Parallel Imports and Price Controls – and Compulsory Licensing!

Rafael Pinho Senra de Morais

Getúlio Vargas Foundation

Toulouse School of Economics

This Version: October 2008

JEL classification: O34, F13, I18

Keywords: Pharmaceuticals, TRIPs, Compulsory Licensing, Price Regulation
Abstract

Grossman and Lai (2008) challenge the orthodox view that parallel trade in pharmaceuticals undermine intellectual property rights and dull the incentives for investment in this research-intensive industry. Their paper show that parallel trade impacts innovation and welfare in the North (innovative region) positively. Although we take the same approach to drugs regulation – namely that price controls create opportunities for international arbitrage when parallel trade is tolerated – we challenge the generality of their results. The authors themselves explicitly recognize that, although a North-South model, their paper is focused on the relation between USA and Canada, or Northern Europe versus Southern Europe.

We adapt their model to represent genuine North-South relationships: those between the innovative USA, Europe or Japan, and the developing world. We do this by allowing the Southern country to issue compulsory licenses on patented medicines in accordance with the TRIPs Agreement. In particular, drugs produced under compulsory licenses are for domestic usage. Our results turn out to challenge both the orthodox wisdom and the results of Grossman and Lai (2008): we show that if compulsory licensing is an instrument available to southern countries, the tolerance of parallel trade by northern countries has absolutely no impact on R&D incentives and innovation. It actually allows the South to recover most of the welfare losses caused by the authorization of parallel trade by the North.

Although we are in principle more concerned about developing countries holding drug manufacturing facilities and therefore capable of imitation like Brazil, South Africa, India or Thailand, the analysis naturally applies to least developed economies, as the recently approved “paragraph 6 system” allows the use of compulsory licensing for importation by countries unable to produce the medicines themselves.
I - Introduction

Intellectual property rights (IPRs) are all about restricting access to inventions, which is necessary for creating economic incentives for investment in Research and Development (R&D). As concerns pharmaceutical innovation, the patent system interacts with the regulation of drug prices for creating those incentives. On top of that, in the WTO era, those policy instruments interact with activities from the international arena, such as parallel trade and (potential and effective) compulsory licensing.

Drug prices differences across countries, as pointed out by Grossman and Lai (2008)\(^2\), emerge from retailer price discrimination, vertical price restraints or different regulations. Their paper and ours focus on the last one. The classic tradeoff about access to medicines and the mechanisms to fund R&D for new and better medicines is the relevant one when a country chooses the price ceiling to impose on some medicine.

However, when parallel trade comes into play, the regulated price cap becomes a strategic variable in an international setting. Parallel trade is an arbitrage mechanism through which the drugs produced in a low-price market flow to the high-price market, reestablishing uniform prices in the presence of a large number of parallel traders and absence of frictions (transportation and transaction costs). The resulting world price is the lowest among all prices set for an individual drug worldwide, if (costless) reimportation is allowed.

As concerns public health and compulsory licensing, the restrictions imposed by the IPRs system on access to (patented) drugs must be reasonable, and they must not create situations where entire populations are denied access to known therapies. Therefore, the WTO TRIPS Agreement contains provisions on safeguards, exceptions and mechanisms to protect essential public health interests. The strongest of those is certainly the possibility of allowing production and sales of a patented medicine without the authorization of the patent owner.

I.1 – The relevance of Compulsory Licensing

As already presented by Reik (1946), compulsory licensing is not only justifiable in the case of non-working (i.e. if the patent holder does not produce the good), but should take place whenever “the immediate need of the nation outweighs the absolute right of the patentee to exclude all others from in any way utilizing the patented invention throughout the lifetime of the patent”.\(^3\) In practice, it is a legal instrument commonly used inside developed

\(^2\) From now on, we might refer to Grossman and Lai (2008) as GL.

