Why Parallel Imports May Raise Producers' Profits

Horst Raff
Institut für Volkswirtschaftslehre
Christian-Albrechts-Universität zu Kiel
D-24098 Kiel, Germany
Email: raff@econ-theory.uni-kiel.de

Nicolas Schmitt
Département d'économie politique
Université de Genève
CH- 1211 Genève 4, Switzerland
Email: schmitt@ecopo.unige.ch

Preliminary and Incomplete: August 2004

Abstract

Contrary to the existing literature, this paper shows that a manufacturer may benefit from parallel imports. This occurs when competitive retailers must order inventories before they know the realization of demand and for products whose sale value drops at the end of the demand period. For these types of products, letting retailers trade unsold inventories reduces the risk of destructive competition among retailers, results in larger orders placed with the manufacturer, higher manufacturer profit and higher consumer surplus. The model provides simple explanations as to why the volume of parallel imports is not only becoming more and more important but also accepted by manufacturers for some products such as automobiles (for instance in the Canada-US market), appliances, consumer electronics, musical recordings or cosmetics and perfumes.

1 Introduction

The purpose of this paper is to show that circumstances exist under which it pays a manufacturer to allow distributors (hereafter called retailers) to engage in parallel imports, that is imports that are not directly controlled by the trademark, the copyright or the patent owner, generally the manufacturer itself. Moreover parallel imports in these circumstances are generally welfare improving.

These results arise when three conditions are met. First, retailers must place orders before they know the state of the demand. Second, the states of the demand are different across markets, and third the products have little value at the end of the demand period (or equivalently, it is costly to maintain them as inventories). There are several markets for which these conditions apply. Many markets (for instance, automobiles, motorcycles, clothes) require manufacturers to have new models every year or so. Other markets, such as toys, have very narrow and well-defined demand periods (e.g., pre-Christmas season). In all these cases, the value of the products decreases significantly at the end of the demand period. Moreover, many goods are not produced on a just-in-time inventory basis and exploiting scale economies often requires significant lags between production and sales. In other words, orders must be placed well before the relevant demand period. This often leads to errors about the future strength of the demand and thus to incentives to ship unsold inventories to markets where the demand turns out to be stronger.

To understand why a manufacturer may have an incentive to allow this type of parallel imports, suppose it does not and is able to enforce effectively a parallel import ban. Retailers may be stuck with significant unsold inventories in which case they have a strong incentive to lower prices. Competition among retailers might even force the price to drop to zero as inventories are essentially sunk investments. Anticipating a possible loss, retailers may be reluctant to order large inventories. Banning parallel imports may thus have detrimental effects on the manufacturer's profitability. Allowing parallel imports on the other hand constitutes a simple mechanism by which the retail price does not fall dramatically when the state of the demand turns out to be low and it provides incentives to retailers to place larger orders than they otherwise would.

Is there evidence that manufacturers like parallel imports? It is notorious that data on parallel imports are very hard to come by simply because trade

statistics do not distinguish between authorized and unauthorized intermediaries (Maskus, 2000). However, the volume of parallel imports has simply become too large to argue seriously that manufacturers either do not know of their existence or know about them but are unable to stop parallel imports even if they lower their profitability. In other words, manufacturers must benefit from parallel imports at least in some markets.[1] This view is held by many experts. For instance, Lipner (1990) writes '[..] some manufacturers, while publicly opposed to gray markets sales of their products, privately do little to inhibit their flow and in some instances even go so far as to encourage these transactions' (p4). Similarly, in a report prepared for the EU Commission (NERA, 1999), the authors write 'Some parallel trade, however, seems to be beneficial to the trademark owner. [..] If, for example, there is for some reason over-production in the source country, and the manufacturer would otherwise be left with an unsold stock, parallel trade may be a means to raise profits through additional sales. Another example is goods such as clothing which are subject to fashion waves. Previous season's clothing in one country can still yield useful revenues in other countries' (p.11). Furthermore, a 'dealer might [..] have over ordered, or might have excess quantities of an older or out-of-date version of the goods' (Lipner, 1990, p7).[2]

This can also be seen with specific examples. In the North American automobile market, retailers selling the products of a manufacturer must sign a contract forbidding them to sell in other countries cars intended for sale in one country. Yet over 200'000 new vehicles intended for the Canadian market were resold south of the border in 2001 (up from 16,000 in 1996; see Kurylko, 2002). Obviously, this no-international trade clause is not enforced even if retailers may face heavy penalties.[3] Parallel trade arises in this market in part because of the quota system adopted by most North American manufacturers. Retailers receive a pre-determined volume allocation per models in order to

[1] In the US, parallel imports were estimated to be worth \$7-10 billions in the mid-1980s (Cespedes, Corey and Rangan, 1988); today, estimates of \$20 billion can be found (Campbell, 2001). In Europe, the volume of parallel trade varies from 5% in markets like appliances, motorcars and consumer electronics to nearly 15% for musical recordings, cosmetics and perfumes (NERA, 1999; House of Commons, 1999). In the UK, the number of parallel imported motorcycles has increased from about 15,000 units in 1997 to 25,000 units in 1998 representing 25% of all new machines (House of Commons, 1999).

[2] Similar points are made in Cespedes et al. (1988) about disk drives, in Campbell (2001) about IT products, and in House of Commons (1999) about motorcycles, clothing/footwear and cosmetics.

