Commodity Taxation and Parallel Imports

Pascalis Raimondos-Møller[1]

Copenhagen Business School, CEPR and CESifo

Nicolas Schmitt[2]

Simon Fraser University and CESifo.

October 2007

[1]Corresponding author: Department of Economics, Copenhagen Business School, Porcelaenshaven 16A, 2000 Frederiksberg, Denmark, Ph: +45 3815 2594, Email: prm.eco@cbs.dk

[2]Department of Economics, 8888 University Drive, Burnaby BC V5A 1S6, Canada, Ph +1 778 782 4582, Fax: +1 778 782 5944, Email: schmitt@sfu.ca

Abstract

We examine the interaction between commodity taxes and parallel imports in a simple two-country model with imperfect competition. While governments determine non-cooperatively their commodity tax rate, the volume of parallel imports is determined endogenously by the retailing sector. We compare the positive and normative implications of having commodity taxes based on destination or origin principle. Origin taxes are shown to have very attractive properties: they lead to lower levels of optimal taxes, they converge as parallel imports increase (while destination taxes diverge), and they lead to higher welfare levels.

    JEL Classification: F12, F15, H21, H24

    Keywords: Parallel import, commodity taxation, market integration

1 Introduction

This paper examines the interaction between commodity taxes and parallel imports in a two-country model of imperfect competition. The main issue investigated is whether increased volumes of parallel imports — for many, a synonym of market integration — lead to a tax convergence or not. The paper shows that more parallel imports do induce a tax convergence if the taxes are based on the origin principle and to a tax divergence if taxes are based on the destination principle. Interestingly, origin taxes are shown to be unambiguously welfare improving.

Parallel imports are legal and highly encouraged within the European Union (EU).[1] To achieve this, the EU has adopted a *regional* exhaustion rule of intellectual property rights. This rule implies that if a EU producer chooses to export within the EU by using local retailers, these retailers can legally re-export the good. The European Court has repeatedly ruled in favour of parallel imports, arguing that they intensify competition and lead to deeper market integration.[2] Recent empirical analyses of the extent of parallel imports within the EU document the importance of these flows.[3]

Existence of parallel imports implies the existence of price differentials that exceed the cost of transporting these products across borders. These price differentials may exist for several reasons. A frequent reason is different consumer's willingness to pay across countries (whether due to tastes or income). This typically induces imperfectly competitive firms to charge different prices across countries (see Malueg and Schwartz, 1994). Prices may also differ, not because the markets are different but because the realization of demands is different across markets. Raff and Schmitt (2007) develop a model where producers and consumers may benefit from parallel imports when retailers must place orders before they learn about the intensity of the demand in their respective market. A third reason for price differentials is government policies and regulations, such as price ceilings on pharmaceutical products (see Kanavos and Costa-Font, 2005; Grossman and Lai, 2006).

---

[1] Parallel imports, also known as grey products, are genuine products imported in a country without the authorization of the intellectual property right owner. See Maskus (2000) for a general discussion.

[2] For example, see the European Commission's decision on the case of Glaxo Wellcome et al. (Official Journal, 2001), and on JCB (Official Journal, 2002) where a 39 million EUR penalty was imposed for attempting to stop parallel imports.

[3] For instance, NERA (1999) reports that for CDs, consumer electronics, auto spare parts, cosmetics, and soft drinks, 5-20% of trade within EU are parallel imports; Ganslandt and Maskus (2004) report that, for some pharmaceutical product brand names, the share of parallel imports reach 50% in Sweden. Additional estimates can be found in OECD (2002), Ahmadi and Yang (2000), and Kanavos and Costa-Font (2005).

The present paper fits within this last category of papers, where government policies affect parallel imports, but turns its attention to *commodity taxes.*

Commodity taxes are important revenue-generating instruments in the EU countries. Different tax rates may naturally lead to different consumer prices that parallel importers seek to profit upon. In turn, parallel imports may affect the non-cooperative choice of these taxes. This simultaneous endogeneity of parallel imports and commodity taxes makes the analysis interesting and by no means straightforward.

Our reference point is the commodity-tax competition literature (see Lockwood, 2001 for a survey). In this literature, it is usually assumed that markets are fully integrated so that the law of one price applies and parallel imports never arise. This literature then focuses on two issues: (i) whether the imposition of tax harmonization, i.e. forcing countries to adopt the same tax, leads to efficiency gains or not, and (ii) which tax regime is best — the destination or the origin tax regime.[4]

Our starting point, however, is somewhat different. We assume that markets are segmented and that all agents behave optimally. The equilibrium is characterized by price differentials that parallel importers exploit given some transaction costs on parallel imports. The policy initiative is to reduce these transactions costs and thus to promote market integration.[5] Parallel imports then increase, affecting retail prices and the taxes set by the governments. In turn, these new prices and taxes affect parallel imports, and so on until a new equilibrium is reached.

Within this set-up we ask a novel question: *does promoting parallel imports bring tax convergence?* Our motive should be clear: while repeated calls for tax harmonization in the EU have had little success, we investigate whether a more 'market-oriented' initiative can bring about a tax convergence through the 'back-door'. Indeed, implementing market integration through parallel imports is arguably more 'market-oriented', and maybe simpler, than asking member countries to harmonize their taxes. Needless to say, such a tax convergence is only desirable if it leads to Pareto efficiency gains, an issue that we also investigate in detail.

To analyze the interaction between commodity taxes and parallel imports, we adopt the Maskus and Chen (2002, 2004) model of parallel imports. This model has the advantage of providing an explicit level of parallel imports and of explaining parallel imports as a by-product of vertical control issues associated with retailing activities. In particular the model assumes a sin-

---

[4]While most papers in this literature focus on one issue, few papers analyse simultaneously the two of them, see Keen et al. (2002) and Behrens et al. (2007)..

[5]One could interpret the consistent rulings of the European Court as steps needed to reduce transactions costs faced by parallel importers.

gle good monopolist that sells in two countries. In the home country, the manufacturer sells directly to consumers and, in the foreign country, it sells through a retailer. This vertical separation in the foreign country creates a double-marginalization problem. To solve it, the manufacturer adopts a two-part tariff, i.e. a low wholesale price and a high fixed fee. However, faced with a low wholesale price, the foreign retailer may find profitable to engage in parallel imports and sell part of its order to the manufacturer's home country. Thus, in trying to avoid the double-marginalization problem, the manufacturer may create a competition problem in its home market (from monopoly to duopoly) and lower its overall profit. In equilibrium, the two-part tariff strategy is set so as to balance the two opposing motives faced by the manufacturer.

