DISCUSSION PAPER SERIES

No. 5022

PARALLEL TRADE, INTERNATIONAL EXHAUSTION AND INTELLECTUAL PROPERTY RIGHTS: A WELFARE ANALYSIS

Tommaso Valletti and Stefan Szymanski

INDUSTRIAL ORGANIZATION

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP5022.asp**

ISSN 0265-8003

PARALLEL TRADE, INTERNATIONAL EXHAUSTION AND INTELLECTUAL PROPERTY RIGHTS: A WELFARE ANALYSIS

Tommaso Valletti, Imperial College, London and CEPR
Stefan Szymanski, Imperial College, London

Discussion Paper No. 5022
April 2005

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 5022

April 2005

ABSTRACT

Parallel Trade, International Exhaustion and Intellectual Property Rights: A Welfare Analysis*

This paper analyses the issue of parallel trade (arbitrage) for products protected by intellectual property rights. Many countries have traditionally allowed owners of intellectual property rights to prohibit arbitrage in the face of international price discrimination. In a well-known paper Malueg and Schwartz (1994) showed that this policy decreases social welfare when the same markets are served in both regimes, with and without arbitrage. Their model considered only the setting of prices, and not investment in product development. We consider a two-stage game where firms choose quality first and then prices. Since the threat of arbitrage *ex post* reduces the incentive to invest *ex ante*, the net benefits of parallel trade may vanish. We also show that the size of the welfare effects is significantly affected by the presence of a 'generic' product, which represents a form of competition for the monopolist. The monopolist will introduce a 'fighting brand' to compete with the generic, which dilutes but does not eliminate the result on the adverse effects of parallel trade on investments.

JEL Classification: F13, L12 and O34
Keywords: investments, parallel trade and price discrimination

Tommaso Valletti
Imperial College London
Tanaka Business School
South Kensington Campus
London
SW7 2AZ
Tel: (44 20) 7594 9215
Fax: (44 20) 7823 7685
Email: t.valletti@ic.ac.uk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=145352

Stefan Szymanski
Imperial College London
Tanaka Business School
South Kensington Campus
London
SW7 2AZ
Tel: (44 20) 7594 9107
Fax: (44 20) 7823 7685
Email: szy@ic.ac.uk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=109002

*We thank seminar participants in Antwerp, Bristol, London, Naples and Toulouse for useful comments. This research has been supported by ESRC research grant no. R022250213. We also thank two anonymous referees and the Editor for their valuable comments.

Submitted 05 April 2005

1. Introduction

Parallel trade of goods protected by intellectual property rights is an important but little understood issue in international trade. Parallel trade (sometimes called "grey market" trade) involves the shipment of *bona fide* goods (i.e. not illegal counterfeits) across international borders in order to exploit price differences[1]. In a free-trade environment parallel trade prevents monopoly suppliers from engaging in international price discrimination. However, where the good is protected by an intellectual property right, such as a patent or trademark, this right may permit the owner to prohibit international arbitrage.

Examples of parallel trade involving an intellectual property right would be:

- The export of pharmaceuticals from the USA to Canada, which are then resold to US consumers by internet pharmacies
- The supply of AIDS drugs at discounted prices for sale in Africa which are then resold by distributors in European markets.
- The purchase of Levi jeans by wholesalers from US distributors for resale in European markets.

Each of these examples involves an infringement of the law, although in each case the rationale is slightly different. In general the US applies what is known as the "first sale doctrine". According to this, once a good has been first placed on market, the seller or manufacturer forfeits all rights to determine how that product will subsequently be disposed of. This effectively rules out price discrimination against American consumers, since a purchaser would always possess the right to resell into the US. However, US law makes an exception of pharmaceutical products since it prohibits the importation of pharmaceutical products by anyone other than the manufacturer (21 USC 381(d)). During the 2004 Presidential election campaign this became a major issue as many States encouraged consumers to buy direct from Canada or from internet pharmacies despite the dubious legality of such acts under Federal law. Most American consumers do not see why they should pay higher prices than consumers in Canada or elsewhere.

[1] Thus parallel trade is a form of arbitrage; the term "parallel" trade is used because in many cases the good is shipped back to the country from which it originated: the good is returning whence it came. Parallel trade might equally involve the shipment of goods produced under licence in one country back to the country of the licensor, the key point is that such trade is motivated by international price differences.

Historically, national intellectual property law in countries such as the UK and France has given exporters who possess a patent or trademark the right to prohibit such parallel trade. This can be done, for example, simply by marking the goods "not for resale in the UK" or "not for resale in France". More recently, the European Union has overridden this national law by adopting a policy known as "Community Exhaustion", meaning that once goods have been placed on the market in the European Union, the holder of the intellectual property right no longer has the right to restrict the further movement of the goods anywhere inside the European Union. The term "exhaustion" is used because the rights of the owner of the intellectual property are said to be "exhausted" once the goods have been placed on the market. Under Community Exhaustion a UK exporter could not prevent the resale of AIDs drugs first sold in France back into the UK, but could prohibit the re-entry of products made available for sale at low prices in Africa[2]. Absent such a guarantee, AIDs drugs would have to be sold at uniform prices across the world, and if these prices were to be affordable for African consumers, there would need to be substantial public subsidies for R&D.

Parallel trade in products such as pharmaceuticals within the EU is now a significant business activity. For example, one recent study estimated that 20% of branded pharmaceuticals sold in the UK in 2002 were parallel sourced (value about £1.3bn)[3]. As one might expect, pharmaceutical companies claim that this is undermining their profitability. As yet the EU policy only extends as far the EU's borders. However, many developing countries are lobbying for the adoption of a policy of "international exhaustion" to be adopted at the WTO. This would mean that goods placed on the market anywhere in the world could then be resold anywhere else. This appears to be motivated by the hope that goods produced under license in developing countries, which are currently prevented from being imported into the licensor's domestic market by the same intellectual property laws, could be exported and generate large profit for the licensees.

That the potential profits from parallel trade of this kind are significant is not in doubt. Prior to entering the EU, Sweden operated a system of international exhaustion, permitting importers to source parallel traded goods from anywhere in the world. For example, if Levi

[2] A number of cases in the European Court of Justice has established this principle, most notably *Silhouette v. Hartlauer (Case C-335/96).* Silhouette, an Austrian company, that sold spectacles under its trademark sold an outdated batch to a Bulgarian company for resale in the Former Soviet Union. However, the distributor then tried to put them on the market in Austria, and the court upheld Silhouette's right to prevent this under its trademark.
[3] York Health Economics Consortium (2003).

jeans could be purchased more cheaply in America than in Europe, a Swedish importer was free to purchase in America and resell in Sweden, even if the goods had originally been intended (by Levi Strauss & Co.) for sale in America[4]. However, on entering the EU Sweden was obliged to respect the rights of any holder of a Community trademark to prevent parallel trade from *outside* the EU. A study by the Swedish Competition Authority (1999) estimated that parallel traded goods in sectors such as motor cycle spare parts, tyres, clothing, footwear, pharmaceuticals, sports equipment and snow scooters were between 10 and 30% cheaper than domestically sourced goods and that the elimination of parallel trade from outside the EU had increased domestic prices by between 0.4 and 5% on average.

This argument raises some interesting economic questions, although there have been surprisingly few papers on the subject[5]. If parallel trade is permitted, why does the producer make low priced goods available in some markets when it knows that they will be parallel traded? If arbitrage undermines profitability in the high price market the producer may choose to offer to sell at a single price across all markets, possibly leaving some markets unserved.

It is clear from this argument that in the short run parallel trade must benefit consumers in high price market but must harm consumers in markets that would have low prices if arbitrage were not possible[6]. The producer's profit also suffers, suggesting that the welfare trade-offs are difficult to judge. This issue is addressed in Malueg and Schwartz (1994) who build on a paper by Varian (1985) and develop a model of a single product sold in a continuum of markets, each characterised by the maximum willingness-to-pay of any consumer in that market. They show that aggregate welfare under international exhaustion (arbitrage) is higher than under international price discrimination (no exhaustion of rights) so long as international exhaustion does not imply the closure of any market. However, if the willingness-to-pay of

[4] This specific issue arose in the European Court in the case of *Davidoff v. A&G Imports* and *Levi Strauss v. Tesco (Case c-414/99).* Tesco, a UK retailer, attempted to circumvent the EU distribution channels of Levi Strauss by buying direct from US wholesalers - the court upheld the right of Levi Strauss under their trademark to prohibit these imports.

[5] However there have been a number of policy oriented reviews - see e.g. Tarr (1985), and Danzon and Towse (2003). Abbot (1998) provides a detailed legal statement of the case for international exhaustion while New Zealand Institute of Economic Research (1999) provides an economic rationale; NERA (1999) provide the case against. The 2002 OECD report provides a useful synthesis. Maskus and Chen (2004) look at parallel trade in relation to vertical price controls and Richardson (2002) makes the simple but important point that prohibiting international exhaustion is not a Nash equilibrium in a game played between governments.

[6] In the presence of increasing returns to scale, price uniformity can negatively affect everybody (Hausman and Mackie-Mason, 1988). Also notice how the welfare implications seem quite different depending on the product - for example, both perfume and AIDS treatments will likely be cheaper in poor countries if parallel trade is

consumers is sufficiently dispersed between markets then those markets where consumer valuations are very low will be closed and aggregate welfare may fall.

We extend their paper in several ways. First we endogenise the quality of the good sold[7]. Most pharmaceutical companies claim that their R&D spending will fall under a regime of international exhaustion. We discuss under what circumstances, when the producer determines quality endogenously, this claim can be supported. Second, we model the case where an alternative product exists, which we label a "generic". Third, we allow the incumbent to supply more than one product.

Under pure monopoly, we discuss a basic trade-off that arises between the *ex post* better allocation that typically occurs under parallel trade when demand dispersion is not too high, and the *ex ante* reduced product quality because of lower investment. We also investigate the claim that the monopolist might react to parallel trade by offering multiple varieties. One might intuitively imagine that introducing a second brand dedicated to low valuation consumers might be a desirable option when arbitrage is possible. That is to say, one could think it would be profitable to restrict countries where a low price is optimal to an inferior brand, so that consumers in countries where a high price is optimal would not then see this inferior version as a substitute, and hence eliminate the negative effects of arbitrage on the monopolist's profit. We do not find support for this view: in our model the monopolist would never choose to introduce a second product absent competition from the generic.