\(^3\) Another paragraph from Reik (1946) also deserves being cited here: “The existence of the public interest has generally been affirmed if the domestic demand for a patented article was not being met to an adequate extent. However, it is held that this applies only to cases in which a serious need of the
economies, including the USA, in an attempt to overcome the existing barriers in new high technology medical treatments based upon biotechnology and gene therapies.

Somehow contradictorily, when it comes to the international context, most large pharmaceutical companies (based in those same developed economies) oppose compulsory licensing of medicines, in particular in poor countries, in spite of public opinion’s concerns over the breadth of access to medicines. The labs believe poor countries will abuse the compulsory licensing system, this will lower returns from research and development (R&D) and undermine their ability to charge high prices in the more developed countries.

Our focus in this paper is on this compulsory licensing which generates an international conflict of interest: one country licensing drugs whose (valid) patent belongs to a foreign firm. Conditions for such licensing are stated in article 31 of the TRIPS Agreement, whose goal is to ensure the interests of patent owners and national governments are balanced. However, concepts like “national emergency or other circumstances of extreme urgency” are quite vague, and moreover are not even required “in cases of public non-commercial use” (art. 31, b). This leaves a huge margin of maneuver for the South, as far as “any such use shall be authorized predominantly for the supply of the domestic market of the Member authorizing such use” (art. 31, f).

I.2 – Grossman and Lai (2008) and the present paper

Grossman and Lai (2008) provide a general model for the game between two regions involving the allowance or not of parallel trade in drugs, the responsiveness of the price cap setting activity to such regime choice and the impact of both variable choices on the pace of innovation and thus welfare in both regions. Their baseline model assumes only the North innovates and only the South sets a binding price ceiling and is solved by backwards induction. Extensions are provided for the cases where both regions innovate, where both exert effective price regulation, and also for the case where there are many countries in the South. Results are qualitatively similar to the baseline case.

If parallel trade is ruled out, northern firms are willing to sell to southern countries at any price above marginal cost. Instead, if parallel trade is allowed, northern firms will hesitate more in serving the South, as an international exhaustion regime implies uniform pricing.

---

4 Grossman and Lai (2008) use both concepts (parallel trade and international exhaustion of IPRs) interchangeably. The Law literature emphasizes that under international exhaustion the imported goods have necessarily been introduced by the patent holder himself or some licensee, while the goods object of parallel importation might have been marketed by a third party without any authorization (in countries in TRIPS transitional period, or even non-WTO members). Such distinction is irrelevant in GL the South is a developed economy but we are more comfortable with such substitutability of concepts by emphasizing that we focus on WTO Members, respectful of TRIPs.
Consequently, there is a required minimum price for serving the South which is the one making northern innovative firms indifferent between selling in both markets at such southern price cap or selling only in the North at monopoly price (and foregoing sales in the South).

The driving force for the results in GL is the incentives faced by the South when setting its price controls, in particular when parallel trade is allowed by the North. There are two such streams pushing in opposite directions: the incentives for the South to free-ride on North’s innovation – already present in Lai and Qiu (2003), Grossman and Lai (2004) and Morais (2005) – and the incentives for the South to ensure its consumers have access to northern drugs, i.e. its fear of not being served, which is introduced by GL.

In a regime permitting parallel trade, the first incentive is mitigated and the second appears, both pushing the regulated price in the South upwards. The South has less incentive to free-ride on North’s innovation, as a low regulated price in the South would trigger parallel trade and undermine worldwide incentives to innovate. Also, the novelty reinforcing such effect is that the South is concerned about ensuring that its consumers are served.

We argue in the present paper that this second effect disappears (even if parallel trade is allowed) when compulsory licensing is truly a policy instrument available to the southern government. It is the fear of having these drugs shipped to the high-income markets that sometimes induce innovative pharmaceutical companies not to serve poor markets in an international exhaustion regime. Such fear vanishes under a credible threat of compulsory licensing. Moreover, by anticipating that a poor country has a true possibility of obtaining a patented good at a low price (possibly as low as at marginal cost) even without the authorization of the IPR holder, the latter will have some incentive to lower her price in order to remain the provider for the poor country.