[3] Interestingly, this clause seems to have rarely been tested in court.

force them to sell a minimum number of cars without restricting retail prices. The unavoidable consequence is a mismatch between demand and supply creating strong incentives for parallel imports between Canada and the US especially when a particular model sells well in one country and not in the other.[4] The automobile producers could easily eliminate this unintended trade.[5] The fact that they do not indicates that they are not particularly hurt by this trade. Other examples of parallel trade arise from the fact that particular models are simply not distributed in one country[6], or that manufacturers deliberately oversupply certain markets because they want to keep a presence there.[7]

The existing literature on parallel imports has mainly focused on price discrimination issues (Malueg and Schwartz, 1999; Richardson, 2002; Chen and Maskus, 2002). This literature finds that manufacturers want to avoid parallel imports. This is the case because manufacturers always prefer market segmentation and price discrimination to arbitrage through parallel imports forcing uniform pricing. The welfare effects of introducing parallel imports are generally ambiguous simply because the elimination of price discrimination may or may not be welfare improving (Tirole, 1988). Does this mean that this literature has pursued the wrong track? Not at all. Our explanation is not a substitute but a complement to price discrimination issues. In our model, the manufacturer's incentive to allow parallel imports is particularly strong for markets with similar retail prices. This incentive may quickly disappear in the presence of significant retail price differences between markets. In this case, a manufacturer has every incentive to keep markets segmented in order to practice (third degree) price discrimination. As parallel imports

[4] 'When a customer wants a car that the dealer cannot supply, the dealer has two choices [...], they can turn the customer away or they can source the car from a Canadian exporter [...]. When a new luxury car hits North American dealer lots, it can generate year-long American waiting lists; [in Canada], they just sit on the lot' (Financial Post, 2001).

[5] This huge increase in parallel imports is also due to wholesale pricing policies adopted by the automobile producers in the two countries. Canadian dealers apparently benefit from lower prices than their US counterparts because they are located in a generally thiner, weaker and more price elastic market (Kurylko, 2002).

[6] In Malysia, Toyota vehicles are imported by unauthorized dealers from Japan and sold as Lexus because Lexus does not yet carry the full range of models in this country (Muller, 2004). See Liang (2003) for similar examples about certain models of cellular phones.

[7] For instance, Mercedes-Benz wanting to keep a presence in Barbados is systematically shipping too many cars there. They invariably end up in the UK (see House of Commons, 1999). According to this report, this implies consent to gray market by the manufacturer.

undermine such an effort, the manufacturer faces a trade off, at least for the products that satisfy our three conditions, between using international price discrimination and allowing parallel imports as a retailer's incentive device. This may explain why the North American automobile manufacturers do not always enforce the no-international trade clause.[8]

Finally, our approach underlines another important point. It places squarely the incentive to have parallel imports on retailers, not on consumers.[9] This point is often lost in the current literature typically based on consumer arbitrage. In this sense, it brings the literature on parallel imports closer to the legal (including law and economics) literature where parallel imports (or gray markets) are mainly viewed as a contractual issue between manufacturers and intermediaries (see Lipner, 1990; Gallini and Hollis, 1999).

The rest of the paper is organized as follows. In Section 2 we illustrate the main points of the paper with the help of a simple example. Section 3 contains a general model and proofs of the main results. Section 4 concludes.

2 An Example

Consider a risk-neutral monopoly manufacturer selling to a continuum of risk-neutral retailers in two different countries, denoted A and B. In turn these competitive retailers sell to consumers. Retailers must deal with two key characteristics of the market: first, they must order and take possession of inventories before demand becomes known, and second, inventories left unsold at the end of the demand period have no value. Retailers face a constant unit cost of distribution that we normalize to zero. The manufacturer's production cost is also assumed to be zero, as is the trade cost; this implies that we do not need to specify in which country the manufacturer is located.

We denote the volume of realized sales in country i by q_i. The volume

[8] Although retail prices for automobiles are lower in Canada than in the US (especially in the luxury car segment), threating retailers may be enough to limit parallel imports while tolerating them. It is only during periods of very significant retail price differences between the US and Canada (whether due to a low Canadian dollar or to a demand boom in the US as during the 1999-2001 period) that manufacturers start enforcing the no-international trade clause.

[9] It is not that consumer international arbitrage may not also help. However, as long as there is no perfect consumer arbitrage, manufacturers and retailers will always prefer parallel imports carried out by retailers than by consumers.

of realized sales may differ from the volume ordered from the manufacturer, denoted by x_i, if after the realization of demand some goods are reexported or inventory is left unsold. Demand in country A is deterministic and given by the simple inverse demand function $p_A = 1 - q_A$. The inverse demand function in B is $p_B = 1 - \frac{q_B}{\hat{\theta}}$, where the random variable $\hat{\theta}$ takes on a value of 1 if demand is high and $\theta < 1$ if demand is low. The low state of demand occurs with probability λ. We have picked these demand functions to illustrate two points. First, parallel imports in our model occur even if the manufacturer does not price-discriminate between markets. In fact, a monopolist selling directly to consumers would set the same price in the two markets. Second, parallel imports may go in both directions, even if one country has a larger market than the other and/or one country has deterministic demand.

The game we study has the following order of moves. The manufacturer first announces a wholesale price w_i for each market $i = A, B$ and whether retailers have the authorization to engage in parallel imports. The retailers then decide how much to sell at the market-clearing price.

2.1 No Parallel Imports

We first derive the benchmark solution when the manufacturer does not authorize the competitive retailers to ship any inventory to the other country (no parallel imports). In this case, the equilibrium is the same as the solution to the flexible price game analyzed by Deneckere et al. (1997). It is the case simply because there is no trade that is contingent on the relative state of the demand in the two countries. In other words, the two markets are completely separate. It is straightforward to show that $\hat{p}_A = \hat{w}_A = 1/2$, $\hat{q}_A = \hat{x}_A = 1/2$ and that the manufacturer's profit in market A is $\pi_A^m = 1/4$.