When competition from a generic is present, we find that *ex post* the generic extends the range of demand dispersion under which international exhaustion is the preferred regime. However the presence of the generic worsens the *ex ante* incentives to invest and the quality reduction is magnified under parallel trade. We also show that the last result depends on the incumbent being restricted to offer only one product. When the patent holder is allowed to introduce a second lower-quality product, we show that it has an incentive to do so *only* when the generic

---

prohibited, yet in the first case the impact on consumers in poor countries does not seem particularly important, while in the latter case it is vital.

[7] With the exception of Rey (2003) and Li and Maskus (2004), to the best of our knowledge there are no papers that look at the dynamic aspects of parallel trade. Rey studies the impact of parallel trade on R&D in a game where governments contribute to finance investments through regulated prices. He argues that parallel trade reduces the ability of a government to make a conscious choice to invest in R&D by paying high prices while permitting foreign governments to negotiate lower prices. Li and Maskus consider a model where a manufacturer engages in cost-reducing R&D and sells his product into another country through a distributor. They find that parallel imports discourage investment in process innovation.

product is present. In such a case, the lower-quality variant is introduced to fight the generic, while profits made from the higher-quality good are protected. This generates an interesting "separation" result: while *ex post* the generic still extends the range of demand dispersion under which international exhaustion is the preferred regime, the *ex ante* incentives to invest are not affected by it but only by the exhaustion regime.

The paper is set out as follows. In the next section we outline our model assumptions and describe the benchmark case where quality is fixed (as in Malueg and Schwartz). For this case aggregate welfare is higher under international exhaustion (arbitrage) as long as not too many markets are closed. Then in Section 3 we show that when quality is endogenous any beneficial effects of arbitrage can be more than offset by the implied reduction in quality of the good in the market, even when all markets are served. In Section 4 we consider the effect that the presence of a non-strategic generic alternative (albeit of a lower quality) has on welfare. Finally, in Section 5 we consider the possibility that the strategic player might introduce more than one quality of product in the face of the non-strategic alternative. Section 6 concludes. All the mathematical proofs are relegated to an Appendix.

2. Model assumptions and preliminary results

We consider the price and product quality choice of a monopolist selling to (or licensing to) a set of distinct markets. We use a model of vertical product differentiation to represent consumer preferences in each market. When a consumer of type t buys a product of quality u at price p, his utility is $tu - p$. t is uniformly distributed between 0 and T with unit density.

There is a continuum of markets characterised by their highest type T. T is uniformly distributed over $[1 - x, 1 + x]$ with density $1/2x$, where $x < 1$ measures demand dispersion: a high value of x implies a large difference between the maximum willingness-to-pay across markets (as may be the case, for example, if income inequalities between countries are large).

The supply of a product with a given quality u involves a fixed cost, increasing in u. At times we will make use of a quadratic functional form: $C(u) = ku^2/2$. Thus this is a model of product innovation where more R&D enables the investing firm to discover "better" products. Once a product has been discovered, its marginal cost of production is not affected by the level of quality. We normalise the marginal cost of production to zero.

We model the problem as a two-stage game in which the firm first sets its quality level and then selects the associated price(s) according to the international exhaustion regime.

In the absence of international arbitrage the monopolist sells the same quality u_d at different prices in the various markets (the subscript *d* stands for price discrimination, while we use a subscript *u* for price uniformity). Let p_i denote the price in market *i*. The firm then sells to all customer types between the marginal type $\underline{t}_i = p_i/u_d$ and the highest type T_i in each market *i*. The firm sets the price to maximise $\pi_i = (T_i - \underline{t}_i)p_i$. If arbitrage cannot be avoided the monopolist sells a quality u_u at a price *p* in every market. If the marginal type $\underline{t} = p/u_u$ is lower than the lowest possible *T* (= 1 – *x*) then all the markets will be served, otherwise the monopolist will serve only those markets with *T* comprised between $\underline{t}$ and 1 + *x*.

PROPOSITION 1 (PURE MONOPOLY, EX POST ANALYSIS): *In the second stage of the game, taking quality as fixed, aggregate consumer surplus will always be higher under international exhaustion (arbitrage), profits will always be lower and aggregate welfare be higher as long as the dispersion of consumer valuations between markets is not too great. There exists a value x* of the dispersion of consumer valuations such that if x > x* then welfare will fall under international exhaustion.*

Proof: See appendix.

Proposition 1 simply gives a benchmark result for our model showing that if quality is taken as exogenous, our model of vertical product differentiation matches up with Malueg and Schwartz. If arbitrage does not induce any market to close it is always welfare enhancing as it improves the allocation of goods among customers, otherwise it depends on the degree of dispersion. When dispersion is high enough (*x* > *x**) the monopolist gives up too many markets, and this effect prevails over the better allocation among served customers, so that the welfare impact of parallel trade is negative.

3. Endogenous quality under monopoly

The stage 2 game is incomplete: in practice firms choose to invest on the basis of expected returns and the amount they invest in seeking and acquiring an intellectual property right may depend on whether they expect international exhaustion to apply. This issue is not addressed in the Malueg and Schwartz paper, but is clearly an important issue for research intensive industries such as pharmaceuticals. By solving for the optimal choice of quality in the stage 1 game, contingent on the expected regime at stage 2, we derive the following result.

PROPOSITION 2 (PURE MONOPOLY, EX ANTE ANALYSIS): *At the first stage of the game the selected quality of the product will always be lower under international exhaustion. With a quadratic cost function, consumer surplus will always be higher but profits and aggregate welfare will always be lower.*

Proof: See appendix.

The main finding is quite intuitive: for a given level of quality, the manufacturer gets higher profits without parallel trade than with parallel trade. Thus the impact of arbitrage is to dilute the incentives of the monopolist to invest in product innovation (quality). As quality also affects consumer surplus, international arbitrage negatively affects buyers as well by delivering an inferior product. The result $u_d > u_u$ is quite general. By applying monotone comparative static methods, it would hold regardless of the cost function given our demand specification. The fact that quality is always lower with parallel trade has a negative impact on consumer surplus *ex ante*. However we know from Proposition 1 that parallel trade has a positive effect on consumer surplus *ex post*.[8] Proposition 2 then says that, for overall welfare, the *ex ante* loss deriving from a lower quality dominates over the *ex post* consumer gains, and welfare is always lower under a uniform pricing regime, regardless of the degree of demand dispersion. This second result depends on the quadratic cost function that we have used. More generally, the size of the welfare loss depends on the convexity of the cost function: if it is very convex - so that quality changes are small in the two regimes - then the *ex ante* analysis would resemble the *ex post* analysis as stated in Proposition 1. On the other hand, if marginal cost is constant or increases at a decreasing rate, then quality differences become magnified and we can construct examples where even consumer surplus is lower with arbitrage.

The negative *ex ante* impact of parallel trade has significant policy implications. In our model a tightening of the regime against parallel trade raises profitability at the margin and therefore investment. Normally we would expect that marginal return to investment is greatest when there is no parallel trade, but counterexamples are possible. Suppose, for example, that there are only two countries. In each country there is a unit demand for the firm's product with willingness-to-pay of two units in country 1 regardless of whether investment takes place or not, while in country 2 the willingness to pay is one if there is no investment and 1.5 if investment occurs. If there is no parallel trade then the profit without the investment is 3 units and with the investment is 3.5 units, thus the change in profit is $\Delta\Pi = 0.5$. On the other hand, if there is parallel trade profits are 2,[9] while after investment the uniform prices is 1.5, both countries are served and profits are 3, so the gain is $\Delta\Pi = 1$. Hence parallel trade increases the incentive to innovate because the investment allows the seller to increase the price in both countries, while without parallel trade it would have affected the price only in country 2.

In summary, our model gives reasonable results on the trade-off between *ex ante* and *ex post* welfare properties of parallel trade. We now exploit its simple structure to investigate how welfare may be affected in the presence of competition and when the incumbent can vary its portfolio of qualities.

4. Competition with (non-strategic) generics

Sellers obtain patents or trademarks in order to protect the market power conferred by the originality or distinctiveness of their product. However, substitutes often exist in the form of less attractive alternatives - "own label" brands in supermarkets or generic treatments in pharmaceuticals[10]. Competition from generics of this kind places an additional constraint on the behaviour of the monopolist.

Suppose the generic is modelled as a product of quality u_g sold at a price p_g. To ensure that the generic product is a competitor 'from below', these are treated as parameters (hence the supplier of the generic is non-strategic).[11] In each market *i*, the marginal type is now determined by the presence of the generic $\underline{t} = p_g / u_g$. Without arbitrage the incumbent good

[8] We are referring to aggregate surplus - customers in some country will always be losers, most notably when they are not served under arbitrage when *x* is high enough.
[9] This profit level can be achieved either by selling at a uniform price of 1 in both countries, or by setting a price of 2 and selling only in one country.
[10] A generic is typically a product that was once protected by a patent which has now expired and can therefore be produced by anyone.
[11] As anybody can produce a product not protected by a patent, one can imagine that the production technology becomes a common knowledge and free entry ensures to drive the price of generic products (with identical quality) down to the marginal cost of production.

(e.g. protected by a patent or a trademark) is sold in each market i to customer types comprised between the highest type T_i and the indifferent type defined as:

$$\tilde{t}_i = \frac{p_i - p_g}{u_d - u_g}.$$

In the absence of international exhaustion the firm maximises its profit $\pi_i = (T_i - \tilde{t}_i)p_i$ with respect to p_i in each market.

If, by contrast, the monopolist cannot discriminate among the various markets (international exhaustion applies) it sells at a common price p its good of quality u_u. If a market is served by both the protected and the generic product, then the indifferent type in each market is:

$$\tilde{t} = \frac{p - p_g}{u_u - u_g}.$$

If this indifferent type is lower than the lowest possible T (= 1 – x) then all the markets will be served, otherwise the monopolist will serve only those markets with T comprised between $\tilde{t}$ and 1 + x, while the remaining markets will be served only by the generic. Hence the firm sets p to maximise $\Pi_u = \int_{\max[1-x,\tilde{t}]}^{1+x} \left(\frac{T - \tilde{t}}{2x}\right) p \mathrm{d}T$.