We augment this model with taxes that governments choose in a non-cooperative way and we examine the implications of two different tax systems: *destination*-based taxes — where taxes are set and collected by the authorities of the country where the good is purchased and consumed (the current EU system) — and the case of *origin*-based taxes — where taxes are set and collected by the authorities of the country where the good is produced (a proposed EU system). Investigating the effects of a reduction of the transaction cost faced by parallel importers on the equilibrium, and in particular on the level of commodity taxes, we show three main results. First, the equilibrium tax rates in the two countries are more similar under origin taxation than under destination taxation. Second, origin-based tax rates converge as parallel trade increases, whereas destination-based taxes diverge. Third, origin-based taxes lead to higher welfare in both countries than destination-based taxes. Thus, market integration and parallel imports can indeed bring both tax convergence and welfare improvements when taxes are origin-based.

Three recent papers have looked at commodity taxation issues in segmented markets. Haufler et al. (2005) use a symmetric reciprocal-dumping trade model to examine how trade cost reductions affect the choice of commodity tax base. They show that, while origin taxes still outperform destination taxes when trade costs are low,[6] the opposite is true when they are high. Haufler and Pflüger (2006) add government revenue requirements to the same reciprocal-dumping model and re-examine the comparison between destination and origin taxes. They show that ambiguity emerges in many of the earlier results. Finally, Behrens et al. (2007) examine how tax (origin-and destination-based) harmonization affects firm location. They use a gen-

---

[6]The superiority of origin-based taxes has been established for the first time by Keen and Lahiri (1993, 1998).

eral equilibrium model with monopolistic competition that allows for variable mark-ups and market segmentation. A main result of their analysis is that countries do not agree on a common tax regime — as firms agglomerate into the large country, this country prefers origin-based taxes while the smaller country prefers destination-based taxes. However, while their analysis introduces the interesting aspect of asymmetry in the location of production, it does not take into account the fact that commodity taxes also adjust to that asymmetry.

Our paper keeps firm location fixed as in Haufler et al. (2005) but allows for asymmetries in the location of production as in Behrens et al. (2007). Like Haufler et al. (2005), we look at how market integration affects the non-cooperative choice of origin and destination-based taxes, but we use a different trade model than theirs, one where vertical links and parallel imports play a key role. We perform our analysis around of what we think is a novel research question ('does market integration lead to tax convergence?') and we show some interesting new results (converging origin taxes are *always* better than the diverging destination taxes). Thus, our simple model shows no ambiguity about which tax base is best (an ambiguity that exists in Haufler et al., 2005) and no conflict between countries on the choice of a common tax regime (a conflict that exists in Behrens et al., 2007).

The rest of the paper is organized as follows. Section 2 develops the model and examines the case where taxes are set according to the destination principle. It derives the optimal taxes and whether they converge or diverge when parallel imports increase. Section 3 analyses the case of origin taxes. Section 4 compares the two tax regimes and investigates the welfare consequences of having more parallel imports. Some concluding remarks are found in section 5.

2 The model with destination-based taxes

As already mentioned, we adopt the Maskus and Chen (2002, 2004) model of parallel imports augmented with endogenously chosen commodity taxes.[7] A manufacturer sells a single product in two countries: its home country, called A, and a foreign country, called B. At home, the manufacturer sells q_A directly to the consumers, while abroad it sells to an independent retailer. This retailer sells q_B units to the consumers in country B and, if it chooses

---

[7]Maskus and Chen (2005) generalize their model in several directions (general demand and cost functions, alternative vertical restraints, price competition etc.) and show that the basic mechanisms and results are preserved. Additional extentions can be found in Li and Maskus (2006) and Ganslandt and Maskus (2007).

to do so, m units as parallel imports to the consumers in country A.

The (inverse) demand function in country A is $p_A = 1 - (q_A + m)$,[8] while in country B it is $p_B = \alpha - q_B$, where p_i $(i = A, B)$ is the consumer price in country i. We assume that the willingness to pay is equal or lower in country B than in country A $(0 < \alpha \leq 1)$. For simplicity, production and retail costs (other than the cost of buying from the manufacturer) are constant and normalized to zero. In addition there is no trade cost associated with authorized trade flows. However, there is an international transaction cost g per unit of parallel imports that the retailer must pay.[9] It is assumed that parallel imports lead to Cournot competition in market A.[10]

In order to solve the double-marginalization problem, the manufacturer charges a wholesale price w and a fixed fee F to the retailer of country B so as to extract the retailer's profit. Let t_A^D (t_B^D) be the specific commodity tax in country A (country B) based on the destination principle (tax collected in the country of consumption). The manufacturer and the retailer profits are then respectively

$$\Pi^M = (p_A - t_A^D)q_A + w(q_B + m) + F; \tag{1}$$
$$\Pi^R = (p_B - t_B^D)q_B + (p_A - t_A^D - g)m - w(q_B + m) - F. \tag{2}$$

The first term of the RHS of (1) is the manufacturer's revenue at home while the second and third terms are the revenues from selling to the retailer in country B. Similarly, the RHS of (2) captures respectively the retailer's revenue in its own market and from parallel trade, while the last two terms represents respectively the variable and the fixed retailer cost.

Since the government in each country imposes a (per unit) tax t_i^D $(i = A, B)$ based on the destination principle, welfare is defined as

$$W_A^D = CS_A + \Pi^M + t_A^D(q_A + m); \tag{3}$$
$$W_B^D = CS_B + t_B^D q_B, \tag{4}$$

where CS_i represents the consumer surplus in country i, Π^M is the manufacturer's profit and the last term in both relationships represents the consumption-tax revenues collected in country i. Since the retailer in B

---

[8]We model parallel imports as perfect substitutes to the product sold directly by the manufacturer. Allowing for some product differentiation would not affect the main results of the paper.

[9]Thus relative to authorized trade, unauthorized trade flows are costly; these costs are assumed to be resource costs such as re-packaging and re-labeling of the goods – costs that parallel importers typically incur in the case of pharmaceutical products.

[10]Since products are assumed to be homogeneous, Bertrand competition in market A would never lead to parallel imports in equilibrium.

earns zero profit in equilibrium, welfare in B includes only the consumer surplus and the tax revenue.