Consider now market B. With competitive retailers, we can associate sales with the volume ordered from the manufacturer because retailers have every incentive to sell whatever volume they have ordered at whatever price supported by the market as the value of the product drops to zero at the end of the demand period. In particular, the retail price may effectively drop to zero when the demand turns out to be low. Hence there are two possibilities when the demand is low: the retail price may be positive or equal to zero. Of course, this will depend on the level of the wholesale price: if it is high enough, inventories will tend to be low leading to a positive retail price. To uncover the circumstance under which the retail price may fall to zero when the demand is low, consider each case separately.

If the retail price in B falls to zero in the low demand state (*case 1*), a retailer earns positive revenue only when the demand is high. Hence, his expected profit is

$$E\pi_B^r = (1-\lambda)(1-q_B)q_B - w_B q_B.$$

Since in equilibrium the retailer's profit is necessarily equal to zero, $q_B = 1 - \frac{w_B}{1-\lambda}$. The manufacturer then chooses w_B that maximizes $\pi_B^m = w_B q_B$. The solution is $\widehat{w}_B = \frac{1-\lambda}{2}$, implying sales and order volumes of $\widehat{q}_B = \widehat{x}_B = \frac{1}{2}$, a retail price in the state of high demand of $\widehat{p}_B^h = \frac{1}{2}$, and a low-demand retail price of $\widehat{p}_B^l = 0$. We observe that the manufacturer sets a low wholesale price in order to induce retailers to order enough inventory to satisfy demand when it is high.

If the retail price in B is positive irrespective of the state of the demand (*case 2*), a retailer expects to earn

$$E\pi_B^r = \lambda\left(1 - \frac{q_B}{\theta}\right)q_B + (1-\lambda)(1-q_B)q_B - w_B q_B$$

so that $q_B = \frac{\theta(1-w_B)}{\theta(1-\lambda)+\lambda}$ ensures that the retailer makes no profit. The manufacturer then chooses $\widehat{w}_B = \frac{1}{2}$ to maximize his expected profit. This implies that $\widehat{q}_B = \widehat{x}_B = \frac{\theta}{2(\theta(1-\lambda)+\lambda)} < \widehat{x}_A$, $\widehat{p}_B^h = \frac{2(\theta(1-\lambda)+\lambda)-\theta}{2(\theta(1-\lambda)+\lambda)}$ and $\widehat{p}_B^l = \frac{2(\theta(1-\lambda)+\lambda)-1}{2(\theta(1-\lambda)+\lambda)}$. Hence the manufacturer sets the same "high" wholesale price he would if demand were deterministic, but accepts that the inventory level will be less than in country A and therefore too low when demand turns out to be high.

Since the profit earned by the manufacturer in market B is $\widehat{\pi}_B^m = \frac{1-\lambda}{4}$ when $\widehat{w}_B = \frac{1-\lambda}{2}$ (*case 1*) and $\widehat{\pi}_B^m = \frac{\theta}{4(\theta(1-\lambda)+\lambda)}$ when $\widehat{w}_B = \frac{1}{2}$ (*case 2*), it is easy to determine the manufacturer's optimal strategy in B.

Result 1: *When the manufacturer does not authorize parallel imports, he chooses $\widehat{w}_B = \frac{1}{2}$ when he expects the demands in the two states to be relatively similar ($\frac{1-\lambda}{2-\lambda} < \theta < 1$) and he chooses $\widehat{w}_B = \frac{1-\lambda}{2}$ otherwise ($\theta < \frac{1-\lambda}{2-\lambda}$). As a result, the manufacturer's total profit is*

$$\widehat{\pi}^m = \begin{cases} \frac{1}{4}\left[1 + \frac{\theta}{\theta(1-\lambda)+\lambda}\right] & \text{if } \frac{1-\lambda}{2-\lambda} < \theta < 1 \\ \frac{1}{4}\left[2-\lambda\right] & \text{if } \theta < \frac{1-\lambda}{2-\lambda} \end{cases}$$

In other words, the manufacturer chooses a high wholesale price ($\widehat{w}_B = \frac{1}{2}$) when he expects the demands to be relatively similar and a low wholesale price ($\widehat{w}_B = \frac{1-\lambda}{2}$) otherwise. We now consider the case where the manufacturer allows the retailers to engage in parallel imports to see how this simple device compares with the case without parallel imports.

2.2 Parallel Imports

The main effect of allowing parallel imports is that it provides retailers in B an opportunity to sell unsold inventories in A when demand is low. Of course, the possibility of having to compete with parallel imports from B means that retailers in A also face random (residual) demand and may end up reexporting part of their inventory to B when demand there turns out to be high. Hence even with demand shocks in only one country, parallel trade may take place in either direction.

Below, we denote the volume of parallel imports originating in B by m^j, where $j = h, l$ denotes the state of demand in B. Since x_i represents the volume of orders to the manufacturer, the volume of sales q_i in the presence of parallel imports is equal to $x_i + m^j$ $(i = A, B;\ j = h, l)$ in the country receiving parallel imports and $x_i - m^j$ in the country of origin of the parallel imports. It should be clear that when the demand in B is low $(j = l)$ then parallel imports necessarily go from B to A $(m^l > 0)$. However, when the demand in B is high $(j = h)$, parallel imports may originate in B or in A. In the first case, $m^h > 0$, while $m^h < 0$ in the second case.