LEMMA 1: *When there is no international exhaustion the presence of the generic drives down prices by more in the richer markets. Under either regime (i.e. with or without arbitrage), when a generic is present then (a) prices chosen by the monopolist are lower (b) consumption of the monopolist's good is higher (c) the profit of the monopolist is lower (d) consumer surplus is higher and (e) welfare is higher than the case where there is no generic.*

Proof: see appendix.

As we say above, the generic represents a form of competition for the monopolist. The generic provides a reservation utility different from zero to every consumer. This reservation utility increases with one's type. Without international arbitrage, the incumbent responds market-by-market to the presence of the generic. As richer markets are made of higher types, this implies that the incumbent responds more aggressively in such markets. This explains the first part of the lemma. The other parts of the lemma simply say that the generic exerts a beneficial effect for consumers in every market under both regimes. In fact the competitive effect is so strong that the monopolist is forced to push its price down and sales of the incumbent good increase with the generic. The monopolist retains those customers who bought in the absence of a generic, but now sells to them at a lower price, while previously unserved customers either buy the incumbent's product or the generic. By revealed preference, customers are better off. As the market is expanded, this is also positive for welfare. We now examine how the welfare effects of parallel trade are affected by the presence of the generic in the stage 2 game where quality is taken as given.

PROPOSITION 3 (COMPETITION FROM GENERIC, EX POST ANALYSIS): *In the second stage of the game, taking quality as fixed, profits fall and consumer surplus increases under international exhaustion. Welfare increases if the dispersion is lower than a fixed threshold* x^{**}*, and decreases otherwise. This threshold value is higher than in the case where there is no generic,* $x^{**} > x^*$*, and the impact of the generic is stronger when the generic is not priced too competitively or the difference in quality is not too great.*

Proof: see appendix.

If dispersion is 'low', the welfare analysis mimics the one conducted in Proposition 1 under pure monopoly: both consumer surplus and welfare increase with parallel trade. The role of the generic is to increase sales, but its effect is basically the same with and without parallel trade: parallel trade improves welfare since it better allocates goods among consumers. If dispersion is 'high', the welfare properties still resemble those of pure monopoly: welfare decreases under parallel trade only if the dispersion is sufficiently high since too many markets would be shut down. However, the positive effect on consumer surplus and welfare due to parallel trade is now extended to a wider range of demand dispersion ($x^{**} > x^*$).

This result on the extended range of x is not obvious. The generic is present under either regime, and it has a positive impact on welfare in each regime (Lemma 1). Thus it is not straightforward that this impact is larger under parallel trade. Proposition 3 finds that the generic matters most in the 'partial' coverage regime. To understand why, we consider how a generic product changes the demand curve that the monopolist faces in each market, and thus how the monopolist's incentive to engage in price discrimination is affected by the generic.

In the absence of a generic product, the demand curve in market *i* where the highest type is T_i is $q_i = T_i - \underline{t}_i = T_i - p_i / u$. This gives rise to the following inverse demand curve in market *i*:

$$p_i = T_i u - q_i u \,.$$

Now consider the presence of the generic product. The monopolist's demand in market *i* is $q_i = T_i - \tilde{t}_i = T_i - (p_i - p_g)/(u - u_g)$ and the associated inverse demand is:

$$p_i = T_i(u - u_g) + p_g - q_i(u - u_g)\,.$$

The presence of the generic decreases both the vertical intercept (it must be the case that $p_g < T_i u_g$ for the generic to be an "effective" competitor, otherwise no one would ever buy it) and the absolute value of the slope of the inverse demand curve in each market. Thus the presence of the generic reduces cross-market differences in consumers valuations perceived by the monopolist, diluting the monopolist's incentive for price discrimination in the absence of parallel trade.

This leads to two sets of results. First, when the degree of demand dispersion is low and the monopolist would supply every market under both regimes, the overall welfare gains to be realised through parallel trade are decreased with the generic since the monopolist would not discriminate much in any case in the absence of parallel trade. However, as this happens when all markets are served, parallel trade still has good welfare properties since it improves the allocation of goods from low- to high-types markets.

Second, the generic has an impact on the monopolist's willingness to give up lower-type markets, which determines the range of dispersion under which parallel trade can improve welfare. One way of looking at the monopolist's incentive to supply everybody is to compare the profit from selling the first unit in the richest and in the poorest markets, e.g. by comparing the ratio of the vertical intercepts. This ratio is $\frac{\max\{T_i u\}}{\min\{T_i u\}} = \frac{1+x}{1-x}$ without the generic. With the generic, it becomes $\frac{\max\{T_i(u-u_g)+p_g\}}{\min\{T_i(u-u_g)+p_g\}} = \frac{(1+x)(u-u_g)+p_g}{(1-x)(u-u_g)+p_g} < \frac{1+x}{1-x}$. In other words, for a given *x*, the degree of "perceived" dispersion by the monopolist is decreased by the presence of the generic good. The generic changes demand in a way that the monopolist gives more weight to consumers with lower valuations in its profit maximisation, which in turn makes it less willing to give up lower-type markets: the less willing the monopolist to give up markets, the wider the range of demand dispersion that has welfare improvement under parallel trade. The range of dispersion for which parallel trade increases welfare is shifted to some value x^{**} that lies to the right of the value x^* found under pure monopoly.

This result on the threshold value at which the welfare effect of parallel trade goes from positive to negative is of negligible importance if either the quality differences are big or the generic is priced very competitively. In fact, if quality differences are big then the generic is not a very effective threat and the monopolist is almost uncontested. Similarly, if the generic is priced very competitively, the monopolist gives up competing against it and what matters is only the quality differential compared to the generic good. In both cases the threshold value is virtually unchanged compared to the case where there is no generic. By contrast, if quality differences are moderate and the generic is priced not 'too' competitively, then the good properties of parallel trade are extended to higher values of demand dispersion in the 'partial' coverage regime by the presence of the generic good. In general, the presence of the generic tends to reinforce the good properties of international arbitrage: it extends the range of demand dispersion compatible with a 'full' coverage regime (i.e. the regime when arbitrage has an unambiguous positive impact on welfare) and contrasts the negative impact for high dispersion values in the 'partial' coverage regime.

We now consider the impact of the generic on welfare effects of parallel trade in the stage 1 game.

PROPOSITION 4 (COMPETITION FROM GENERIC, EX ANTE ANALYSIS): *In the presence of a generic, at the first stage of the game the selected quality of the product will* *always be lower under international exhaustion. The reduction in quality due to arbitrage is increased by the presence of a generic. With a quadratic cost function, aggregate welfare will always be lower and consumer surplus will be higher as long as dispersion is not too great.*

Proof: see appendix.

The main finding is that, for any cost function, quality is reduced by international arbitrage even in the presence of a generic product. In fact, the impact is more negative than under pure monopoly and the quality reduction due to parallel trade is magnified.

When read together, Propositions 3 and 4 describe a reasonable picture of the impact of the generic on parallel trade. Proposition 3 obtains the non-obvious result that parallel trade improves its allocative properties *ex post* in the presence of the generic as this extends the range of parameters of demand dispersion such that parallel trade is beneficial. However this positive result comes at a cost: the better *ex post* properties must be traded off against worse *ex ante* properties (Proposition 4). On top of the quality reduction due to parallel trade *per se*, the impact of the generic is to reduce quality further. The last result depends on one embedded assumption: the monopolist is restrained to offer only one product.

5. The monopolist offers more than one product

Under what circumstances will the producer be willing to introduce a second product, of lower quality? A second product might be introduced for two main reasons. First, in the presence of arbitrage, a lower-quality variant may become a tool for the monopolist to increase his ability to discriminate between markets, given that price discrimination would be undone by parallel trade. Second, it may be intended to compete with the generic, where the quality of the monopolist's leading brand has already been fixed. We begin by showing that it is *never* optimal in our model to introduce a low quality variant for the purposes of discriminating between markets when there is no competition form a generic. We then show that a "fighting brand" is an efficient response to competition from a generic product.

5.1 No competition from a generic

To model the introduction of a second product we assume that once a certain quality u has been produced at the investment stage, the producer is also able to obtain a lower-quality product of quality u_l at a zero extra cost, as a by-product of the first innovation, $u_l \leq u$. This assumption is clearly extreme but it allows us to focus only on the discrimination and competition issues, forgetting about problems associated with costs. This might also be a plausible assumption for the case of a pharmaceutical product where R&D is driven by the search for novel treatments and considerations of product positioning are secondary at the innovative stage.

LEMMA 2: *Whether there is or is not international exhaustion the unconstrained monopolist is better off by never selling the low-quality product, as it can ensure at least the same profit by selling only the high-quality product.*

We do not provide a formal proof of this result here since it is a manifestation of the so-called "no haggling" result (Stokey, 1979). The intuition comes from the fact that preferences are linear in quality and quality is costless to produce. Hence the monopolist's objective function is also linear and there is always a corner solution where the monopolist supplies only one product.[12] Although by introducing a second product of lower quality can enlarge the market, the availability of such product also induces some customers that were previously buying the high-quality product to purchase the low-quality product at a cheaper price. The reduced revenues from such 'switching' customers always prevail over the increased revenues from the expanded bottom end of the market, hence the monopolist should never downgrade its own quality.

If quality also involved a variable cost of production, then this result would not hold for a sufficiently convex cost function since in this case we know from Mussa and Rosen (1978) that it would be optimal to induce self-selection among heterogeneous customers by offering a menu of qualities. We notice two aspects here. First, in our context it makes sense to assume that variable costs do not play a big role *after* R&D expenses have been sunk. In fact, Lemma 2 would still hold for constant marginal cost of production since the monopolist's problem would still be linear in quality. Second, our main point is that arbitrage and product variety

[12] Formal proofs are available from the authors. Stokey's original problem refers to inter-temporal price discrimination. She finds that, although price discrimination is feasible, it is never advantageous to charge declining prices over time and it is always optimal to pre-commit to a fixed price. Her problem is linear since the undiscounted cost of producing/delivering a good is assumed to be constant. As shown by Salant (1989), this problem is isomorphic to second-degree price discrimination where the "probability of sale" corresponds to

are not related. Arbitrage does not cause *additional* products to be introduced: if the monopolist is offering a single product without arbitrage, then it would still offer a single product in the presence of arbitrage. This result does not depend on the uniform distribution of willingness-to-pay in the various countries.[13]

5.2 Competition from a generic

The nature of the problem changes quite dramatically in the presence of the generic. The first important result is that the monopolist now wants to introduce the lower-quality variant.