We investigate the following three-stage game. In the first stage, governments choose simultaneously their tax rates t_A^D, t_B^D. In the second stage, the manufacturer sets w and F. In the third and final stage, the retailer sets q_B and m while the manufacturer sets q_A.

Starting with the last stage of the game, we maximize (1) with respect to q_A, (2) with respect to q_B and m, and solve for the Cournot outputs q_A and m. The results are:

$$q_A(w, t_A^D) = \frac{1 - t_A^D + g + w}{3}; \qquad m(w, t_A^D) = \frac{1 - t_A^D - 2g - 2w}{3};$$

(5)

$$q_B(w, t_B^D) = \frac{\alpha - t_B^D - w}{2}.$$

To determine the wholesale price w, we move to the second stage of the game where the manufacturer chooses w and F. In the absence of parallel imports,[11] the manufacturer's optimal pricing strategy is to set the wholesale price equal to the marginal cost of production, i.e. $w = 0$, and a fixed fee equal to the retailer's profit. However, when parallel imports are allowed, a low wholesale price leads to parallel imports and thus to competition in country A. To mitigate this effect, the manufacturer sets its wholesale price higher than marginal cost. The optimal choice of w balances the double marginalization and the increased competition problems. Since the manufacturer extracts the entire profit of the retailer, it sets $F = q_B^2 + m^2$. Substituting F and (5) into (1) and maximizing with respect to w gives

$$w(t_A^D) = \frac{2 - 2t_A^D + 8g}{13}.$$

(6)

Substituting (6) into (5) gives us the equilibrium values of outputs for given taxes:

$$q_A(t_A^D) = \frac{5(1 - t_A^D) + 7g}{13}; \qquad m(t_A^D) = \frac{3(1 - t_A^D) - 14g}{13};$$

(7)

$$q_B(t_A^D, t_B^D) = \frac{13(\alpha - t_B^D) - 2(1 - t_A^D) - 8g}{26}.$$

Hence, a higher tax t_A^D discourages total sales in A by raising the consumer price. As a consequence, parallel imports decrease. The manufacturer is now

---

[11] That is, when exclusive territory constraints at the country level can be enforced.

able to focus more on the double marginalization problem in country B by reducing w. This increases sales in country B. A higher tax in country B naturally affects negatively sales in that country, but it does not affect parallel imports as they do not depend on t_B^D (an outcome due to the constant retail cost assumption).

We now move to the first stage of the game where governments choose taxes by maximizing domestic welfare. Substituting (6)-(7) into (3)-(4) and solving for the optimal tax choices gives the following best-reply functions:

$$t_A^D = -\frac{1}{66}(38 + 9g) \tag{8}$$

$$t_B^D = \frac{1}{39}\left[13\alpha - 2 + 2t_A^D - 8g\right] \tag{9}$$

Hence, while taxes are strategic complements for country B, country A has a dominant tax rate strategy. The intuition is the following: we showed that, as t_A^D increases, the sales in country B rise. As a result, the manufacturer earns higher profits from country B (through the fixed fee charged to the retailer). The government in country B has then an incentive to raise its tax rate so as to extract additional rents from the manufacturer. However, country B's tax changes do not feed back to country A as the manufacturer's wholesale price, and thus the volume of parallel imports, do not depend on t_B^D.

We can now solve for the Nash equilibrium tax rates $(\widehat{t}_A^D, \widehat{t}_B^D)$ and obtain:

$$\widehat{t}_A^D = -\frac{1}{66}\{38 + 9g\} \qquad \text{and} \qquad \widehat{t}_B^D = \frac{\alpha}{3} - \frac{1}{99}\{8 + 21g\}. \tag{10}$$

It is worth noting that the consumption tax in country A is always negative, i.e. a consumption subsidy. The purpose of this subsidy is, of course, to mitigate the monopoly distortion in country A both directly by lowering the consumer price, and indirectly by inducing a higher volume of parallel imports.

Proposition 1 follows:

Proposition 1 *Under destination-based commodity taxes, parallel imports take place in equilibrium if and only if $g < 0.348$. In this case, the equilibrium tax $\widehat{t}_A^D$ is always negative, while the equilibrium tax $\widehat{t}_B^D$ is positive provided $\alpha > \frac{1}{33}(8 + 21g)$. Moreover, as the international transaction cost g falls below 0.348, the tax differential $(\widehat{t}_B^D - \widehat{t}_A^D)$ increases.*

Proof: Substituting $\widehat{t}_A^D$ into (7), $m > 0$ if $g < 0.348$. Thus, in the equilibrium with parallel imports, taxes can take the following values:

$$\widehat{t}_A^D \in \underset{(g=0)}{(-0.575,} \underset{(g=.348)}{-0.623)} \qquad \text{and} \qquad \widehat{t}_B^D \in (\underset{(g=0)}{\frac{\alpha}{3} - 0.081,} \underset{(g=.348)}{\frac{\alpha}{3} - 0.1546}). \tag{11}$$

Using (10), $\widehat{t}_B^D > 0$ when $\alpha > \frac{1}{33}(8 + 21g)$. Moreover, the tax differential unambiguously increases as g falls since $\frac{d(\widehat{t}_B^D - \widehat{t}_A^D)}{dg} < 0$. $\qquad\qquad QED$

It is interesting to compare the above equilibrium with the one without parallel imports, i.e. when $g > 0.348$. In this case, the manufacturer and the government in country A have well defined objectives. The manufacturer corrects the double marginalization problem that it faces in country B by setting its wholesale price equal to the marginal cost of production $(w = 0)$,[12] and the government in country A focuses on the monopoly distortion by setting $\widehat{t}_A^D = -1$. This leads to marginal cost pricing and thus to a consumer price that is equal to zero. In country B, there is no production and so the only goal of the government is to extract rents from the foreign manufacturer. As a result, the equilibrium tax rate is $\widehat{t}_B^D = \alpha/3$.

Figure 1 below illustrates the behavior of the Nash tax differential $(\widehat{t}_B^D - \widehat{t}_A^D)$ as a function of g. For $g > 0.348$, the tax differential is highest and independent of g (and thus of parallel imports). As g falls below its critical level, higher volumes of parallel imports have a pro-competitive effect in country A. Anticipating the manufacturer's discrete change of w at $g = 0.348$, governments adjust their tax rate by reducing both the optimal subsidy in country A (from 1 to 0.623) and the optimal tax in country B (from $\frac{\alpha}{3}$ to $\frac{\alpha}{3} - 0.1546$). As a result, the tax differential falls discretely to its lowest level. However, further decreases in g (and thus higher volumes of parallel imports[13]) reduce even more the subsidy in country A, but increase the tax in country B (due to the tax externality described above). The combination of these two decisions unambiguously raise the tax differential.

