovation, $\gamma$ and $x$. 

Proposition 1:
1) 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 $x$.

2) 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.

3) When elasticity of innovation and $x$ are high, parallel trade decreases the innovating country welfare as well as global welfare.

4) 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 1 indicates that parallel imports can usually make the North’s welfare improve whenever the difference between the North and the South is small. On the other hand, if the difference between the two markets is significant, parallel imports will make the North worse. Thus, Proposition 1 provides a reasonable interpretation of why some countries permit parallel importing, but some do not.

Another finding is that global welfare decreases under the case of low elasticity of innovation at any given values of $x$, which is quite different from the results in section 2, where global welfare always increases under a single product model with linear demand function. It is mainly resulted from the reduced number of innovations which is accompanied with the reduced profit. For the case of low elasticity of innovation, global welfare is improved if the difference between the two markets is not significant. The profit of monopolist is reduced which is a direct consequence of parallel trade reduces. Therefore, parallel trade reduces the incentives for innovation, which results in a decreasing number of innovations. Although parallel trade in a single product case improves the welfare of the importing country and global welfare concluded in Section 2, the welfare effects of parallel trade become ambiguous in the more general model.

4. Conclusion

We introduce a model of parallel trade with innovation to explain the phenomena why there exist various policies on parallel trade in the real world. One reasonable interpretation is that allowing parallel imports may reduce producers’ incentives for innovation, and thus reduce the number of new innovation. We show that the welfare effect is not only related to the difference between the two markets, but also related to the elasticity of innovation. Parallel trade improves the welfare of importing country, which, however, also reduces the incentives to innovate new products. Overall, the welfare effects of parallel imports are ambiguous. It provides an interpretation for why some countries allow parallel import but some do not. Furthermore, when elasticity of innovation and $x$ is high, parallel trade decreases innovating country welfare as well as global welfare. When elasticity of innovation is low, allowing parallel trade improves the welfare of the innovating country as well as global welfare. However, the increase in global welfare is not significant.

REFERENCES