In this paper, we focus explicitly on the pharmaceutical market and compulsory licensing of northern medicines by southern governments. We extend the analysis in Grossman and Lai (2008) to allow the South to potentially issue compulsory licenses (for domestic usage) and analyze the effects of the existence of such trade policy instrument on the regulated drug prices, innovation and welfare.

Drug price regulation is of particular interest to developing and least-developed economies, which are commonly keen on constraining such prices in order to avoid monopoly pricing by foreign laboratories. In these poor countries, compulsory licensing is increasingly seen as an alternative measure for obtaining access to existing drugs at affordable prices. Although more costly at a first moment (price regulation only requires imposing a price cap, while licensing requires bargaining with the IPR holder, notification of the WTO and effective production or domestic licensing), compulsory licensing can be preferred to price regulatory measures: the latter carries a risk for the South of not being served, while the
This paper focuses on the interaction between those instruments and their use by the South in face of a reactive North where the innovative firms are located.

Finally, the model presented is more directly applicable to developing countries holding drug manufacturing facilities – those capable of performing drugs imitation like Brazil, South Africa, India or Thailand. However, the analysis applies naturally to least-developed economies, since the recent approval of the TRIPS “paragraph 6 system”. This refers to paragraph 6 of the Doha Declaration on the TRIPS Agreement and Public Health, agreed to at the 2001 WTO ministerial in Doha, Qatar. In such “system”, any country is allowed to issue compulsory licenses and produce essential medicines for exportation to countries unable to produce the medicines themselves. A waiver allowing for this was agreed to in 2003, and this was made a permanent amendment to the TRIPS Agreement on 6 December 2005. Although by August 2008, only one WTO notification has been filed\(^6\), we believe this instrument will increasingly be used. In any case, only its existence is enough for the purposes of this paper.

II - Model setting

In this section we develop our model, which builds directly on Grossman and Lai (2008). As a first step, we present the basic setup of their model and link it to our setup, where developing countries have compulsory licensing of drugs as a potential policy instrument, on top of price cap regulation.

II.1 – Grossman and Lai’s general and baseline models

Grossman and Lai (2008) propose the following three-stage game (the reader is asked to refer to Grossman and Lai (2008) for full details on their setup and solution):

1\(^{st}\) stage: North chooses whether or not to allow for parallel importation of drugs\(^7\)

---

\(^5\) One could argue that regulation requires much information-gathering and expertise. The potential use of international reference pricing invalidates such argument. International reference pricing refers to establishing the price ceiling in a given country for a given medicine as a (weighted) average of its (regulated) prices in a bundle of relevant countries.

\(^6\) Rwanda filed submission IP/N/9/RWA/1 dated 19 July 2007, concerning the importation of TriAvir – a fixed-dose combination product against Aids, whose original patent belongs to GlaxoSmithKline – to be manufactured in Canada for exportation to Rwanda by Apotex Inc., the largest Canadian drug firm.

\(^7\) The decision in the 1\(^{st}\) stage is a long-term commitment by the northern government, sustained by an appropriate reputation mechanism, the same way as in GL. For the equivalence of the existence of parallel trade and international exhaustion of IPRs, see a previous footnote.
2\textsuperscript{nd} stage: Each region sets a potentially binding price cap on drugs $\overline{p}_i$ for $i = N, S$ (The North sets a price cap $\overline{p}_N = \infty$ in the baseline model)

3\textsuperscript{rd} stage: The innovative firms (located only in the North in the baseline model) set prices $p_S$ and $p_N$, which have to be lower than the price ceilings established in the 2\textsuperscript{nd} stage if they want to sell their products in a given market.

The game is solved by backwards induction. We shall consider the baseline model now. The North compares its welfare in each exhaustion regime of IPRs, anticipating the subsequent choices of the South. The South maximizes its welfare $W^j_S$ where $j \in \{ne, ie\}$ by choosing its price cap, given the exhaustion regime of IPRs chosen by the North ($ne$ means national exhaustion or no parallel trade while $ie$ stands for international exhaustion).