---

[12]Note that there is a strong discontinuity at $g = 0.348$ concerning the wholesale price. This is due to the fact that, even when evaluated at $m = 0$, the impact of a change in w has a marginal effect on m and on the manufacturer's profit for a value of g arbitrarily smaller than 0.348 but has no such effect for g marginally above this value.

[13]By substituting (10) into (7) it is easy to show that the relation between parallel imports and transaction costs is monotonically negative, i.e. $dm/dg < 0$ when $g < 0.348$.
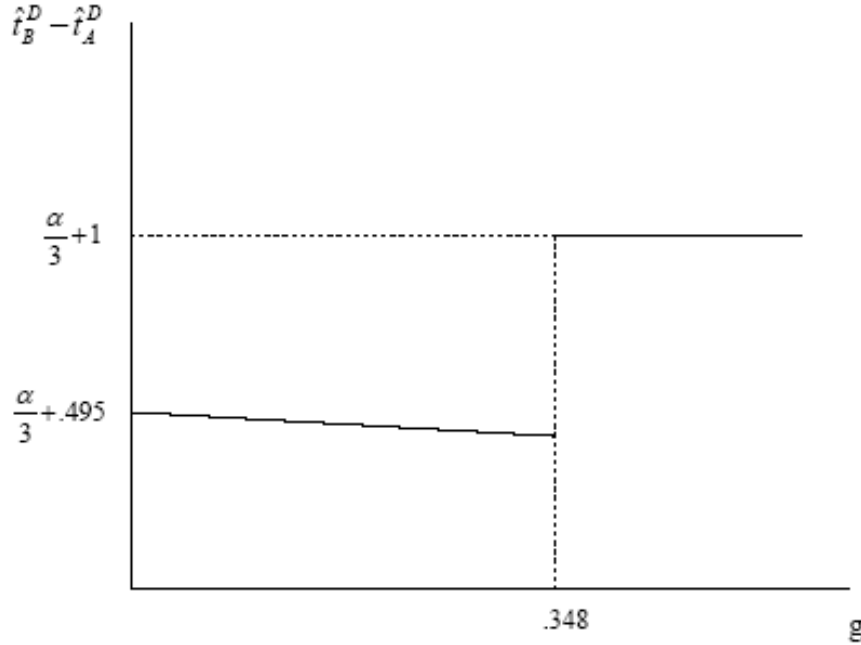
Figure 1: Destination taxes and parallel imports

To summarize, the model predicts that, when g falls below its critical value, an initial tax convergence takes place. However, further decreases in g, while increasing parallel imports, introduce fiscal externalities that lead to the divergence of the two countries' destination taxes.

3 The model with origin-based taxation

Having developed in detail the case of destination taxes, we can be brief with the origin-based taxes. Assuming that taxes are imposed and collected in the country where production takes place,[14] the profit and welfare equations are now

$$\Pi^M = (p_A - t_A^O)q_A + (w - t_A^O)(q_B + m) + F; \qquad (12)$$

$$\Pi^R = p_B q_B + (p_A - g)m - w(q_B + m) - F, \qquad (13)$$

---

[14]We carry the analysis under the same assumption as in the rest of the literature, viz. that taxes are based on production units. We will alter this assumption later on.

and

$$
\begin{aligned}
W_A^O &= CS_A + \Pi^M + t_A^O(q_A + q_B + m); & (14) \\
W_B^O &= CS_B, & (15)
\end{aligned}
$$

where t_A^O is the origin tax in country A. Country B has no tax revenue as it has no production to tax.

Using the same game as in the previous section, the output and parallel import decisions as functions of taxes, the wholesale price, and the parameters of the model are

$$
q_A(w, t_A^O) = \frac{1 - 2t_A^O + g + w}{3}; \quad m(w, t_A^O) = \frac{1 + t_A^O - 2g - 2w}{3};
$$
(16)

$$
q_B(w) = \frac{\alpha - w}{2}.
$$

From the second stage of the game,

$$
w(t_A^O) = \frac{2 + 11t_A^O + 8g}{13}, \tag{17}
$$

which, once substituted into (16), gives

$$
q_A(t_A^O) = \frac{5(1 - t_A^O) + 7g}{13}; \quad m(t_A^O) = \frac{3(1 - t_A^O) - 14g}{13};
$$
(18)

$$
q_B(t_A^O) = \frac{13\alpha - 2 - 11t_A^O - 8g}{26}.
$$

Thus, a higher t_A^O decreases again sales and parallel imports in country A but, *contrary to the destination case*, it increases the wholesale price w.[15] This difference stems from the fact that, while the retailer still decides about m, the manufacturer now pays the origin-based (production) tax. The only way the manufacturer is now able to reduce parallel imports is by increasing its wholesale price, and thus by increasing the retailer's cost in B.

Moving to the optimal choice of taxes, it is clear that country B's optimal tax is $\widehat{t}_B^O = 0$, as the country has no tax base and is thus unable to extract

---

[15]It is interesting to note that, despite different manufacturer's objective functions, the solutions of the second stage of the game for q_A and m are the same in the two tax regimes (see (7) and (18)). This is because the manufacturer uses two tools (w and F) allowing it to 'neutralize' the effect of the tax regime (but of course not of the tax level) on the quantity decisions.

rents from the foreign manufacturer (we alter this assumption later on). In country A, the welfare-maximizing tax is

$$\widehat{t}_A^O = -\frac{2}{249}[51 + 61g].\tag{19}$$

Proposition 2 follows:

Proposition 2 *Under origin-based commodity taxes, parallel imports take place in equilibrium if and only if $g < 0.337$. The equilibrium tax $\widehat{t}_A^O$ is always negative, while the equilibrium tax for country B is always zero. Moreover, as the international transaction cost g falls, country A's subsidy decreases. Consequently, the equilibrium tax differential $(\widehat{t}_B^O - \widehat{t}_A^O)$ also decreases.*

Proof: Substituting (19) in (18), $m > 0$ if $g < 0.337$. Accordingly, the welfare-maximizing tax rate in country A is negative and takes the values

$$\widehat{t}_A^O \in \underset{(g=0)}{(-0.41,} \underset{(g=.337)}{-0.575)}.\tag{20}$$

It is straightforward to see from (19) that $\frac{d\widehat{t}_A^O}{dg} < 0$ and thus that $(\widehat{t}_B^O - \widehat{t}_A^O)$ necessarily decreases when g falls. $\qquad\qquad QED$

The intuition should be straightforward by now: as g falls, parallel imports increase,[16] and thus the government in country A reduces its subsidy as the competition induced by parallel imports helps correcting the monopoly distortion. Given that country B has a zero tax, the tax differential between the two countries necessarily gets smaller as g falls.