Given these assumptions, the expected retail profit in B when $m^h > 0$ is equal to

$$\begin{aligned} E\pi_B^r = (1-\lambda)\{(1-\left(x_B - m^h\right))\left(x_B - m^h\right) + (1-(x_A + m^h))m^h\} \\ + \lambda\{(1 - \frac{x_B - m^l}{\theta})\left(x_B - m^l\right) + (1-\left(x_A + m^l\right))m^l\} - w_B x_B, \end{aligned} \quad (1)$$

and when $m^h < 0$, the expected retail profit in B is

$$\begin{aligned} E\pi_B^r &= \lambda\{(1 - \frac{x_B - m^l}{\theta})\left(x_B - m^l\right) + (1-\left(x_A + m^l\right))m^l\} \\ &\quad +(1-\lambda)(1-\left(x_B - m^h\right))x_B - w_B x_B. \end{aligned} \quad (2)$$

We make the further assumption that the difference in demands that triggers parallel imports between the two countries is sufficiently small that the volume of parallel imports equalizes retail prices. Hence

$$1 - \frac{x_B - m^l}{\theta} = 1 - \left(x_A + m^l\right)$$

or

$$1 - \left(x_B - m^h\right) = 1 - (x_A + m^h).$$

It follows that the parallel import volumes are

$$m^l = \frac{x_B - \theta x_A}{1+\theta} \quad \text{or} \quad m^h = \frac{x_B - x_A}{2}. \tag{3}$$

Substituting (3) into (1) or (2) and setting the retailer's profit to zero (competition at the retail level), the relationship between wholesale price and order volume such that a retailer expects zero profit is

$$w_B = 1 - \frac{1+\lambda+\theta(1-\lambda)}{2\,(1+\theta)}\,(x_A + x_B)\,, \tag{4}$$

irrespective of the direction of the parallel imports when the demand is high in B.

Using the same reasoning as above, it is easy to verify that the relationship in market A between w_A and $x_A + x_B$ is identical to (4). This means that the wholesale price in each country depends on the manufacturer's overall shipment only and that $w_A = w_B$. Of course, this is due to the assumption that parallel trade equalizes retail prices across countries. As a result, we can only determine the total volume of orders that maximizes the manufacturer's overall profit, $\Sigma\pi_i^m = w_B(x_A + x_B)$, where w_B is given by (4). Denoting the optimal overall volume by $\bar{X}$, we get:

Result 2: *When the manufacturer authorizes parallel imports, he chooses* $\widehat{w}_B = \frac{1}{2}$ *irrespective of the expected demand in* B*. This results in an overall shipment of* $\bar{X} = \frac{1+\theta}{(1+\lambda+\theta(1-\lambda))}$ *and a manufacturer's overall profit equal to*

$$\bar{\pi}^m = \frac{1+\theta}{2\,(1+\lambda+\theta(1-\lambda))}. \tag{5}$$

Figure 1 illustrates the manufacturer's total equilibrium profit as a function of θ with and without parallel imports. Not surprisingly, the manufacturer's expected profit is increasing in θ. Except at $\theta = 1$, the profit with parallel imports in case 1 is always higher than without parallel imports. In case 2, the profit from parallel imports is also higher, at least if $\theta > \frac{1}{4}$ given $\lambda = \frac{1}{2}$ (more generally, if $\theta > \frac{1-\lambda}{3-\lambda}$) and thus if the demand in the low state is not too low. The lower manufacturer's profit without parallel imports illustrates well the effect of the destructive competition. If the demand is sufficiently high as compared to the low demand (*case 1*), the manufacturer facing competitive retailers is forced to give up revenues in the low demand state to make sure retailers order enough inventories to satisfy demand when

it is high. The problem is that competition among retailers drives the retail price to zero when the demand is low. Figure 1 shows that allowing parallel imports decreases (but does not eliminate) the impact of destructive competition. Its main advantage is to be a particularly simple and cheap incentive device with respect to retailers.

2.3 Welfare

We now investigate the welfare effects of allowing parallel imports, starting with the effect on consumer surplus. Suppose first that parallel imports are not allowed and let CS_i^{npi} denote the expected consumer surplus in country i in this case. In country A consumer surplus is $CS_A^{npi} = \frac{1}{8}$. The expected domestic consumer surplus in B is

$$CS_B = \lambda(1 - p_B^l)\frac{q_B^l}{2} + (1-\lambda)(1-p_B^h)\frac{q_B^h}{2} \tag{6}$$

since the probability that the demand is low is λ and the demands are linear. The specific level of the expected consumer surplus will of course depend on whether the retail price drops to zero when the state of the demand is low. If it does (case 1), then $CS_B^{npi} = \frac{1-\lambda}{8}$ since in this case, $\widehat{q}_B^h = \widehat{q}_B^l = \frac{1}{2}$ and $\widehat{p}_B^h = \frac{1}{2}$. If the retail price does not drop to zero (case 2), we have $CS_B^{npi} = \frac{\theta(\theta+\lambda(1-\theta))}{8(\theta(1-\lambda)+\lambda)^2}$, since $\widehat{p}_B^h = \frac{2(\theta(1-\lambda)+\lambda)-1}{2(\theta(1-\lambda)+\lambda)}$, $\widehat{p}_B^l = \frac{2(\theta(1-\lambda)+\lambda)-\theta}{2(\theta(1-\lambda)+\lambda)}$ and $\widehat{q}_B^l = \widehat{q}_B^h = \frac{\theta}{2(\theta(1-\lambda)+\lambda)}$.