If the North had decided in the 1\textsuperscript{st} stage for no parallel trade, northern firms will be willing to serve the southern market at any price above marginal cost. In case of international exhaustion of IPRs, if a firm charges different prices in the two markets, parallel traders will cover the high-priced market with products from the low-priced market.\textsuperscript{8} As parallel trade eliminates the power to discriminate prices across markets, northern firms will hesitate more in serving the southern market.

The South has no interest in leaving rents to northern firms. However, if the regime is international exhaustion, it anticipates the 3\textsuperscript{rd} stage and incorporates an extra constraint to its maximization problem for ensuring its consumers are served. On the other hand, since through price controls (2\textsuperscript{nd} stage) the South can impose its preferred price level (constrained by parallel trade or not) upon northern firms, its price ceiling is always binding.

As for the North, as its price is not controlled, under national exhaustion we have $p_N = p_M$ (the monopoly price firms choose) and under international exhaustion $p_N = p_S$ (due to binding price caps and no arbitrage imposed by parallel trade).

The main results of GL’s baseline model are analysed below with the support of their Figure 1, reproduced below. The graph shows the optimal choice of the binding price ceiling $p_S$ as a function of the relative market sizes $m = \frac{M_N}{M_S}$ in both IPRs regimes.

The values of $\tilde{p}$ and $p^*$ (in the baseline model) are defined as follows:

$\tilde{p}$ is the price yielding the same profit for the monopolist if he sells in both North and South at a uniform price (below the monopoly price, equal to the price ceiling in

\textsuperscript{8}Transportation and marketing costs are ignored for simplicity, just like in GL. Therefore – in GL’s words – “price arbitrage is perfect in a regime that permits parallel trade”.
the South), or if he foregoes sales in the South and sells only in the North at monopoly price (as there is no price control in the North in the baseline model).

\[ p^* \] is the price ceiling that the South would enforce in a regime with parallel trade if it were not afraid of not being served, i.e. if it were not subject to the constraint \( p_s \geq \tilde{p} \).

For the lower part of Figure 1, \( p_s^{ne} \) shows that in the absence of parallel trade, the South sets a price equal to marginal cost if its market size is small enough (for \( m > m_1 \)), as northern firms are willing to serve the South at any price higher or equal to cost. For \( m < m_1 \), the South is large enough to be concerned about the impact of its price ceiling on overall profits of northern firms and their incentives to innovate, as these determine the number of goods invented at steady state, which benefit (also) southern consumers. For \( m < m_1 \), the larger the market in the South, the higher is the price cap chosen by the Regulator in South. In the limit, if the South is huge, being the only region responsible for providing incentives to innovation (i.e. if the size of the northern market is insignificant), the price cap chosen in

---

\[ 9 \text{ In GL's paper own words: "It cannot be optimal for South's government to allow its market for differentiated products to be foreclosed. Accordingly, the government of South chooses its price ceiling } p_s \text{ subject to the constraint that } p_s \geq \tilde{p}." \] In this paper, however, the South can use price controls strategically: South might prefer choosing a too low price cap in order not to be served, allowing it to use compulsory licenses in a further stage. This case will be explored later on.
South will be the same as the one it would choose under international exhaustion (as shown by the intersection of both bold curves at $p^*$. This is the first crucial insight.

As for the upper part of Figure 1, when parallel trade is allowed by the North, the price cap imposed by the South is the uniform price worldwide (if the South is served). A parallel trader will profit from any arbitrage possibility and such threat will force innovative firms to charge the same price in both regions, in case they wish to serve both markets. The incentives for the South, though, when setting its price cap $p_s$ will incorporate such externality on northern prices and overall R&D incentives. The southern government thus has no incentive to set any price cap below $p^*$ irrespective of how tiny it is, as the South knows its price control will determine incentives to worldwide R&D.\(^{11}\)

On the other hand, in case of international exhaustion, in GL the South is also concerned about being served. This is expressed by $\tilde{p}$, the regulated price in the South making northern firms indifferent between selling in both regions at that price or only in the North at monopoly price. Clearly, $\tilde{p}$ is increasing in the northern relative market size: as the market in the North becomes insignificant ($m \to 0$), firms in the North are willing to serve the South at any price above marginal production cost; as the market in the South becomes insignificant ($m \to \infty$), firms would only accept to serve the South at monopoly price.