LEMMA 3: *In the presence of the generic the incumbent always introduces a second, low quality product. This raises both its profits and consumer surplus. Sales of the high-quality good are the same one as without competition from the generic. This holds both with and without international exhaustion.*

Proof: see appendix.

The lower-quality variant is introduced for strategic reasons rather than to discriminate more efficiently. As we showed in Lemma 1, the presence of the generic introduces a type-dependent reservation utility that obliges the monopolist to expand its total output. With only one product on sale, all its customers would have to be offered the same deal. However, the monopolist can now price the low-quality variant in a way to contrast competition from below (a 'fighting brand'; see Johnson and Myatt, 2003) while protecting the higher margins on the high-quality good. That the monopolist is now making higher profits compared to the case of competition against a generic using a single brand is not surprising as the firm has two goods at its disposal and decides to supply both. But consumers are also better off: those who still buy the high-quality good or the generic are paying the same price as before, so their welfare is unchanged. Those who decide to switch instead to the lower-quality variant must be better off by revealed preferences. Overall consumer surplus increases.

"quality" discrimination in a model *à la* Mussa and Rosen (1978). Che and Gale (2000) discuss how the "no haggling" result also depends on binding budget constraints.

[13] To see this, imagine the extreme case where a mass m of customers in concentrated in a "high" country where everybody is of type t_h, and a mass $(1 - m)$ is concentrated in a "low" country where everybody is of type $t_l < t_h$. If the monopolist sells only the high quality good, his profit is $\pi^1 = \max[t_l u, m t_h u]$, according to whether both markets are served or only the "high" country. On the other hand, if it offers two products, by solving the

The other interesting implication for our purposes is that the patent holder supplies the high-quality good to the same number of customers as under pure monopoly. This holds under both pricing regimes. In terms of *marginal* incentives to conduct R&D the analysis is now identical to the one presented in Section 3 when there was no competition from a generic. By comparing the allocations with and without parallel trade we obtain our final result.

PROPOSITION 5 (MULTI-PRODUCT INCUMBENT, COMPETITION FROM GENERIC, EX POST AND EX ANTE ANALYSIS): *In the second stage of the game, taking quality as fixed, welfare increases under international exhaustion if the dispersion is lower than a fixed threshold x^{***}, and decreases otherwise. This threshold value is higher than in the case where there is no generic, $x^{***} > x^{*}$. At the first stage of the game, the selected quality of the product is the same as in the monopoly case, with and without parallel trade. Thus quality is always lower under international exhaustion but this reduction is not affected by the presence of a generic. With a quadratic cost function, ex ante aggregate welfare is always lower and consumer surplus is higher as long as dispersion is not too great.*

Proof: see appendix.

Proposition 5 should be contrasted with Propositions 3 and 4. The *ex post* welfare effects still resemble those of the 1-product case: the presence of the generic product extends the good welfare properties of parallel trade to a wider range of demand dispersion. The reasons are the same ones as those discussed in the 1-product case. Unlike the 1-product case, however, there is no trade-off between improved *ex post* welfare properties and worse *ex ante* properties. Once the option to introduce a fighting brand is allowed, there is no further incentive to reduce investment in quality caused by the presence of the generic on top of the typical quality reduction implied by parallel trade under pure monopoly. The incumbent reacts to the presence of competition by adjusting its portfolio of varieties and is able to 'protect' marginal revenues from the high-quality product. The quality chosen is exactly the same as under pure monopoly since marginal revenues in the different regimes are the same. Thus we obtain a "separation" result on the impact of parallel trade in the presence of competition from generics and endogenous quality choices. The sign and magnitude of the reduction in *ex ante* investments is the same one that occurs under pure monopoly (Section 3), while *ex post* the

---

standard screening problem that satisfies incentive compatibility for the high types and individual rationality for incumbent introduces a lower-quality variant. The impact of the generic is to extend the positive welfare properties of parallel trade to a wider range of demand dispersion, with no impact on *ex ante* allocations.

6. Discussion and conclusions

It is clear that a monopolist, such as the owner of a valuable patent or trademark, will choose to engage in international price discrimination. In this paper we have considered whether social welfare would increase if the monopolist were prevented from doing so. One way to bring this about would be to permit arbitrage. We have extended the analysis of Malueg and Schwartz to include not just the price setting decision of the monopolist but also the initial decision to invest in developing a product of a certain quality. This part of the story is critical when we are discussing the supply of goods protected by patents and trademarks - the very reason for granting such intellectual property rights is to encourage investment.

We have considered two constraints on monopoly power: the inability to prevent arbitrage (international exhaustion) and the presence of a generic, and in both cases we have found that the incentive of the monopolist to supply quality is diminished. However, when we considered these two constraints simultaneously we have found that the incentive to supply quality is not reduced relative to benchmark monopoly case as long as the monopolist can offer an inferior version on its own product (a fighting brand). Our results are summarised in Table 1. We have analysed both *ex post* welfare (i.e. taking into account only price effects) and *ex ante* welfare (taking into account investments) in four cases. Case A is the benchmark case analysed in Malueg and Schwartz (*ex post*), case B permits competition from a generic when the incumbent is constrained to sell a single product. Case C concerns the situation where there is no competition from the generic but the monopolist can offer two products and case D considers competition from a generic when the incumbent offers two products.

Case A makes the basic but important point that once investment in quality is considered then international exhaustion can reduce welfare. Case B shows an additional trade-off: the presence of the generic both extends the range of demand dispersion with positive *ex post* welfare properties of international exhaustion and exacerbates the negative *ex ante* impact on investment. Case C shows that in the absence of competition from a generic product the

---

the low types, the equilibrium profit is: $\pi^2 = t_l u_l + mt_h(u - u_l) < \pi^1$.

opportunity to introduce a second quality of product does not produce a different result from case A. This is because it is a dominant strategy for the monopolist to offer a single product even in the presence of arbitrage. Case D is perhaps the most interesting from a theoretical perspective. When a generic is available the monopolist will choose to offer more than one product, because the second product will act as a fighting brand against the generic. The presence of the generic extends the range of demand dispersion with positive *ex post* welfare properties of international exhaustion (as in case B), but will not lead to a further reduction of investment in quality (contrary to case B) because of the presence of the fighting brand.

Table 1: Summary of effects of parallel trade

	No competition, 1 product	Comp. from generic, 1 product	No competition, 2 products	Comp. from generic, 2 products
Case	A	B	C	D
Ex post	Malueg and Schwartz	Better properties than A	As with 1 product. Case C is identical to A	Better properties than A
Ex ante	Lower quality. Welfare worsens	Worse properties than A: even lower quality		Same as A: lower quality. Welfare worsens

These results also have significant policy implications. The desirability of international exhaustion is a subject of great controversy. On the face of it parallel trade brings significant benefits to consumers in high valuation markets, but our analysis shows that any *ex post* consumer benefit must be traded off against the reduction in *ex ante* incentives to invest. However, this trade-off is diluted when the monopolist is able to offer more than one product variety and generics are present in the market. The nature of the generic in this example is therefore crucial. When dealing with goods protected by trademark generics are quite common, for example unbranded jeans, or coffee, or perfumes with lesser brand names. Moreover, in such cases it is not unusual to see the trademark owner offer branded alternatives that differ significantly in quality. However, in the case of goods protected by patents then during the life of the patent generics that copy the protected good are in fact illegal. When the patent expires the owner attempts to differentiate its brand from generic copies, but in fact the expiry of the patent means that generic copies can exactly mimic the original product. Thus the possibility of a reasonable generic substitute for a good protected by patent is relatively small, unlike the case of branded goods. This suggests that policy on international exhaustion might optimally differ according to the type of intellectual property right considered.

Another way that one might think about the generic product is to think of it as a constraint on the breadth of the patent granted to a monopolist. When the generic is close in quality terms to the monopolist's product it is as if the scope of the patent were quite narrow, and when the generic is of significantly inferior quality it is as if the scope of the patent were quite broad. Following this analogy, our analysis suggests that international exhaustion is more likely to increase consumer welfare where patents are narrowly defined as long as the patent owner introduces more than one version of the product.

It is easy to extend our model to multiple periods, an initial phase of "pure" monopoly followed by periods when inferior generics are allowed to develop. Our results from Section 5 then say that the incumbent offers a single product during the monopoly phase, while it reacts to entry by offering two variants with the purpose of competing against generic rivals. This theoretical result is supported by empirical evidence. Hollis (2002) reports that brand-name pharmaceutical firms routinely introduce "pseudo-generics" in Canada. Pseudo-generics are the innovator drug company's generic version of its own brand name drug. They are generally manufactured by the innovator company and distributed through a third party. He also observes that in countries where this practice takes place, pseudo-generics are invariably a response to the entry of true generics when a patent expires.[14] This is indeed a case of fighting brands as explained by our model. Our results also imply that the incentives to innovate are only affected by the regime of international exhaustion and not by the entry of the generics.

The issues discussed here explore the underlying tension that exists between intellectual property rights and the principles of free competition. The paper makes a contribution by extending the analysis to the case where the choice of product quality is endogenous. This is an issue that deserves more attention by economists, not merely because of its theoretical complexities, but because of the significant practical implications for welfare.

[14] This practice was also common in the United States during the early and mid-90s, until it was halted following an investigation of the Federal Trade Commission.

References

Abbott, F.M., 1998, "First Report (Final) to the Committee on International Trade Law of the International Law Association on the Subject of Parallel Importation," *Journal of International Economic Law*, 1 (4), 607-636.

Bae, S.H. and J.P. Choi, 2002, "A Model of Piracy," *mimeo*, Michigan State University.

Che, Y.-K. and I. Gale, 2000, "The Optimal Mechanism for Selling to a Budget-Contrained Buyer," *Journal of Economic Theory*, 92, 198-233.