Suppose now that parallel imports are allowed, and let CS_i^{pi} denote consumer surplus in country i under parallel imports. Above, we only found that $x_A + x_B = \bar{X} = \frac{1+\theta}{(1+\lambda+\theta(1-\lambda))}$. Hence, we do not know the volume of orders from each country and the volume of parallel imports. To compute consumer surplus, all we need, however, is individual country sales which can be found in the following way. Suppose there is a low demand in B. Since parallel imports equalize retail prices, it must be true that $1 - \frac{q_B}{\theta} = 1 - q_A$ and thus $q_A = \frac{q_B}{\theta}$. In addition, parallel imports make sure that, across the two countries, total sales must be equal to total orders (i.e., $q_A + q_B = x_A + x_B$). Combining these two conditions, we get,

$$q_B^l = (x_A + x_B)\,\frac{\theta}{1+\theta}$$

when the demand in B is low. Using a similar reasoning, we get

$$q_B^h = \frac{x_A + x_B}{2}$$

when demand in B is high. Once retail prices in B are found, we can use (6) to get

$$CS_B^{pi} = \frac{4\lambda\theta + (1-\lambda)(1+\theta)^2}{8(1+\lambda+\theta(1-\lambda))^2}.$$

The same method yields for country A:

$$CS_A^{pi} = \frac{4\lambda + (1-\lambda)(1+\theta)^2}{8(1+\lambda+\theta(1-\lambda))^2}.$$

A comparison of expected consumer surplus in the two countries with and without parallel imports yields the following result.

Result 3: *Consumer surplus in each country is higher with parallel imports, if the demands in the two states are sufficiently similar (i.e. θ close enough to 1).*

The top of Figure 2 illustrates consumer surplus with and without parallel imports for country B. It shows that the consumer surplus in B is higher with than without parallel imports provided the demand in B is not too low. This is expected. When the demand is very low, the price in B drops to zero and this can only benefit consumers. Allowing parallel imports makes this possibility less likely and this effect is detrimental to consumers. However it also induces retailers to place larger orders than without parallel imports and this effect by being pro-competitive benefits consumers. It is this effect that explains why consumer surplus is higher with parallel imports than without for most values of the demand parameter and especially when the price does not drop to zero. Hence, allowing parallel imports lines up the manufacturer's incentive with the consumers' best interests for most states of demand.

Finally we turn to social welfare, which consists of the sum of consumer surplus and the manufacturer's profit. The retailers expect zero profit. Since we have not specified where the manufacturer is located, we can compute total social welfare, $W = CS_A + CS_B + \pi$. This welfare must increase with parallel trade for θ close enough to 1, since parallel trade raises profits and consumer surplus in both countries. Alternatively, we can compute π_B as resulting from the manufacturer's activity in B and then obtain social welfare in B: $W = CS_B + \pi_B$. Figure 2 (bottom panel) plots social welfare in B with and without parallel imports.[10] For $\theta > .165$, the welfare ranking of the two scenarios is straightforward: parallel imports lead to higher welfare.

[10] This means that Figure 1 and Figure 2 cannot be compared since Figure 1 illustrates the manufacturer's total profit over both markets.

For $\theta < .165$, the opposite occurs. The main conclusion is that there is no real disagreement between consumers and the manufacturer. Except for a small range of the demand parameter, the manufacturer and the consumers both either prefer parallel imports or prefer no parallel imports. When the manufacturer likes parallel imports it is because its profit is higher and the consumers like them because the volume of the product available on the market is higher (or equivalently the expected retail price is lower). This simple result is in sharp contrast with the conclusion of the literature on parallel imports using price discrimination since it typically produces a disagreement between consumers and producers.

3 Generalization

In this section we demonstrate that the main results of the paper go through for a more general specification of the demand and the distribution of the demand shock. Demand in country A is given by the function $D(p)$ with $D' < 0$. Demand in B is $D_B(p, \theta) = \theta D(p)$, where p is price and $\theta \in [\theta^-, \theta^+]$ is the (non-negative) realization of a random variable with density $f(\theta)$ and cumulative distribution $F(\theta)$. Like in the example, this specification implies that ex ante the monopolist would not want to price discriminate. We also still assume that in the presence of parallel trade prices are strictly positive for all realizations of demand. Note, however, that we no longer restrict the market in B to be smaller than the one in A.

. In particular, we will prove the following result:

Proposition 1 *(a) If retail prices without parallel trade are positive for all realizations of demand, then parallel trade raises the manufacturer's expected profit, expected consumer surplus and expected social welfare. (b) If the retail price falls to zero without parallel trade for some realizations of demand, then parallel trade raises the manufacturer's expected profit, expected consumer surplus and expected welfare, provided that there is sufficiently little probability mass on these realizations.*

Proof: Writing the inverse demand functions as $p_A = p_A(q_A)$ and $p_B = p_B(\frac{q_B}{\theta})$, and assuming that in the presence of parallel trade prices are strictly positive for all realizations of demand, the expected retail revenue in B, ER_B^r,

reads

$$\begin{array}{ll} \int_{\theta^-}^{\theta^+} \left\{ p_B(\frac{x_B - m_B}{s})(x_B - m_B) + p_A(x_A + m_B)m_B \right\} f(s)ds, & \text{if } m_B \geq 0 \\ \int_{\theta^-}^{\theta^+} p_B(\frac{x_B - m_B}{s}) x_B f(s)ds, & \text{if } m_B < 0. \end{array} \tag{7}$$

In country A, the expected retail revenue ER_A^r is

$$\begin{array}{ll} \int_{\theta^-}^{\theta^+} p_A(x_A + m_B)x_A f(s)ds, & \text{if } m_B \geq 0 \\ \int_{\theta^-}^{\theta^+} \left\{ p_A(x_A + m_B)\,(x_A + m_B) + p_B(\frac{x_B - m_B}{s})\,(-m_B) \right\} f(s)ds, & \text{if } m_B < 0. \end{array} \tag{8}$$