In GL, if under international exhaustion the South charges a price below $\tilde{p}$, the South is sure of not being served. This is the second main insight, pushing the South in the presence of parallel trade to be constrained to setting its price cap at $p_s^{ie} = \max \{ p^*, \tilde{p} \}$. This is exactly what will be crucially different in our model, when the South is allowed to issue compulsory licenses for drugs it is not being provided – and this will be anticipated.

In Appendix A we state and comment the main results of the baseline model in GL, to which our results will refer later on. It should be emphasized, however, that GL has results concerning welfare in North and South individually but none on overall welfare.

II.2 – Our baseline model

While Grossman and Lai (2008) are concerned about intra-European or North-American (USA-Canada) parallel trade, in the present paper we focus on a non-cooperative sequential game involving a developed North and a developing South. For this reason, at first we assume – similarly to GL in their baseline model – that only the North innovates and only

\(^{10}\) An underlying assumption is that the South is not subject to any capacity constraint, i.e. facilities in the South can potentially produce drugs in quantities sufficient to cover both regions.

\(^{11}\) GL’s words: “$p^*$ is independent of $m$ under $ie$; this is because arbitrage renders the effect of price controls on the incentives for innovation independent of the geographic distribution of world demand.”
the South sets binding price controls. Firms in the North invent the relevant medicines and are allowed to price freely in their home country, while the South does not innovate and regulates the domestic prices of prescription drugs. As our paper deals with genuine North-South trade, these assumptions are arguably more representative of real situations here, as compared to GL, where the baseline model works as a benchmark to facilitate intuition.

We keep the same notation and whole structure of the model in Grossman and Lai (2008), except that we add an extra stage in the game they propose, in order to incorporate the potential use of compulsory licensing by southern governments.\textsuperscript{12} Our game thus includes four stages, where the first three stages are precisely the ones in Grossman and Lai (2008):

1\textsuperscript{st} stage: North chooses whether or not to allow for parallel importation of drugs

2\textsuperscript{nd} stage: Each region sets a potentially binding price cap on drugs $p_i$ for $i = N, S$

(The North sets a price cap $p_N = \infty$ in the baseline model)

3\textsuperscript{rd} stage: The innovative firms (located only in the North in the baseline model) set prices $p_S$ and $p_N$, which have to be lower than the price ceilings established in the 2\textsuperscript{nd} stage if they want to sell their products in a given market.

4\textsuperscript{th} stage: If not served, the South decides whether to issue a compulsory license and serve domestic consumers at the exact same production cost $c$, incurring however some positive (small) fixed cost of compulsory licensing.

Compulsory licensing refers to covering the domestic market only, which is in accordance with the TRIPs Agreement. We follow GL where “the formerly patented good can be produced by firms anywhere in the world at a unit cost of $c$”.\textsuperscript{13} We extend their approach \textit{vis-à-vis} expired patents and imitation towards compulsory licensing. We thus do not even need to argue on facilities in the South being equally efficient to northern ones, any difference in productivity being reflected in wages. We need, however, to assume that those southern facilities are public or that there is competition among southern labs for the production of imitations, in a way that the market is served at $p_S = c$ when a compulsory license is issued.\textsuperscript{14}

On top of that, there is nonetheless some fixed cost associated with the issuing of compulsory licenses. These are due to administrative procedures, but also (potentially) some reputational loss and retaliation – and eventually fixed royalties to the patent owner. The

\textsuperscript{12} As mentioned previously, the reader is asked to refer to Grossman and Lai (2008) for further details on the setup.