In the absence of parallel imports ($g > 0.337$), the government in country A mitigates the monopoly distortion by setting a production subsidy equal to $\widehat{t}_A^O = -1/3$. Note that this subsidy is lower than in the destination case $(\widehat{t}_A^D = -1)$. The reason for this is that a production subsidy increases not only the output the manufacturer sells in the home country, but also the output sold to the foreign country. Thus, the government now takes into account the terms of trade effect of its policy (aimed at correcting the monopoly distortion).

Figure 2 illustrates the tax differential $(\widehat{t}_B^O - \widehat{t}_A^O)$ as a function of g. As the international transaction cost falls below its prohibitive value, parallel imports increase in country A, and the country reacts by raising its optimal subsidy (from 0.333 to 0.575). This initial jump of the subsidy can be explained by the fact that the existence of parallel imports induces the manufacturer to increase its wholesale price, and thus country A's terms of trade,

---

[16]To prove that parallel imports increase as g falls, substitute (19) into (18) and compute dm/dg.

which reduces sales to country B. With the terms-of-trade effect now been taken care of, the government in country A chooses its optimal subsidy to address the monopoly distortion by increasing its production subsidy. As before, further decreases in g lead to lower subsidy in A since more competition brought by parallel imports reduces the need for such a subsidy. With $\widehat{t}_B^O$ being always zero, the tax differential $(\widehat{t}_B^O - \widehat{t}_A^O)$ then clearly falls.
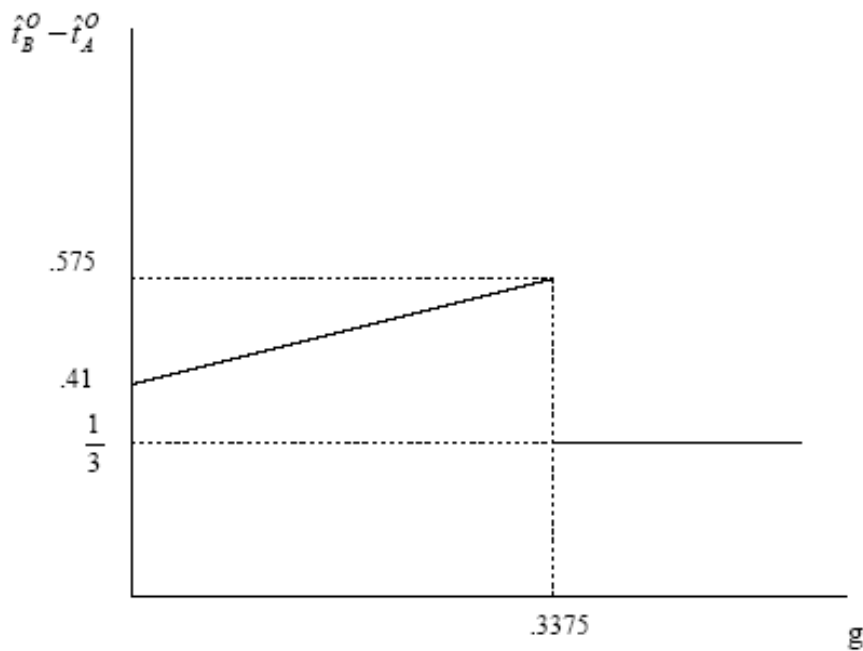

Figure 2: Origin taxes and parallel imports

To summarize, the model predicts that when g falls below a critical value, an initial tax divergence takes place. However, further decreases in g lead to a convergence of the two countries' origin taxes.

Before comparing the two tax regimes in detail, we should note that, although our tax instrument (per unit tax) is the same as the one used in the literature, the origin tax regime boils down to a relatively trivial outcome in our model as no production takes place in the retailer country. We therefore briefly consider the case where country B has a tax base. To do so, we assume that the government in country B can tax the value that the retailer is adding to that country (i.e., the retailer's gross profit at rate t'_B) while the government in country A still imposes a consumption origin-based tax t_A^O.

Thus, while the manufacturer attempts to capture the retailer's gross profit through its two-part pricing strategy, the government in country B taxes the retailer's gross profit before it is 'exported' to country A. This means that (13) and (14) become respectively:

$$\Pi^R = (1 - t'_B)[p_B q_B + (p_A - g)m - w(q_B + m)] - F; \qquad (21)$$
$$W'_B = CS_B + t'_B[p_B q_B + (p_A - g)m - w(q_B + m)]. \qquad (22)$$

This modification does not affect the third stage of the game. It does affect, however, the manufacturer's wholesale pricing decision which now turns out to be:

$$w(t_A^O, t'_B) = \frac{2 + 11t_A^O + 8g + t'_B(9\alpha + 8 + 8t_A^O - 16g)}{13 + 25t'_B}.$$

Irrespective of its incentive about parallel imports, the manufacturer has now an incentive to raise w in order to decrease the retailer's gross profit that is taxed by government B. Not surprisingly, in equilibrium, government B chooses a strictly positive tax rate $\widetilde{t}'_B > \widehat{t}_B^O = 0$ (where tilda denotes the Nash values when country B taxes the retailer on its gross profit). Government A now selects an even higher subsidy rate in comparison to the case where country B does not have a tax base, i.e. $\widetilde{t}_A^O < \widehat{t}_A^O$. This unambiguously results in a higher tax rate differential, i.e. $\widetilde{t}'_B - \widetilde{t}_A^O > \widehat{t}_B^O - \widehat{t}_A^O$. Although several equilibria might exist, the most reasonable one has the same convergence properties as illustrated in Figure 2.[17],[18]

4   Origin vs. destination taxes

We now compare in detail the destination and the origin taxes and examine the welfare implications of the two tax regimes.

Starting with the tax levels, consider the case without parallel imports ($g > 0.348$). In this case, the incentives in country A are quite different in the two tax systems. With origin taxation, the manufacturer's cost of

---

[17]The existence of several equilibria is due to the fact that $w(t_A^O, t'_B)$ is a non-linear function of t'_B, and thus the best replies in (t_A^O, t'_B) are non-linear as well. By most reasonable equilibrium, we mean the one involving the smallest absolute deviation with respect to $(\widehat{t}_A^O, \widehat{t}_B^O)$.