Since the expected retail profits, $E\pi_i^r = ER_i^r - w_i x_i$, must equal zero, we can write the manufacturer's problem of maximizing the expected value of $\pi^m = w_A x_A + w_B x_B$ as

$$\max_{x_A, x_B} \int_{\theta^-}^{\theta^+} \left\{ p_A(x_A + m_B)\,(x_A + m_B) + p_B(\frac{x_B - m_B}{s})\,(x_B - m_B) \right\} f(s)ds. \tag{9}$$

The volume of parallel trade, $m_B = m_B(x_A, x_B, \theta)$, is the flow of goods that equalizes retail prices across the two countries ex post for each level of x_A, x_B and θ. It is implicitly defined by

$$p_A(x_A + m_B) = p_B(\frac{x_B - m_B}{\theta}). \tag{10}$$

Seeking a contradiction, suppose that the manufacturer's profit is maximized for $m_B = 0$, and let the corresponding profit-maximizing order levels be denoted by x_A^* and x_B^*. Assume for the time being that prices are non-negative at these volumes for all realizations of θ. Using the envelope theorem we then must have

$$\frac{d\pi^m(m_B)}{dm_B} = \int_{\theta^-}^{\theta^+} \left\{ p_A(x_A^*) + x_A^* p_A'(x_A^*) - p_B(\frac{x_B^*}{s}) - \frac{x_B^*}{s} p_B'(\frac{x_B^*}{s}) \right\} f(s)ds = 0. \tag{11}$$

That is, x_A^* and x_B^* are chosen to equalize expected marginal revenues in A and B. In addition, marginal revenue must equal marginal cost, which is zero. Hence

$$p_A(x_A^*) + x_A^* p_A'(x_A^*) = \int_{\theta^-}^{\theta^+} \left\{ p_B(\frac{x_B^*}{s}) + \frac{x_B^*}{s} p_B'(\frac{x_B^*}{s}) \right\} f(s)ds = 0. \tag{12}$$

However, the probability that marginal revenues are equalized ex post for x_A^* and x_B^* is zero. That is, for almost all $\theta \in [\theta^-, \theta^+]$ we must have:

$$(p_A(x_A^*) + x_A^* p_A'(x_A^*)) = 0 \gtrless p_B(\frac{x_B^*}{\theta}) + \frac{x_B^*}{\theta} p_B'(\frac{x_B^*}{\theta}), \tag{13}$$

and hence market-clearing prices will also differ across the countries. Permitting retailers to undertake parallel trade will then lead to a flow of goods from the country with the lower price (and marginal revenue) to the country with the higher price (and marginal revenue). thereby raising the overall retail profits in A and B ex post for almost all realizations of θ. Integrating over $\theta \in [\theta^-, \theta^+]$ to obtain expected retail profit, we observe that allowing parallel trade raises this profit holding fixed the volumes ordered from the manufacturer at x_A^* and x_B^*. Since retailers are perfectly competitive and face given wholesale prices, the quantity ordered from the manufacturer has to increase to keep expected retail profits at zero. This implies for the manufacturer that $d\pi^m(m_B)/dm_B > 0$, which contradicts the assumption that $m_B = 0$ was optimal.

Next, suppose that in the absence of parallel trade the choice of x_A^* and x_B^* causes the retail price in B to drop to zero for low realizations of demand. In particular, let $\tilde{\theta} \equiv \tilde{\theta}(x_B^*)$ be defined by $p_B(x_B^*/\tilde{\theta}) = 0$, so that price is zero for $\theta \in [\theta^-, \tilde{\theta}]$. We then have to modify (12) as follows:

$$p_A(x_A^*) + x_A^* p_A'(x_A^*) = \int_{\tilde{\theta}}^{\theta^+} \left\{ p_B(\frac{x_B^*}{s}) + \frac{x_B^*}{s} p_B'(\frac{x_B^*}{s}) \right\} f(s) ds = 0. \tag{14}$$

Consider what happens if we allow parallel trade. For any $\theta \in [\theta^-, \tilde{\theta}]$, goods will flow from country B to country A. This will have no effect at the margin on retail profit in B, because the price there is zero. The retail profit in A, however, must fall as successive units are shipped there, since x_A^* was chosen optimally to start with, and additional units cause marginal revenue to become negative. Thus parallel trade may lower total ex-post retail profit for low values of θ. For $\theta \in (\tilde{\theta}, \theta^+]$, parallel trade will raise overall ex-post retail profit for the reasons given in the preceding paragraph. Integrating over $\theta \in [\theta^-, \theta^+]$, we observe that the effect of parallel trade will depend on the distribution of θ. In particular, if there is sufficiently little probability mass on $[\theta^-, \tilde{\theta}]$, the manufacturer's expected profit will increase with parallel trade.

Expected consumer surplus rises with parallel trade, if the volume ordered from the manufacturer does. This is the case, if the retail price in B does not

drop to zero in the absence of parallel trade. Since the manufacturer's profit rises in this case, parallel trade unambiguously increases expected welfare. If without parallel trade the retail price in B drops to zero for low realizations of demand, then expected consumer surplus and welfare will rise, provided there is not too much probability mass on these realizations.■

4 Conclusions

This paper demonstrated that there are circumstances under which it pays a manufacturer to allow retailers to engage in parallel trade. Specifically, parallel trade helps avoid destructive competition between retailers and gives the latter an incentive to place larger orders than they otherwise would. Moreover, it was shown that parallel imports in these circumstances are generally welfare improving. These results arise when three conditions are met: first, retailers must place orders before they know the state of the demand; second, the states of the demand are different across markets; and third the products have little value at the end of the demand period (or equivalently, it is costly to maintain them as inventories).