Danzon, P.M. and A. Towse, 2003, "Differential pricing for pharmaceuticals: reconciling access, R&D and patents," *International Journal of Health Care Finance and Economics*, 3 (3), 183-205.

Hausman, J.A. and J.K. MacKie-Mason, 1988, "Price discrimination and patent policy," *RAND Journal of Economics*, 19 (2), 253-265.

Hollis, A., 2002, "The importance of being first: evidence from Canadian generic pharmaceuticals," Health Economics, 11, 723-734.

Johnson, J.P. and D.P. Myatt, 2003, "Multiproduct Quality Competition: Fighting Brands and Product Line Pruning," *American Economic Review*, 93 (3), 748-774.

Li, C. and K. Maskus, 2004, "Parallel Imports and Cost Reducing Research and Development," *mimeo*, University of Colorado.

Maskus, K. and Y. Chen, 2004, "Vertical Price Control and Parallel Imports: Theory and Evidence," *Review of International Economics*, forthcoming.

Malueg, D. and M. Schwartz, 1994, "Parallel Imports, Demand Dispersion and International Price Discrimination," *Journal of International Economics*, 37 (3-4), 167-195.

Mussa, M. and S. Rosen, 1978, "Monopoly and Product Quality," *Journal of Economic Theory*, 18, 301-17.

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.

New Zealand Institute of Economic Research, 1999, "Parallel Importing: A Theoretical and Empirical Investigation", Report to Ministry of Commerce.

OECD, 2002, "Synthesis Report on Parallel Imports," Paris.

Richardson, M., 2002, "An elementary proposition concerning parallel imports," *Journal of International Economics*, 56, 233-245.

Rey, P., 2003, "The Impact of Parallel Imports on Prescription Medicines," *mimeo*, IDEI, University of Toulouse.

Salant, S., 1989, “When is Inducing Self-Selection Suboptimal for a Monopolist?” *Quarterly Journal of Economics*, 104, 391-397.

Stokey, N., 1979, “Intertemporal price Discrimination,” *Quarterly Journal of Economics*, 93, 355-371.

Swedish Competition Authority, 1999, *Parallel Imports - Effects of the Silhouette Ruling*, Swedish Competition Authority, Report Series 1999:1.

Tarr, D.G., 1985, “An Economic Analysis of Grey Market Imports,” *mimeo*, U.S. Federal Trade Commission, Washington, D.C.

York Health Economics Consortium, 2003, “Benefits to Payers and Patients from Parallel Trade,” University of York.

Varian, H., 1985, “Price Discrimination and Social Welfare,” *American Economic Review*, 75, 870-75.

Appendix

Proof of proposition 1

When price discrimination is possible, then the optimal price and the associated marginal type and profit in each market are respectively:

$$p_i = T_i u_d / 2$$
$$\underline{t}_i = T_i / 2$$
$$\pi_i = T_i^2 u_d / 4$$

From this we can derive expressions for aggregate profit and consumer surplus:

$$(1) \quad \begin{cases} \Pi_d = \int_{1-x}^{1+x} \frac{\pi_i}{2x} \mathrm{d}T = u_d (3 + x^2)/12 \\ CS_d = \int_{1-x}^{1+x} \left(\frac{\int_{\underline{t}_i}^{T} (tu_d - p_i)\mathrm{d}t}{2x} \right) \mathrm{d}T = u_d (3 + x^2)/24 \end{cases}$$

When (costless) arbitrage is possible the monopolist sets a single price p to maximise:

$$\Pi_u = \int_{\max[1-x, p/u_u]}^{1+x} \frac{T - p/u_u}{2x} p \mathrm{d}T \ .$$

This gives rise to the following solutions under price uniformity:

$$(2) \text{ If } x \leq 1/2 \begin{cases} p = u_u / 2 \\ \underline{t} = 1/2 \\ \Pi_u = u_u / 4 \\ CS_u = \int_{1-x}^{1+x} \int_{\underline{t}}^{T} \frac{tu_u - p}{2x} \mathrm{d}t\mathrm{d}T = \frac{u_u (3 + 4x^2)}{24} \end{cases} \qquad \text{If } x > 1/2 \begin{cases} p = u_u (1+x)/3 \\ \underline{t} = (1+x)/3 \\ \Pi_u = u_u (1+x)^3 / (27x) \\ CS_u = \int_{\underline{t}}^{1+x} \int_{\underline{t}}^{T} \frac{tu_u - p}{2x} \mathrm{d}t\mathrm{d}T = \frac{2u_u (1+x)^3}{81x} \end{cases}$$

Comparing (1) and (2), when quality is not affected by arbitrage ($u_d = u_u = u$), we get:

$$(3) \qquad \text{If } x \leq \frac{1}{2} \begin{cases} \Delta\Pi = \Pi_u - \Pi_d = -ux^2 / 12 < 0 \\ \Delta CS = ux^2 / 8 > 0 \\ \Delta Welfare = ux^2 / 24 > 0 \end{cases}$$

$$(4) \qquad \text{If } x > \frac{1}{2} \begin{cases} \Delta\Pi = -u[x^2 / 12 - (2x-1)^2 (4+x)/(108x)] < 0 \\ \Delta CS = u[x^2 / 8 - (2x-1)(16 - x + 46x^2)/(648x)] > 0 \\ \Delta Welfare = u[x^2 / 24 - (2x-1)(40 - 43x + 34x^2)/(648x)] > 0 \text{ if } x < x^* \approx 0.625 \end{cases}$$

The following table summarises the impact of arbitrage *ex post*:

Number of markets served	Same if $x \leq$ ½, smaller if $x >$ ½
Consumer surplus	Higher
Profits	Lower
Welfare	Higher if $x < x^*$, lower if $x > x^*$

Proof of proposition 2
Using the quadratic cost function we get the following equilibrium values for the quality:

$$(5) \qquad u_d = \frac{3+x^2}{12k}, \; u_u = \begin{cases} 1/(4k) \text{ if } x \leq 1/2 \\ \dfrac{(1+x)^3}{27kx} \text{ if } x > 1/2 \end{cases}$$

Plugging these values back into eq. (1) and (2) and taking their difference we can derive the following *ex ante* implications of allowing international arbitrage with endogenous quality:

$$\text{If } x \leq \frac{1}{2} \begin{cases} \Delta\Pi = -x^2(6+x^2)/(288k) < 0 \\ \Delta CS = x^2(6-x^2)/(288k) > 0 \\ \Delta Welfare = -x^4/(144k) < 0 \end{cases}$$

$$\text{If } x > \frac{1}{2} \begin{cases} \Delta\Pi = \dfrac{(1+x)^6}{1458kx^2} - \dfrac{(3+x^2)^2}{288k} < 0 \\ \Delta CS = \dfrac{2(1+x)^6}{2187kx^2} - \dfrac{(3+x^2)^2}{288k} > 0 \\ \Delta Welfare = \dfrac{7(1+x)^6}{4374kx^2} - \dfrac{(3+x^2)^2}{144k} < 0 \end{cases}$$

Number of markets served	Same if $x \leq$ ½, smaller if $x >$ ½
Quality	Lower
Consumer surplus	Higher
Profits	Lower
Welfare	Lower

Proof of lemma 1
(i) No international exhaustion (no arbitrage)
We can write equilibrium prices, the marginal type and profits as:

$$p_i = [T_i(u_d - u_g) + p_g]/2$$
$$\tilde{t}_i = T_i/2 - p_g/[2(u_d - u_g)]$$
$$\pi_i = [T_i(u_d - u_g) + p_g]^2/[4(u_d - u_g)].$$

The impact on price due to the presence of the generic good is bigger in the richer markets ($\partial p_i / \partial u_g = -T_i/2$) - this is the first part of the lemma. To ensure that the generic is 'effective' (i.e. it has a positive market share in every market) we need the following condition, which we assume to hold:

(A) $p_g < \frac{(u_d - u_g)u_g(1-x)}{2u_d - u_g}$.

Assumption (A) plays only a technical role as it limits the cases to be considered (without it, if x tends to 1 the generic would not be sold in some markets). It will be convenient at times to express this inequality as a limit on the degree of dispersion:

$$x \leq x_{\max} = 1 - p_g \frac{2u_d - u_g}{u_g(u_d - u_g)} < 1.$$

Aggregate profits for the patent holder, for the generic and consumer surplus are:

$$(6)\ \begin{cases} \Pi_d = \int_{1-x}^{1+x} \frac{\pi_i}{2x} dT = \frac{(p_g + u_d - u_g)^2}{4(u_d - u_g)} + \frac{(u_d - u_g)x^2}{12} \\ \Pi_{g,d} = \int_{1-x}^{1+x} \frac{(\tilde{t}_i - p_g / u_g) p_g}{2x} \mathrm{d}T = \frac{p_g}{2}[1 - \frac{p_g(2u_d - u_g)}{u_g(u_d - u_g)}] \\ CS_d = \int_{1-x}^{1+x} (\frac{\int_{\tilde{t}_i}^{T}(tu_d - p_i)\mathrm{d}t + \int_{\underline{t}}^{\tilde{t}_i}(tu_g - p_g)\mathrm{d}t}{2x}) \mathrm{d}T = \frac{u_d(3 + x^2)}{24} + \frac{3(u_g - p_g)^2}{8u_g} + \frac{p_g^2 u_d}{8u_g(u_d - u_g)} + \frac{x^2 u_g}{8} \end{cases}$$

Compared to the case without generic competition (eq. (1)), the monopolist is forced to lower the price and its profits go down. The market is also expanded: more customers buy from the incumbent and some new customers buy the generic. Hence, for a given level of quality, consumer surplus unambiguously goes up with the presence of the generic. This also implies that welfare must go up with the presence of the generic as market coverage gets closer to the efficient allocation (the first best occurs when everybody consumes the high-quality good).