[18]We should note here that we tried to model origin based, value added taxes as they would be applied in the EU, viz. with an invoice-credit procedure attached to it (where imported inputs are exempted from value-added taxation — see Ebrill et al., 2001, pp. 180). However, our model of vertical control does not provide much insights as to how commodity taxes should be designed — the best-reply functions with respect to taxes are 3rd degree polynomials with no simple solutions.

selling abroad one unit of output is equal to the tax rate t_A^O, and thus the manufacturer sets $w = t_A^O$ to the foreign retailer. Hence, a change in t_A^O not only affects consumer surplus and production sold in country A but also exports to country B. This induces country A's government to reduce the optimal subsidy to the manufacturer (as compared to the destination base case where the destination tax does not influence the marginal cost of selling a unit abroad). As a consequence, country A's destination-based subsidy is higher than its origin-based subsidy, i.e. $\widehat{t}_A^D = -1 < \widehat{t}_A^O = -1/3$. For country B, the fact that it uses the destination-based tax as a rent extraction instrument, while it has no tax base under the origin principle, makes the comparison obvious, i.e. $\widehat{t}_B^D = \alpha/3 > \widehat{t}_B^O = 0$.

A similar result can be derived when parallel imports take place under both tax regimes ($g < 0.337$). In this case, A's destination-based subsidy is always higher than its origin-based subsidy, i.e. $\widehat{t}_A^D < \widehat{t}_A^O$. Similarly for country B, the destination-based rent extraction tax is positive,[19] while it is zero under the origin principle. Hence, independently of whether parallel imports take place or not, the following relation is true:

$$\widehat{t}_B^D > \widehat{t}_B^O > \widehat{t}_A^O > \widehat{t}_A^D.$$

It immediately follows that, regardless of g and α, the tax rate differential is systematically lower under origin taxation than it is under destination taxation, i.e. $\widehat{t}_B^O - \widehat{t}_A^O < \widehat{t}_B^D - \widehat{t}_A^D$.

Moreover, and as already shown, the tax rate differential decreases with lower values of g under origin taxation, and it increases under destination taxation, i.e. $\frac{d(\widehat{t}_B^O - \widehat{t}_A^O)}{dg} > 0$ and $\frac{d(\widehat{t}_B^D - \widehat{t}_A^D)}{dg} < 0$. Proposition 3 summarizes the above findings.

Proposition 3 *Origin tax rates are unambiguously more similar than the corresponding destination tax rates. Moreover, as g falls and parallel imports increase, taxes/subsidies converge under the origin principle and diverge under destination principle.*

The comparison of the tax levels, however, says nothing about the desirability of one tax regime as compared to the other. Therefore, we now investigate the welfare properties of the two commodity tax regimes in the presence or not of parallel imports.

A and B's equilibrium welfare levels can be found by using (3) and (4) for the destination tax case, and (14) and (15) for the origin tax case once the equilibrium values of the endogenous choices are taken into account (see

---

[19]At least as long as $\alpha > \frac{1}{33}(8 + 21g)$ (see Proposition 1).

Appendix for details). Figures 3 and 4 illustrate the resulting welfare levels
as a function of g when the two countries have the same willingness to pay
(or size), i.e. $\alpha = 1$ (similar figures can be drawn for $a < 1$).
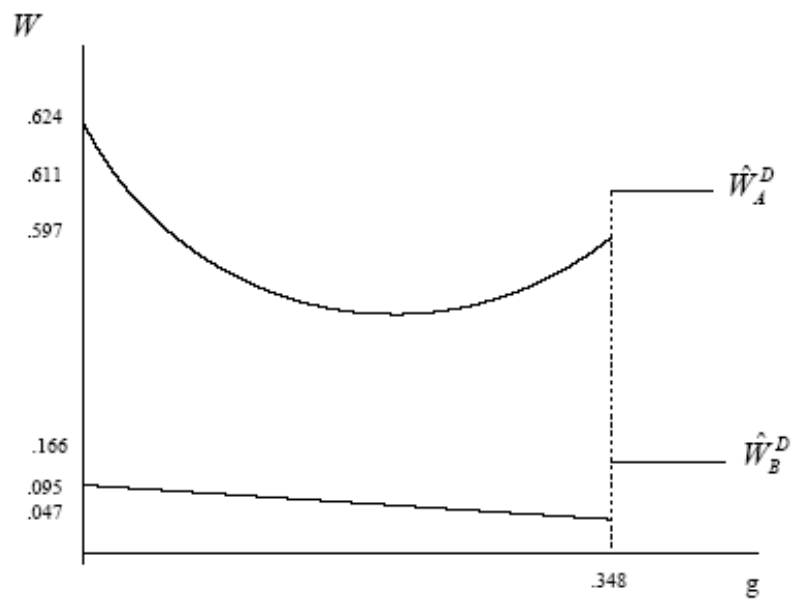

Figure 3: Welfare under destination taxes $(a = 1)$
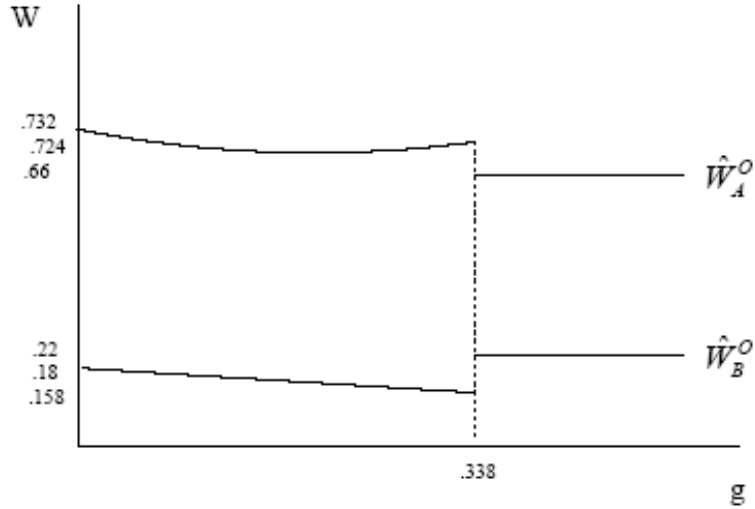
Figure 4: Welfare under origin taxes $(a = 1)$