There are other possible mechanisms that would allow a manufacturer to reduce destructive competition, including resale price maintenance (see Deneckere et al., 1997), a return policy for unsold inventories, or vertical integration with retailers. The point of the paper is that parallel trade constitutes a particularly simple mechanism to achieve this goal. We would expect this to be true, especially if the manufacturer has less information than retailers about local market conditions, as seems especially likely if the manufacturer is located overseas. Resale price maintenance, for instance, would involve considerably more checking and monitoring of retailers than allowing parallel imports. Similarly a manufacturer's return policy is costly, not only because it might allow well informed local retailers to shirk on sales effort, but also because a foreign manufacturer may have no particular physical facilities in its export markets to handle returned merchandise. Finally, vertical integration may be difficult for a foreign manufacturer again due to informational asymmetries, but also because the market volume may be too small to justify the investment.

References

[1] Campbell, Scott (2001), 'Vendors Aim to Squash Gray Market', *Computer Reseller News*, Issue 963, September 17.

[2] Cespedes, F.; E.R. Corey abd V.K. Rangan (1988), 'Gray Markets: Causes and Cures', *Harvard Business Review,* July-August, 75-82.

[3] Chen, Yongmin and Keith Maskus (2002), 'Vertical Pricing and Parallel Imports', mimeo

[4] Deneckere, Raymond, Howard P. Marvel and James Peck (1997), 'Demand Uncertainty and Price Maintenance: Markdowns as Destructive Competition' *American Economic Review 87, 4, 619-41.*

[5] Financial Post (2001), 'Wanted: Canadian Cars', November 3, 2001

[6] Gallini, Nancy and Aiden Hollis (1999), 'A Contractual Approach to the Gray Market', *International Review of Law and Economics,* 19(1), 1-21.

[7] Ganslandt, Mattias and Keith Maskus (2001), 'Parallel Imports of Pharmaceutical Products in the European Union', Working paper 546, Research Institute of Industrial Economics, Stockholm.

[8] House of Commons. 1999. *Trademarks, Fakes and Consumers.* Eight Report, Trade and Industry Committee. http://www.parliament.the-stationery-office.co.uk/pa/cm199899/cmselect/cmtrdind/380/38007.htm

[9] Kurylko, Diane. (2002). 'Gray Market Yields Gold.' *Automotive News* 76, (5972), February 25.

[10] Liang, Ann (2003), '30 Percent of Malaysian Cells Purchased Through Grey Market', World IT Report, December 28

[11] Lipner, Seth. 1990. *The Legal and Economic Aspect of Gray Market Goods.* London: Quorum Books.

[12] Maskus, Keith. 2000. Parallel Imports. *World Economy* 23, 9. 1269-84.

[13] Malueg, D.A. and Marius Schwartz (1994), 'Parallel Imports, Demand Dispersion and International Price Discrimination', *Journal of International Economics 37, 167-96.*

[14] Muller, Urs (2004), 'No Matter What Color that Lexus is Asia Quite Often Is Gray', *Automotive News* 78, 6078, February 2.

[15] NERA (1999), 'The Economic Consequences of the Choice of a Regime of Exhaustion in the Area of Trademarks', Final report for DGXV of the European Commission, February.

[16] Richardson, Martin (2002), 'An Elementary Proposition Concerning Parallel Imports', *Journal of International Economics* 56, 1, 233-45.