(ii) International exhaustion (arbitrage)

The monopoly price, the threshold type and the monopoly profit depend on the critical value of x, the degree of dispersion between markets:

$$(7)\quad \text{If } x \leq \hat{x} \begin{cases} p = (p_g + u_u - u_g)/2 \\ \tilde{t} = \frac{1}{2} - \frac{p_g}{2(u_u - u_g)} \\ \Pi_u = (p_g + u_u - u_g)^2 / [4(u_u - u_g)] \end{cases}$$

$$(8)\quad \text{If } x > \hat{x} \begin{cases} p = [p_g + (u_u - u_g)(1+x)]/3 \\ \tilde{t} = \frac{1+x}{3} - \frac{2p_g}{3(u_u - u_g)} \\ \Pi_u = [p_g + (u_u - u_g)(1+x)]^3 / [27(u_u - u_g)^2 x] \end{cases}$$

where $\hat{x} = \frac{1}{2} + \frac{p_g}{2(u_u - u_g)}$.

When there is exhaustion we observe again the market expansion effect on the incumbent product due to the presence of the generic product. Eq. (8) represents the solution when the monopolist drops some markets ('partial' coverage). It is easy to verify that condition (A) always ensures that $\min[T,\tilde{t}\,] > \underline{t}$, i.e. the generic product always commands a positive market share in both regimes. Eq. (7) is valid if $p_g < (u_u - u_g)u_g/(2u_u - u_g)$, while eq. (8) is valid as long as $p_g < u_g(1-x)$, which are both satisfied under the more stringent condition (A) (Condition (A) has to be re-written substituting u_u for u_d under international exhaustion). We can also write down the expressions for the profits of the supplier of the generic and for consumer surplus under uniform pricing:

$$\text{If } x \le \hat{x} \begin{cases} \Pi_{g,u} = \int_{1-x}^{1+x} \frac{(\tilde{t} - p_g/u_g)p_g}{2x}\mathrm{d}T = \frac{p_g}{2}[1 - \frac{p_g(2u_u - u_g)}{(u_u - u_g)u_g}] \\ CS_u = \int_{1-x}^{1+x} (\frac{\int_{\tilde{t}}^{T}(tu_u - p)\mathrm{d}t + \int_{\underline{t}}^{\tilde{t}}(tu_g - p_g)\mathrm{d}t}{2x})\mathrm{d}T = \frac{u_u(3+4x^2)}{24} + \frac{3(u_g - p_g)^2}{8u_g} + \frac{p_g^2 u_u}{8u_g(u_u - u_g)} \end{cases}$$

$$\text{If } x > \hat{x} \begin{cases} \Pi_{g,u} = \int_{\tilde{t}}^{1+x} \frac{(\tilde{t} - p_g/u_g)p_g}{2x}\mathrm{d}T + \int_{1-x}^{\tilde{t}} \frac{(T - p_g/u_g)p_g}{2x}\mathrm{d}T = \\ = \frac{p_g}{9}[\frac{7x - 1 - x^2}{x} - \frac{9p_g}{u_g} - \frac{2p_g(1+x)}{x(u_u - u_g)} - \frac{p_g^2}{x(u_u - u_g)^2}] \\ CS_u = \int_{\tilde{t}}^{1+x} (\frac{\int_{\tilde{t}}^{T}(tu_u - p)\mathrm{d}t + \int_{\underline{t}}^{\tilde{t}}(tu_g - p_g)\mathrm{d}t}{2x})\mathrm{d}T + \int_{1-x}^{\tilde{t}} (\frac{\int_{\underline{t}}^{T}(tu_g - p_g)\mathrm{d}t}{2x})\mathrm{d}T = \frac{2u_u(1+x)^3}{81x} + \\ u_g\frac{69x - 12x^2 + 23x^3 - 4}{162x} + p_g[\frac{2 - 23x + 2x^2}{27x} + \frac{2p_g(1+x)}{27x(u_u - u_g)} + \frac{p_g}{2u_g} + \frac{2p_g^2}{81x(u_u - u_g)^2}] \end{cases}$$

Compared to eq. (2) consumer surplus is positively affected by the presence of the generic. As in the case without arbitrage, efficiency is also unambiguously increased.

Proof of proposition 3

Fix a given quality level ($u_u = u_d = u$). Denote by d the quality differential between the patented good and the generic good ($d = u - u_g$). Welfare comparisons give:

$$(9) \qquad \text{If } x \le \hat{x} \begin{cases} \Delta\Pi = \Pi_u - \Pi_d = -dx^2/12 < 0 \\ \Delta\Pi_g = 0 \\ \Delta CS = dx^2/8 > 0 \\ \Delta Welfare = dx^2/24 > 0 \end{cases}$$

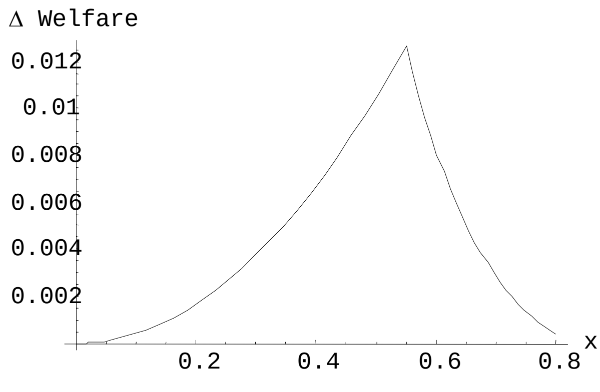
The previous expressions take the same form as those obtained in Proposition 1 without generic when $x \le 1/2$ (eq. (3)): their sign w.r.t. x varies in exactly the same way and they would coincide when $d = u$. Also notice how the 'full' coverage regime is extended to values of x greater than ½ and that $\partial\hat{x}/\partial p_g > 0$ and $\partial\hat{x}/\partial d < 0$. When dispersion is 'high' the signs of the profit differences are derived as follows:

$$\frac{\partial \Delta\Pi}{\partial x} = -\frac{2p_g^3 + 6p_g^2 d + 6p_g d^2(1-x^2) + d^3(2-6x^2+5x^3)}{54d^2x^2} < 0, \Delta\Pi\Big|_{x=\hat{x}} = \frac{-(p_g+d)^2}{48d} < 0$$

$$\frac{\partial \Delta\Pi_g}{\partial x} = \frac{p_g[d(1+x)+p_g][(d(1-x)+p_g]}{9d^2x^2} > 0, \Delta\Pi_g\Big|_{x=\hat{x}} = 0$$

$$(10)\ \text{If } x > \hat{x} \begin{cases} \Delta\Pi < 0; \Delta\Pi_g > 0 \\ \Delta CS = d[x^2/8 - (2x-1)(16 - x + 46^2)/(648x)] + \Omega(p_g, x, d) > 0 \\ \Omega(p_g, x, d) = \dfrac{p_g(8-11x+8x^2)}{108x} + \dfrac{p_g^2(16-11x)}{216dx} + \dfrac{2p_g^3}{81d^2x} \\ \Delta Welfare = d[x^2/24 - (2x-1)(40-43x+34x^2)/(648x)] + \Psi(p_g, x, d) \\ \Psi(p_g, x, d) = \dfrac{p_g(8-11x+8x^2)}{108x} + \dfrac{p_g^2(19x-8)}{216dx} - \dfrac{4p_g^3}{81d^2x} \end{cases}$$

If dispersion is 'high', the expressions for the differences in consumer surplus and in welfare are made of several terms. The first term in each expression takes the same form as the expression derived in the corresponding case without generics (eq. (4)) and have the same sign as x varies, while the other terms reflect the impact of the generic. If d is big or if $p_g \to 0$, then only the first terms matter, $\hat{x} \to 1/2$, and welfare analysis w.r.t. demand dispersion is identical to monopoly. In all other cases, it is possible to show that $\Omega > 0$ and $\Psi > 0$ when $\hat{x} < x < x_{\max}$. By denoting by x^{**} the root that makes $\Delta Welfare = 0$ in eq. (10), this means that $\Delta Welfare(x^*) > 0$ and that welfare decreases only in the 'partial' coverage regime if the dispersion of x is sufficiently high, with $x^{**} > x^*$ as long as this root falls in the range of admissible dispersion implied by Assumption (A). In fact, it is possible to construct examples where welfare increases for *any* degree of admissible dispersion: a case is depicted in the following figure ($u = 3$, $u_g = 2$, $p_g = 0.1$, implying $x_{\max} = 0.8$).



Number of markets served by the incumbent product	Same if $x \leq \hat{x}$, smaller if $x > \hat{x}$ ($\hat{x} > 1/2$)
Number of markets served by the generic product	Same
Consumer surplus	Higher
Profits (incumbent)	Lower
Welfare	Higher if $x < x^{**}$ ($x^{**} > x^*$)

Proof of proposition 4

The expressions for stage-2 gross profit are not linear in u. As a first step note that:

(11)
$$\begin{cases} \dfrac{\partial \Pi_d}{\partial u} = \dfrac{(3+x^2)}{12} - \dfrac{p_g^2}{4(u-u_g)^2} \\ \dfrac{\partial \Pi_u}{\partial u} = \begin{cases} \dfrac{1}{4} - \dfrac{p_g^2}{4(u-u_g)^2} \text{ if } x \le \hat{x} \\ \dfrac{(1+x)^3}{27x} - \dfrac{p_g^2[3(u-u_g)(1+x)+2p_g]}{27x(u-u_g)^3} \text{ if } x > \hat{x} \end{cases} \end{cases}$$

The presence of the generic reduces the incentives to invest in quality in all regimes. The level of marginal revenues is pushed down in both regimes. However our main interest is to look at their difference.[15] To answer the question whether the presence of arbitrage affects our earlier result that arbitrage reduces quality relative to the no arbitrage regime, it suffices to show that the difference between marginal revenues with and without arbitrage is negative:

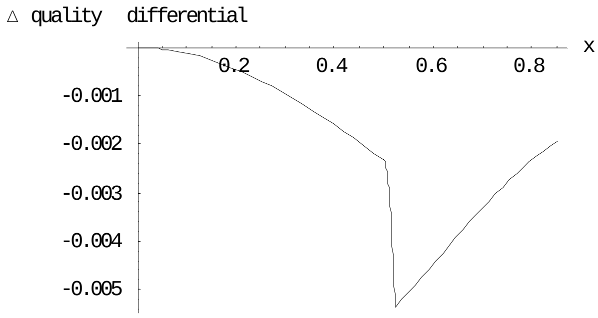
$$\frac{\partial(\Pi_u - \Pi_d)}{\partial u} = \begin{cases} -x^2/12 < 0 \text{ if } x \le \hat{x} \\ \dfrac{4-15x+12x^2-5x^3}{108x} - \dfrac{p_g^2[3(u-u_g)(4-5x)+8p_g]}{108x(u-u_g)^3} < 0 \text{ if } x > \hat{x} \end{cases}$$

Hence we have proven again that - also in the presence of a generic good - the incentive to produce quality goods is reduced with parallel trade, for any cost function.