\textsuperscript{13} As in the general equilibrium model in GL, here too “any wage gap that exists between North and South reflects a difference in the productivity of labor in all of its uses”, which implies the same production cost worldwide.

\textsuperscript{14} As argued previously, in the era of the Paragraph 6 system, this becomes a reasonable assumption, even for least developed countries.
underlying idea is that it is cheaper for the South to regulate the prices of northern drugs than to engage in a whole WTO procedure for the license. As such, compulsory licensing is not efficiency-enhancing per se.\textsuperscript{15} Instead, it is a loophole for the South to avoid high northern prices despite it being cheaper to produce in the North after the drug has been discovered, as marginal cost is the same but there is no fixed cost (of compulsory licensing) in the North.

We assume the same Condition 1 holds and solve the game by backwards induction. As the South can potentially license patented drugs without the authorization of the patent holder – differently from Grossman and Lai (2008) – this region is therefore not thrilled by the perspective of not being served if setting a too low regulated price. However, if issued, a compulsory license implies some fixed cost. Moreover, it is restricted to domestic use: those drugs licensed compulsorily are constrained by the South’s boundaries, and thus can not be object of parallel trade, even if the northern region had decided for an international exhaustion regime and charges a higher price. As such, compulsory licensing represents a way to ensure southern consumers are served and incentives to northern firms are dulled only to the extent of the market size in the South, as compulsory licensing represents an exception to the rule of uniform pricing in a regime allowing parallel trade.

II.3 – South’s strategic behavior

The South is called to play twice in the game we present in this paper. In the 2\textsuperscript{nd} stage, knowing the IPR regime chosen by the North, it sets the ceiling on the drug prices the northern firms can charge. In the 4\textsuperscript{th} (and new) stage, it decides whether or not to issue compulsory licenses, in case northern firms have preferred not to sell in the South.

The South will not be that concerned about not being served (due to a low regulated price in the second stage), as compulsory licenses can follow. On top of that, by knowing that compulsory licenses only undermine incentives to R\&D to the extent of its market size, the South will not hesitate much in issuing compulsory licenses in the final stage, although it has to bear some (small) fixed cost.

II.3.1 – South’s optimal choices with National Exhaustion

If the North rules out parallel trade, any patented drug reimportation is illegal per se. In such case, there is no need to emphasize that compulsory licensing refers to serving the domestic market only – which will be essential in our analysis of international exhaustion.

\textsuperscript{15} Compulsory licensing has some transaction cost that is not present in setting price caps. Regulation, thus, is a cheaper way (more efficient “technology”) to achieve the same goal, as argued previously.
In the 4th stage of the game, the South will issue compulsory licensing if it is not served, and will have consumers getting the drug at \( p_S = c \) and will bear some small fixed cost. (If the non-served South does not issue compulsory licenses, it gets zero welfare.)

However, in the 3rd stage, northern firms are willing to serve the South at any price above marginal cost \( c \), and the South knows it when it chooses the price control in the 2nd stage. As such, the South has to compare \( W_S^{CL} \) and \( W_S^{ne}(p_S) \), where the former is the welfare when compulsory license is used and the latter is the welfare obtained under national exhaustion by setting a price cap \( p_S \) (as in GL and Figure 1). If welfare in the South were higher under compulsory licensing, then the South would (strategically) set a very low price ceiling (\( p_S^{ne} < c \)), in order not to be served, which is a requisite for compulsory licensing. However, our first result shows that this is never the case.

**Proposition 1:** Suppose that the South does not innovate and the North imposes no price controls. In a national exhaustion regime, the South’s optimal choice is not influenced by the potential use of compulsory licensing.

A corollary of Proposition 1 is that Figure 1 remains valid in its lower part (ne).