A careful look at the figures reveals some interesting results:

1. Welfare levels in both countries are highest under origin taxes. That is, for any value of g, both countries prefer origin taxes to destination taxes, i.e. $W_i^O > W_i^D, i = A, B$. *In other words, if the tax regime was a choice, the two countries would choose origin taxes.*[20]

2. Parallel imports are beneficial to the manufacturing country if taxes are origin based (i.e., $\widehat{W}_A^O(g < .338) > \widehat{W}_A^O(g \geq .338)$). If taxes are destination based, the manufacturing country gains from parallel imports only if the resource cost g is very low (i.e., $\widehat{W}_A^D(g) > \widehat{W}_A^D(g \geq .348)$ only for g small enough). For the retailer country, the introduction of parallel imports is detrimental regardless of the tax regime. *Thus, if the level of transaction cost on parallel import was a choice, the retailer country would choose g high enough to avoid parallel imports under both tax regimes. The manufacturing country however would always choose g so as to have parallel imports: under the origin-based tax regime, it sets $g < 0.338$, and under the destination-based regime, it sets a sufficiently low g.*

---

[20]See Keen and Lahiri (1993, 1998) for a similar result where perfect market integration is assumed.

3. The introduction of parallel imports leads to a jump up of *total* welfare under origin taxes. Further increases in parallel imports first reduce total welfare slightly and then increase it. Still, the reduction is not as large as the initial increase so that total welfare is unambiguously higher with parallel imports than without them. Exactly the opposite changes occur with destination taxes, i.e. total welfare first decreases and then increases, but this increase is not enough to compensate for the initial fall. Thus, total welfare is unambiguously lower with parallel imports than without them. *In other words, parallel imports improve total welfare only when taxes are origin based.*

Proposition 4 summarizes the above points:

Proposition 4 *In the presence of parallel imports, both countries prefer having an origin-based regime to a destination-based regime. However, if countries could choose whether or not to reduce the parallel imports transaction cost, they would generally disagree irrespective of the tax regime. A supranational authority (such as the EU) would promote parallel imports, and thus would reduce the transaction cost, if and only if taxes were origin based.*

Proof: See Appendix.

The above proposition makes clear, that if the supranational authority promoted parallel imports, the member countries would have an incentive to adopt origin-based taxes. In that case, the sum of countries' welfare will increase with the retailer country incurring a loss and the producing country incurring a gain. Clearly, some redistribution would be necessary to get strict Pareto improvements.

Before we conclude, we should note that parallel imports under destination taxes are *larger* than they are under origin taxes ($m^D > m^O$).[21] One might have expected the opposite outcome given the facts that origin taxation is preferred to destination taxation, and that parallel imports promote market integration. However, one should remember that, in the present model, parallel imports do not have only positive effects. Specifically, parallel imports create two sources of inefficiency and one source of efficiency. The two inefficiencies are the resource cost (g) incurred with each unit of parallel import, and the double marginalization associated with the manufacturer's pricing decision. The efficiency source is the pro-competitive effect generated by parallel imports. Whether parallel imports are good or not is thus a matter of the relative size of these effects. In particular, like in Maskus and Chen, we also find a U-shaped relationship between total welfare and g independently of whether taxes are destination- or origin-based.

---

[21] To compute $m^D - m^O$, substitute first (10) and (19) into (7) and (18).

5 Conclusions

Economic integration can take several forms. One of them is to encourage parallel imports as the EU has repeatedly done over the last decades. If this form of economic integration has been considered as important, it is because many markets are vertically related and thus not only manufacturers take decisions about given products, but wholesalers and retailers as well. Indeed, there is little point to liberalize standard barriers to trade if, at the same time, contractual arrangements, contribute to segment markets and prevent or limit trade from taking place. Making parallel imports legal at the level of the European Union has made this market segmentation more difficult and thus has made more effective the liberalization of barriers to trade.

This paper shows that, in addition, the existence of parallel imports is not without consequences on the choice of commodity tax regime. In particular, we have shown that, in a world in which free trade prevails for authorized flows of products, the encouragement of parallel imports brings about a tax convergence between trading countries and a higher collective level of welfare (as compared to the case without parallel imports) if commodity taxation is origin based. This is in sharp contrast with the case of destination-based taxes where parallel imports are shown to bring both a tax divergence and a lower total welfare (as compared to the case without parallel imports). Thus, by including non-cooperative taxes into a model of parallel-imports, we have shown that origin taxation is unambiguously preferable to destination taxation *even* in a set up where markets are not perfectly integrated. More importantly, this superiority of origin taxes comes with an extra windfall: countries' taxes rates convergence. Thus, the superiority of origin taxes does not conflict with the desire to bring tax rates closer to each other. To our knowledge, this is a novel result to the literature.

The research agenda that we have proposed here points towards tax competition models that lead to a tax convergence/harmonization as an endogenous reaction to the countries' choices. Tax harmonization as a policy that the countries would happily implement, has not been very successful in practice. Countries naturally protect their right to choose taxes, and perhaps we should try to find mechanisms that give them an incentive to choose similar taxes. Market integration through the encouragement of parallel imports is the mechanism that we have proposed in this paper.

6   Appendix

In the equilibrium with parallel imports, the welfare levels in the destination- and in the origin-based regimes are respectively

$$\widehat{W}_A^D(g) = \frac{1}{99}[\frac{4768}{99} + \frac{3}{11}(2784g - 1457) + \frac{\alpha}{3}(8 + 33\alpha) + 7\alpha g];$$

$$\widehat{W}_B^D(g) = \frac{3}{2}(\frac{\alpha}{3} - \frac{8}{99} - \frac{7}{33}g)^2;$$

$$\widehat{W}_A^O(g) = \frac{\alpha^2}{4} + \frac{1}{83}[40 - 28g + \frac{461}{6}g^2];$$

$$\widehat{W}_B^O(g) = \frac{1}{2}(\frac{\alpha}{2} + \frac{8}{83} - \frac{25}{249}g)^2.$$

It can then be checked that the following results hold:

1. $\widehat{W}_A^O(g) - \widehat{W}_A^D(g) > 0$ and $\widehat{W}_B^O(g) - \widehat{W}_B^D(g) > 0$ for all feasible values of α and g consistent with parallel imports.