Under full coverage the analysis is similar to the case without the generic. The difference in marginal incentives is unaffected by the presence of the generic. However, this does not imply that the reduction in quality is the same as under full monopoly. In fact, the difference between marginal revenues is unaffected by the generic only for the same value of u. Notice that, in eq. (11), the extra negative term compared to the absence of generic is bigger the smaller is u. As $u_u < u_d$, then there is an *extra* incentive to reduce quality due to the presence of the generic also in the full coverage case. In the case of partial coverage the first (negative) term is the same one that we derived in the case of pure monopoly. There is also a second term (negative as well: recall that the analysis makes sense as long as $x < x_{\max}$). Hence the decrease in quality due to parallel trade is more pronounced under this regime by the presence of the generic product.

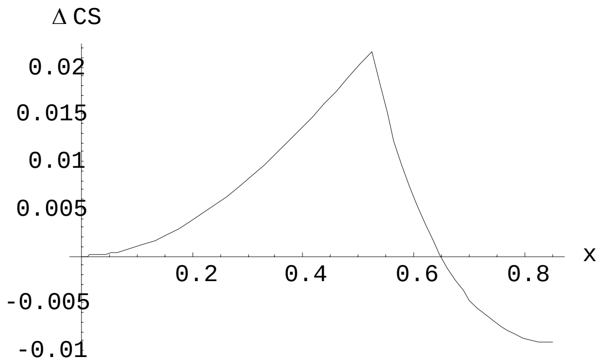
In the figure below we report the *difference in quality differential* with and without generic, i.e. we report $(u_u - u_d)_{generic} - (u_u - u_d)_{monopoly}$ for the case of quadratic costs (parameters: k = 0.06, $p_g = 0.1$, $u_g = 2$).

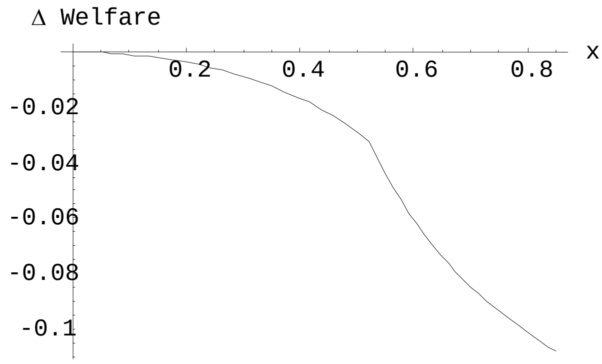
[15] As a side remark, it can be shown that $\frac{\partial^2 \Pi_d}{\partial u \partial p_g} < 0, \frac{\partial^2 \Pi_d}{\partial u \partial u_g} < 0, \frac{\partial^2 \Pi_u}{\partial u \partial p_g} < 0, \frac{\partial^2 \Pi_d}{\partial u \partial u_g} < 0$. In other words, under both exhaustion regimes, the way a monopolist reacts in terms of its own investment to a 'better' generic depends on whether 'better' means having a higher quality u_g (the monopolist's u decreases) or a lower price p_g (u increases). These two side-results are also found in Bae and Choi (2002).



A full welfare analysis depends on the convexity of the cost function. With the quadratic cost function it is possible to obtain closed-form solutions but they are not reported here because they take cumbersome expressions. We report drawings to show the main features. They show that the reduced quality worsens the *ex ante* welfare properties of parallel trade (parameters: $k = 0.06$, $p_g = 0.1$, $u_g = 2$).

Number of markets served by incumbent product	Same if $x \leq \hat{x}$, smaller if $x > \hat{x}$
Number of markets served by generic product	Same
Quality	Lower
Consumer surplus	See figure below
Profits (incumbent)	Lower
Welfare	See figure below





Proof of lemma 3

(i) No international exhaustion

Imagine that the monopolist offers two products of different quality. The higher-quality product u_d is sold at a price p_{di} and the lower-quality product $u_l \geq \underline{u}$ is sold at a price $p_{li} < p_{di}$ in each market i. We restrict the analysis to the case of a "poor" generic that still commands a positive market share ($\underline{u} > u_g$) in order for the following indifferent types to emerge:

$$\tilde{t}_i = (p_{di} - p_{li})/(u_d - u_l)$$

$$\underline{\underline{t}}_i = (p_{li} - p_g)/(u_l - u_g)$$

$$\underline{t} = p_g / u_g$$

so that in a given market i, the incumbent goods are sold to the higher-end of the market ($\tilde{t}_i \geq \underline{\underline{t}}_i \geq \underline{t}$). The expression of the profit then is $\pi_i = (T_i - \tilde{t}_i)p_{di} + (\tilde{t}_i - \underline{\underline{t}}_i)p_{li}$ resulting in:

$$
\begin{aligned}
p_{di} &= [T_i(u_d - u_g) + p_g]/2 \\
p_{li} &= [T_i(u_l - u_g) + p_g]/2 \\
\tilde{t}_i &= T_i/2 \\
\underline{\underline{t}}_i &= T_i/2 - p_g/[2(u_l - u_g)] \\
\pi_i &= \frac{[T_i(u_l - u_g) + p_g]^2}{4(u_l - u_g)} + \frac{T_i^2(u_d - u_l)}{4}.
\end{aligned}
$$

Notice that the price of the high-quality good is unchanged compared to the corresponding expressions with only one good (Lemma 1), while its sales are identical to the case without generic (Proposition 1). Aggregate profits and consumer surplus are:

$$
(12)\quad \begin{cases}
\Pi_d = \int_{1-x}^{1+x} \frac{\pi_i}{2x} dT = \frac{(u_d - u_g)(3 + x^2)}{12} + \frac{p_g[p_g + 2(u_l - u_g)]}{4(u_l - u_g)} \\
\Pi_{g,d} = \int_{1-x}^{1+x} \frac{(\underline{\underline{t}}_i - \underline{t})p_g}{2x} \mathrm{d}T = \frac{p_g}{2}[1 - \frac{p_g(2u_l - u_g)}{(u_l - u_g)u_g}] \\
CS_d = \int_{1-x}^{1+x} (\frac{\int_{\tilde{t}_i}^{T}(tu_d - p_{di})\mathrm{d}t + \int_{\underline{\underline{t}}_i}^{\tilde{t}_i}(tu_l - p_{li})\mathrm{d}t + \int_{\underline{t}}^{\underline{\underline{t}}_i}(tu_g - p_g)\mathrm{d}t}{2x}) \mathrm{d}T = \\
\quad \frac{u_d(3 + x^2)}{24} + \frac{3(u_g - p_g)^2}{8u_g} + \frac{p_g^2 u_l}{8u_g(u_l - u_g)} + \frac{x^2 u_g}{8}
\end{cases}
$$

Since the monopolist uses the two variants when it could always mimic the single-product offer, by revealed preference the monopolist makes higher profits with two variants. If we compare eq. (12) to eq. (6) we can also see that the generic profit and consumer surplus are higher than when the monopolist supplies no second product.

(ii) With international exhaustion

When the firm cannot discriminate among the various markets it sells at a common price p_u the higher-quality product u_u and the lower-quality product u_l is sold at a price p_l. These are now the indifferent types ($\tilde{t} \geq \underline{\underline{t}} \geq \underline{t}$): $\tilde{t} = (p_u - p_l)/(u_u - u_l)$, $\underline{\underline{t}} = (p_l - p_g)/(u_l - u_g)$, $\underline{t} = p_g/u_g$. Depending on the dispersion of the willingness-to-pay, the incumbent maximises:

$$
\Pi_u = \int_{\max[1-x,\tilde{t}]}^{1+x} \frac{T - \tilde{t}}{2x} p_u \mathrm{d}T + \int_{\max[1-x,\tilde{t}]}^{1+x} \frac{\tilde{t} - \underline{\underline{t}}}{2x} p_l \mathrm{d}T + \int_{\max[1-x,\underline{\underline{t}}]}^{\max[1-x,\tilde{t}]} \frac{T - \underline{\underline{t}}}{2x} p_l \mathrm{d}T.
$$

This gives rise to the following solutions, which can be divided into three regions:

$$(13)\qquad \text{If } x \le \frac{1}{2} \begin{cases} p_u = (p_g + u_u - u_g)/2 \\ p_l = (p_g + u_l - u_g)/2 \\ \tilde{t} = 1/2 \\ \underline{\underline{t}} = 1/2 - p_g/[2(u_l - u_g)] \\ \Pi_u = (p_g + u_l - u_g)^2/[4(u_l - u_g)] + (u_u - u_l)/4 \end{cases}$$

$$(14)\qquad \text{If } \frac{1}{2} < x \le \hat{x} \begin{cases} p_u = [3(p_g - u_g) + 2u_u(1+x) + u_l(1-2x)]/6 \\ p_l = (p_g + u_l - u_g)/2 \\ \tilde{t} = (1+x)/3 \\ \underline{\underline{t}} = 1/2 - p_g/[2(u_l - u_g)] \\ \Pi_u = (p_g + u_l - u_g)^2/[4(u_l - u_g)] + (u_u - u_l)(1+x)^3/(27x) \end{cases}$$

$$(15)\qquad \text{If } x > \hat{x} \begin{cases} p_u = \dfrac{p_g + (u_u - u_g)(1+x)}{3} \\ p_l = \dfrac{p_g + (u_l - u_g)(1+x)}{3} \\ \tilde{t} = (1+x)/3 \\ \underline{\underline{t}} = (1+x)/3 - 2p_g/[3(u_l - u_g)] \\ \Pi_u = \dfrac{[p_g + (u_l - u_g)(1+x)]^3}{27(u_l - u_g)^2 x} + \dfrac{(u_u - u_l)(1+x)^3}{27x} \end{cases}$$

where $\hat{x} = \frac{1}{2} + \frac{p_g}{2(u_l - u_g)}$.