**Proof:**

- For \( m \geq m_1 \), the South is small enough for not being concerned about generating incentives to R&D in the North. As no parallel trade is allowed here, the South knows northern firms will be willing to serve the southern market at any price above marginal production cost. The South can already induce \( p_S = c \) at the 2nd stage of the game by setting \( p_S = c \) and it will be served. Using compulsory licenses would simply mean obtaining the drugs at a higher cost, as southern facilities produce at the same marginal cost \( c \) but there is a (small) fixed cost of using compulsory licenses. Consequently, \( W_S^{CL} \) is lower than \( W_S^{ne}(c) \).

- For \( m < m_1 \), the South could induce \( p_S = c \) at the 2nd stage but achieves higher welfare by setting \( p_S > c \) (as shown by \( p_S^{ne} \) in Figure 1). Its market is large enough for the South to benefit more from the dynamic gains of (more) innovation than it suffers from the static losses of a price higher than marginal cost. As here \( W_S^{ne}(p_S) > W_S^{ne}(c) \), and \( W_S^{ne}(c) \geq W_S^{CL} \) as argued in the previous paragraph, it is straightforward that South’s welfare for \( m < m_1 \) is larger under the price caps in figure 1 than if the South uses compulsory license.

II.3.2 – South’s optimal choices with International Exhaustion

In this regime, the South’s choice is different from GL as the possible use of compulsory licenses changes the South’s incentives in its regulatory activity. The South is not
concerned anymore about being served when choosing price controls in the 2nd stage of the game, i.e. about setting a price at least as high as \( \tilde{p} \) in Figure 1. However, as in GL, the South cares about worldwide innovation (no matter the South’s market size) as it knows parallel trade implies uniform price, i.e. the importation by the North of the southern price controls.

- For \( m > m_2 \) the now unconstrained South would prefer setting \( p_S = p^* \) in the presence of parallel trade. However, if it does so, it will not be served as northern firms make a higher profit by only serving the North at monopoly price (check Figure 1). And if the South is not served, its optimal choice in the last stage of the game is to issue compulsory licenses. Knowing this, the South’s choice is between setting a high price control, being served and not issuing compulsory licenses, obtaining \( W_S^{ie}(\tilde{p}) \), or setting a low price control, not being served and issuing compulsory licenses, achieving \( W_S^{CL} \). The crucial point is that \( W_S^{ie}(p^*) \) is not achievable and both regions know that.\(^{16}\)

For any relative market sizes, GL’s Proposition 3 shows that southern welfare is always higher under national exhaustion (as \( p_S^{ie} \) could be chosen under national exhaustion, but some different \( p_S^{ie} \) (strictly lower) was preferred; thus \( p_S^{ie} \) necessarily generates a higher welfare). This holds for low values of \( m \) – where \( p_S^{ie} \) and \( p_S^* \) are close – and with more reason for high \( m \), when the small South has lower incentives to boost northern R&D in the absence of parallel trade and is pushed to provide such protection by uniform pricing. We thus know that \( W_S^{ie}(c) \gg W_S^{ie}(\tilde{p}) \) for \( m > m_2 \). Consequently, we can be sure that for a small fixed cost (SFC) of compulsory licenses, \( W_S^{CL} > W_S^{ie}(\tilde{p}) \), since \( W_S^{CL} = W_S^{ie}(c) – SFC \).\(^{17}\)

As a consequence, the South sets any \( p_S < \tilde{p} \) as it anticipates it can issue compulsory licenses in the last stage of the game if it is not served. The northern firm will set \( p_N = p_M \) and any \( p_S > p_S^* \) and only serve the North at monopoly price. Compulsory licensing ensures southern consumers are served at \( p_s = c \) and the South bears some SFC.

- For \( m_1 < m < m_2 \) being served or not is not an issue as the preferred choice of South under international exhaustion \( p^* \) is larger than the minimum required price ceiling for being served \( \tilde{p} \). In this case, the incentives for the South to issue compulsory licenses do not come

\(^{16}\) Somehow strikingly, in this range of \( m \) northern firms prefer compulsory licenses to being obliged to serve both regions at \( p^* \) (due to parallel trade and uniform pricing), as \( (M_N + M_S)\pi(p^*) < M_N\pi(p_M) \).