2. Irrespective of α, $sign\frac{\partial \widehat{W}_A^O}{\partial g} < 0$ for $0 \le g < .1822$ and $sign\frac{\partial \widehat{W}_A^O}{\partial g} > 0$ for $.1822 < g < .338$, while $sign\frac{\partial \widehat{W}_B^O}{\partial g} < 0$ for $0 \le g \le 0.338$.

3. $sign\frac{\partial \widehat{W}_A^D}{\partial g} < 0$ for $0 \le g < \frac{1}{1856}(\frac{1457}{3} - 77\alpha)$ and $sign\frac{\partial \widehat{W}_A^D}{\partial g} > 0$ for $\frac{1}{1856}(\frac{1457}{3} - 77\alpha) < g \le 0.348$, while $sign\frac{\partial \widehat{W}_B^D}{\partial g} < 0$ for $0 \le g < \frac{1}{7}(11\alpha - \frac{8}{3})$ and $sign\frac{\partial \widehat{W}_B^D}{\partial g} > 0$ for $\frac{1}{7}(11\alpha - \frac{8}{3}) < g \le 0.348$.

These conditions imply that $\widehat{W}_A^D(g)$ and $\widehat{W}_A^O(g)$ are U-shaped with respect to g, while $\widehat{W}_B^D(g)$ and $\widehat{W}_B^O(g)$ are negatively sloped with respect to g (unless $\alpha < 0.464$ in which case $\widehat{W}_B^D(g)$ is U-shaped as well).

In the absence of parallel imports, the equilibrium levels of welfare in the destination and in the origin cases are respectively

$$\widehat{W}_A^D(m = 0) = \frac{1}{2} + \frac{\alpha^2}{9}; \ \widehat{W}_B^D(m = 0) = \frac{\alpha^2}{6};$$

$$\widehat{W}_A^O(m = 0) = \frac{7}{18} - \frac{\alpha}{6} + \frac{(3\alpha + 1)^2}{36}; \ \widehat{W}_B^O(m = 0) = \frac{(3\alpha + 1)^2}{72}.$$

It can then be checked that:

4. $\widehat{W}_A^O(m = 0) > \widehat{W}_A^D(m = 0)$ holds for $\alpha > .774$ and $\widehat{W}_B^O(m = 0) > \widehat{W}_B^D(m = 0)$ for all feasible values of α.

5. $\widehat{W}_A^D(g) < \widehat{W}_A^D(m = 0)$ unless α is high and g is low. $\widehat{W}_B^D(g) < \widehat{W}_B^D(m = 0)$ for all feasible values of α.

6. $\widehat{W}_A^O(g) > \widehat{W}_A^O(m = 0)$ irrespective of the value of α. Similarly, $\widehat{W}_B^O(g) < \widehat{W}_B^O(m = 0)$ for all feasible values of α.

References

[1] Ahmadi, R. and B.R. Yang, 2000. Parallel Imports: Challenges from Unauthorized Distribution Channels, *Marketing Science* 19, 279-94.

[2] Behrens, K., J. Hamilton, G. Ottaviano, and J.-F. Thisse, 2007. Destination- versus origin-based commodity taxation and the location of industry, *Journal of International Economics* 72, 271-291.

[3] Ebrill, L., M. Keen, J.-P. Bodin, and V. Summers, 2001. *The Modern VAT*, International Monetary Fund, Washington D.C.

[4] Ganslandt, M. and K.E. Maskus, 2004. The price impact of parallel trade in pharmaceuticals: evidence from the European Union, *Journal of Health Economics* 23, 1035-1057.

[5] Ganslandt, M. and K.E. Maskus, 2007. Vertical distribution, parallel trade, and price divergence in integrated markets, *European Economic Review* 51, 943-970.

[6] Grosmann, G. and E. Lai, 2006. Parallel imports and price controls, CEPR DP 5779, London.

[7] Haufler, A. and M. Pflüger, 2006. International oligopoly and the taxation of commerce with revenue constrained governments, *Economica*, forthcoming.

[8] Haufler, A., G. Schjelderup, and F. Stahler, 2005. Barriers to trade and imperfect competition: the choice of commodity tax base, *International Tax and Public Finance* 12, 281-300.

[9] Kanavos, P. and J. Costa-Font, 2005. Pharmaceutical parallel trade in Europe: stakeholder and competition effects, *Economic Policy*, October, 751-798.

[10] Keen, M. and S. Lahiri, 1993, Domestic tax reform and international oligopoly, *Journal of Public Economics* 51, 55-74.

[11] Keen, M. and S. Lahiri, 1998. The comparison between destination and origin principles under imperfect competition, *Journal of International Economics* 45, 323-350.

[12] Keen, M. S. Lahiri, and P. Raimondos-Møller, 2002. Tax principles and tax harmonization under imperfect competition: a cautionary example, *European Economic Review* 46, 1559-1568.

[13] Li, C. and K. Maskus, 2006. The impact of parallel imports on investments in cost-reducing research and development, *Journal of International Economics* 68, 433-455.

[14] Lockwood, B., 2001. Tax competition and tax coordination under destination and origin principles: a synthesis, *Journal of Public Economics* 81, 279-319.

[15] Malueg, D.A., Schwartz, M., 1994. Parallel imports, demand dispersion and international price discrimination, *Journal of International Economics* 37, 167-96.

[16] Maskus, K.E., 2000. Parallel Imports, *The World Economy* 23, 1296-1284.

[17] Maskus, K.E. and Y. Chen, 2002. Parallel imports in a model of vertical distribution: theory, evidence, and policy, *Pacific Economic Review* 7, 319-334.

[18] Maskus, K.E. and Y. Chen, 2004. Vertical price control and parallel imports: theory and evidence, *Review of International Economics* 12, 551-570.

[19] Maskus, K.E. and Y. Chen, 2005. Vertical pricing and parallel imports, *Journal of International Trade and Economic Development* 14, 1-18.

[20] NERA, National Economic Research Associates, 1999. *The Economic Consequences of the Choice of Regime in the Area of Trademarks*, London: National Economic Research Associates.

[21] OECD, 2002. *Synthesis Report on Parallel Imports*, Paris.

[22] Official Journal, 2001. Commission Decision of 8 May 2001, IV/36.957/F3, *Official Journal of the European Communities* L302. Vol. 44, November 17.

[23] Official Journal, 2002. Commission Decision of 21 December 2000, COMP.F.1/35.918, *Official Journal of the European Communities* L69. Vol. 45, March 12.

[24] Raff, H. and N. Schmitt, 2007. Why parallel trade may raise producers' profits, *Journal of International Economics* 68, 485-503.