Eq. (13) corresponds to full coverage, eq. (14) to partial coverage of the high-quality good (but full coverage of the low-quality good) and eq. (15) to partial coverage of both variants. Since in each regime the incumbent decides to introduce the second variant against the non-strategic generic, it must be that the incumbent is better off and that profits in eq. (13)-(15) are higher than profits in eq. (7)-(8) for a given *x*. Notice that sales of the high-quality good are again identical to the case without generic (eq. (2)). We can also derive expressions for aggregate profits and consumer surplus depending on the threshold values for dispersion:

$$\text{If } x \le \frac{1}{2} \begin{cases} \Pi_{g,u} = \displaystyle\int_{1-x}^{1+x} \frac{(\underline{\underline{t}} - \underline{t})p_g}{2x} \mathrm{d}T = \frac{p_g}{2}[1 - \frac{p_g(2u_l - u_g)}{2(u_l - u_g)u_g}] \\ CS_u = \displaystyle\int_{1-x}^{1+x} (\frac{\int_{\tilde{t}}^{T}(tu_u - p_u)\mathrm{d}t + \int_{\underline{\underline{t}}}^{\tilde{t}}(tu_l - p_l)\mathrm{d}t + \int_{\underline{t}}^{\underline{\underline{t}}}(tu_g - p_g)\mathrm{d}t}{2x})\mathrm{d}T = \\ \dfrac{u_u(3+4x^2)}{24} + \dfrac{3(u_g - p_g)^2}{8u_g} + \dfrac{p_g^2 u_u}{8u_g(u_l - u_g)} \end{cases}$$

$$
\text{If } \frac{1}{2} < x \le \hat{x}
\begin{cases}
\Pi_{g,u} = \int_{1-x}^{1+x} \frac{(\underline{\underline{t}}-\underline{t})p_g}{2x}\mathrm{d}T = \frac{p_g}{2}[1-\frac{p_g(2u_l-u_g)}{2(u_l-u_g)u_g}] \\
CS_u = \int_{\tilde{t}}^{1+x} (\frac{\int_{\tilde{t}}^{T}(tu_u-p_u)\mathrm{d}t + \int_{\underline{\underline{t}}}^{\tilde{t}}(tu_l-p_l)\mathrm{d}t + \int_{\underline{t}}^{\underline{\underline{t}}}(tu_g-p_g)\mathrm{d}t}{2x})\mathrm{d}T + \\
+ \int_{1-x}^{\tilde{t}} (\frac{\int_{\underline{\underline{t}}}^{T}(tu_l-p_l)\mathrm{d}t + \int_{\underline{t}}^{\underline{\underline{t}}}(tu_g-p_g)\mathrm{d}t}{2x})\mathrm{d}T = \\
\frac{2u_u(1+x)^3}{81x} + \frac{u_l(2x-1)(16-x+46x^2)}{648x} + \frac{3(u_g-p_g)^2}{8u_g} + \frac{p_g^2 u_u}{8u_g(u_l-u_g)}
\end{cases}
$$

$$
\text{If } x > \hat{x}
\begin{cases}
\Pi_{g,u} = \int_{\underline{\underline{t}}}^{1+x} \frac{(\underline{\underline{t}}-\underline{t})p_g}{2x}\mathrm{d}T + \int_{1-x}^{\underline{\underline{t}}} \frac{(T-\underline{t})p_g}{2x}\mathrm{d}T = \\
\frac{p_g}{9}[\frac{7x-1-x^2}{x} - \frac{9p_g}{u_g} - \frac{2p_g(1+x)}{x(u_l-u_g)} - \frac{p_g^2}{x(u_l-u_g)^2}] \\
CS_u = \int_{\tilde{t}}^{1+x} (\frac{\int_{\tilde{t}}^{T}(tu_u-p_u)\mathrm{d}t + \int_{\underline{\underline{t}}}^{\tilde{t}}(tu_l-p_l)\mathrm{d}t + \int_{\underline{t}}^{\underline{\underline{t}}}(tu_g-p_g)\mathrm{d}t}{2x})\mathrm{d}T + \\
+ \int_{\underline{\underline{t}}}^{\tilde{t}} (\frac{\int_{\underline{\underline{t}}}^{T}(tu_l-p_l)\mathrm{d}t + \int_{\underline{t}}^{\underline{\underline{t}}}(tu_g-p_g)\mathrm{d}t}{2x})\mathrm{d}T + \int_{1-x}^{\underline{\underline{t}}} (\frac{\int_{\underline{t}}^{T}(tu_g-p_g)\mathrm{d}t}{2x})\mathrm{d}T = \\
\frac{2u_u(1+x)^3}{81x} + u_g\frac{69x-12x^2+23x^3-4}{162x} + \\
p_g[\frac{2-23x+2x^2}{27x} + \frac{p_g}{2u_g} + \frac{2p_g(1+x)}{27x(u_l-u_g)} + \frac{2p_g^2}{81x(u_l-u_g)^2}]
\end{cases}
$$

Comparisons with the expressions for consumer surplus with arbitrage but only one quality (Lemma 1) show that the introduction of the lower-quality variant has a positive impact on it.

Proof of proposition 5

We first conduct welfare comparisons *ex post*, for a given quality level ($u_u = u_d = u$) and with the same low-quality patented good. Denote by d (respectively d_l) the quality differential between the high-quality good and the generic good, $d = u - u_g$ (respectively between the low-quality good and the generic good, $d_l = u_l - u_g$). It results:

$$
\text{If } x \le \frac{1}{2}
\begin{cases}
\Delta\Pi = \Pi_u - \Pi_d = -dx^2/12 < 0 \\
\Delta\Pi_g = 0 \\
\Delta CS = dx^2/8 > 0 \\
\Delta Welfare = dx^2/24 > 0
\end{cases}
$$

$$
\text{If } \frac{1}{2} < x \le \hat{x} \begin{cases} \Delta\Pi = -dx^2/12 + d_l(2x-1)^2(4+x)/(108x) < 0 \\ \Delta\Pi_g = 0 \\ \Delta CS = dx^2/8 - d_l(2x-1)(16-x+46x^2)/(648x) > 0 \\ \Delta Welfare = dx^2/24 - d_l(2x-1)(40-43x+34x^2)/(648x) > 0 \end{cases}
$$

$$
\text{If } x > \hat{x} \begin{cases} \Delta\Pi = d[-\frac{x^2}{12} + \frac{(2x-1)^2(4+x)}{108x}] + \frac{p_g(2-5x+2x^2)}{18x} + \frac{p_g^2(4-5x)}{36d_l x} + \frac{p_g^3}{27{d_l}^2 x} \\ \Delta\Pi_g = \frac{d_l(4\hat{x}-x)(x-\hat{x})(2\hat{x}-1)}{9x} > 0 \\ \Delta CS = d[x^2/8 - (2x-1)(16-x+46x^2)/(648x)] + \Omega(p_g, x, d_l) \\ \Delta Welfare = d[x^2/24 - (2x-1)(40-43x+34x^2)/(648x)] + \Psi(p_g, x, d_l) \end{cases}
$$

Looking at the corresponding expressions in Proposition 3 (eq. (9) and eq. (10)), it turns out that a similar analysis applies now. Thus the *ex post* welfare effects resemble those of the 1-product case as found in Proposition 3[16]. By denoting by x^{***} the root that makes $\Delta Welfare = 0$ in the equation above, this means that $\Delta Welfare(x^*) > 0$ and that welfare decreases only in the 'partial' coverage regime if the dispersion of x is sufficiently high, with $x^{***} > x^*$. If p_g is small, it is also easy to see that $\Psi(p_g, x, d_l) > \Psi(p_g, x, d)$, and thus $x^{***} > x^{**}$, otherwise the inequality can be reversed.

Ex ante, quality is exactly the same as under pure monopoly with no generic. In fact, from eq. (12) and from eq. (13)-(15) one derives the marginal revenues in the different regimes in the presence of a generic and two variants:

$$
\frac{\partial \Pi_d}{\partial u} = (3+x^2)/12; \quad \frac{\partial \Pi_u}{\partial u} = \begin{cases} 1/4 \text{ if } x \le 1/2 \\ (1+x)^3/(27x) \text{ if } x > 1/2 \end{cases}
$$

that are identical to marginal revenues under pure monopoly (these are derived from eq. (1) and (2)) for a given degree of dispersion. Thus, contrary to eq. (11) that was derived assuming the incumbent would stick to one product, the generic now has no impact on investments.

We have considered u_l as exogenous so far. It is easy to check that all the expressions for the monopolist's profits are decreasing in the lower-quality variant. Hence the patent holder will set the lowest possible one. If we imagine for computational simplicity that the incumbent can mimic the generic quality, its best reply is to offer $u_l = u_g$ at a price just below p_g. In this way the incumbent displace the generic producers and supplies the generic itself. With the quadratic cost function, the *ex ante* welfare properties of parallel trade are:

[16] There are some minor differences as there are now three regimes. 'Full' coverage of both variants happens for a smaller range of the dispersion parameter; the partial coverage regime (i.e. when the monopolist drops all its products) also arises for a more limited range of dispersion as $\hat{x} = 1/2 + p_g/2d_l > \hat{x} = 1/2 + p_g/2d$.

$$\Delta Welfare = \begin{cases} -x^4/(144k) - 6u_g x^2/(144k) < 0 \text{ if } x \le 1/2 \\ \dfrac{7(1+x)^6}{4374kx^2} - \dfrac{(3+x^2)^2}{144k} + u_g \underbrace{\dfrac{41x^3 - 120x^2 + 123x - 40}{648x}}_{A} < 0 \text{ if } x > 1/2 \end{cases}$$

These expressions should be contrasted with those in the pure monopoly case (Proposition 2). Recalling that the impact of the generic is to reduce the incentives to discriminate in the absence of parallel trade, the overall welfare analysis is quite clear. When every market is served, the generic has no impact on investment ex ante but there are less gains ex post from the reallocation of goods thanks to parallel trade. When the monopolist drops the superior product from some markets ($x > ½$), again there is no impact on investment but ex post the negative role of parallel trade is diluted (the term A above has one root at x^*, and is positive for $x > x^*$). With the quadric cost function, taking into account that u_g has to be lower than the quality of the high-quality product, the negative effect of parallel trade on investment always prevails.