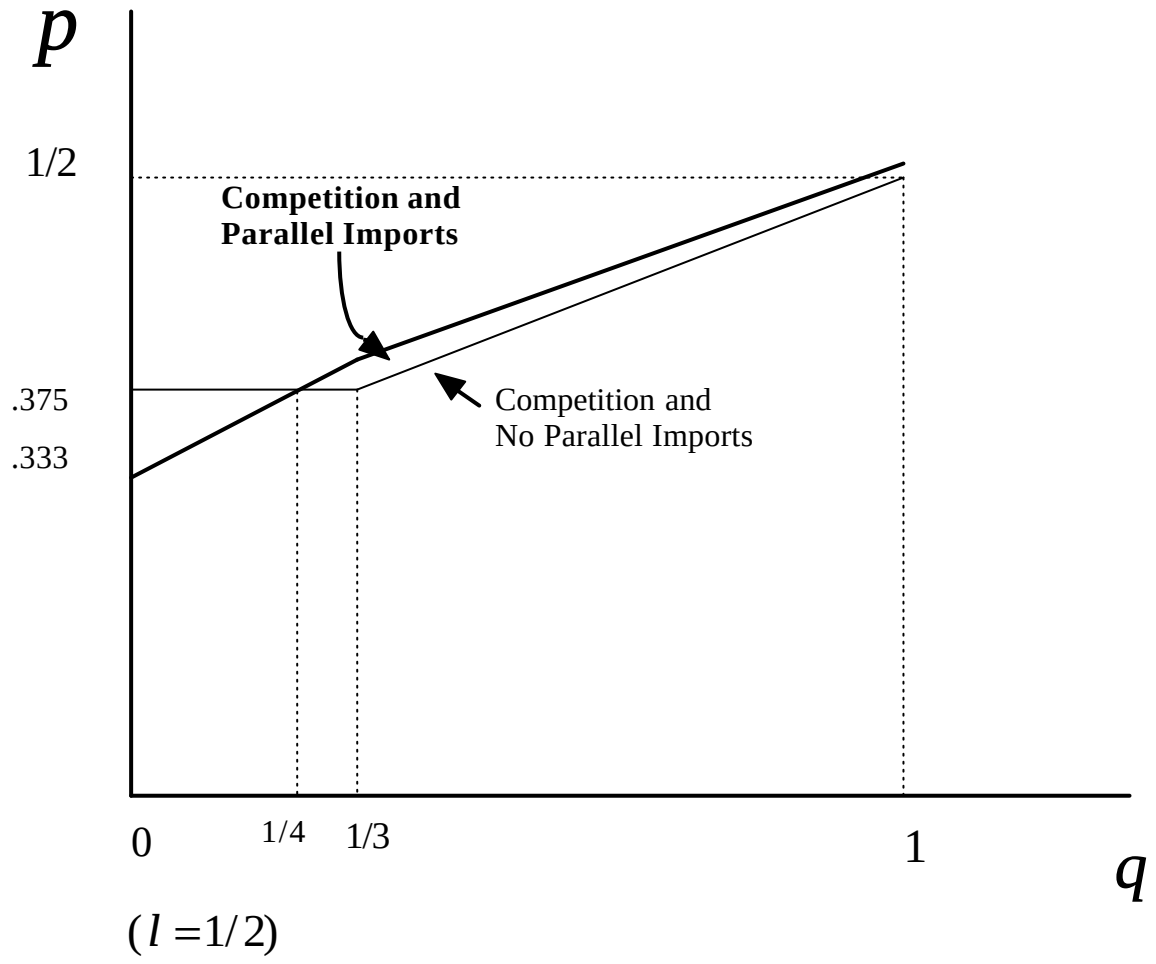

Figure 1: Manufacturer's Overall Profit

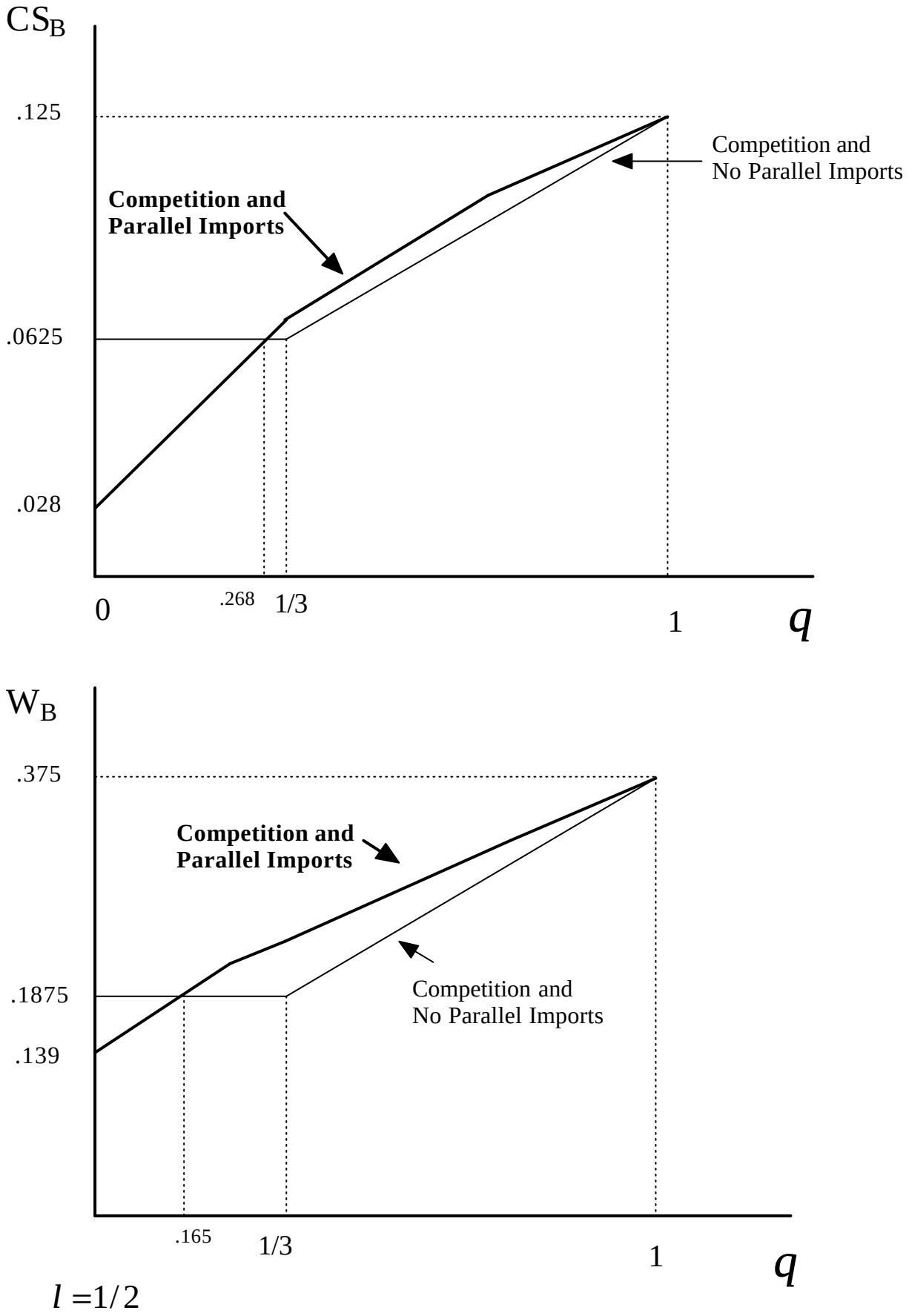

Figure 2: Consumer Surplus and Welfare in B