\(^{17}\) We repeatedly emphasize the fixed cost associated with compulsory licensing is small. The alternative of a high, prohibitive fixed cost is totally uninteresting for the purposes of our paper.
from any fear of not being served. Instead, it comes strictly from the desire not to take on the burden of incentivizing worldwide innovation (which comes from the uniform pricing imposed by parallel trade). The South’s choice is between \( W_S^{ie} (p^*) \) or \( W_S^{CL} \). Again, due to the small size of the fixed cost associated with compulsory licensing \( (W_S^{CL} = W_S^{ne} (c) - SFC) \) and the fact that welfare in South is always smaller under \( ie \) than \( ne \) in the absence of compulsory licensing (Proposition 3 in GL), we have that \( W_S^{CL} > W_S^{ie} (p^*) \).

**Lemma 1:** For some medium range of relative market sizes, although the South is always served, it strategically sets too low a price control and issues compulsory licensing.

In other words, for \( m_1 < m < m_2 \) the availability of compulsory licensing changes the regulatory behaviour of the South. In spite of always being served, the South changes its regulatory practice strategically, in order not to be served and afterwards be able to use compulsory licenses.

- For \( m < m_1 \) things are trickier. Compulsory licensing implies \( p_S = c \) but this was not the preferred choice of South under national exhaustion of IPRs\(^{18} \) as the South is itself concerned about the incentives to northern innovation, due to its high relative market size.

For \( m = m_1 \): \( W_S^{CL} > W_S^{ie} (p^*) \) for a sufficiently small fixed cost of compulsory licensing, for the reasons explained above (related to GL’s Proposition 3).

For \( m = 0 \): welfare is the same in both regimes since the North is insignificant and the South charges the same price in both regimes\(^{19} \): \( W_S^{ie} (p^*) = W_S^{ne} (p^*) \). We also know that \( W_S^{ne} (p^*) > W_S^{ne} (c) \) as \( p_S = c \) could have been chosen under \( ne \) but the South preferred a higher price ceiling, yielding a necessarily higher welfare. Moreover, \( W_S^{CL} = W_S^{ne} (c) - SFC \) by definition. We then obtain that \( W_S^{ie} (p^*) > W_S^{CL} \) for \( m = 0 \).

**Lemma 2:** The smaller the fixed cost associated with issuing compulsory licenses, the closer \( m_0 \) will be to 0, i.e. the smaller will be the range of relative market sizes for which compulsory licenses do not happen.

From the continuity of the welfare function, there is some (new) threshold strictly between 0 and \( m_1 \) for which \( W_S^{ie} (p^*) = W_S^{CL} \). It does not depend on SFC being strictly positive: this new threshold \( m_0 \) is always larger than 0, although closer to 0 the larger SFC.

\(^{18}\) Further on we deal with the case where the South can discretionally remunerate the patent owner.

\(^{19}\) When \( m = 0 \) the South is virtually the whole world and the distinction \( ne \) and \( ie \) loses its sense.
In other words, there always exists a range of relative market sizes \((0 < m < m_0)\) for which the South prefers to provide itself incentives to northern R&D, instead of issuing compulsory licenses.

**Proposition 2:** Suppose that the South does not innovate and the North imposes no price controls. In an international exhaustion regime, the South’s optimal choice is not always to use compulsory licensing, even in the absence of any costs associated with such practice. Compulsory license never happens when the relevant market in the South is large enough.

As seen, there is no interest to issue compulsory license when parallel trade is not allowed. When there is parallel trade, for a large enough market size in the South \((m < m_0)\), the South prefers a regulated price higher than the price that would induce northern firms not to serve the southern market.

In other words, through compulsory licensing, the small South is able to reverse back to its welfare under national exhaustion (minus some small fixed cost associated with the use of compulsory licenses) in spite of the North having chosen international exhaustion (and thus parallel trade and uniform pricing). The key feature of compulsory licensing is that the targeted drugs have to be used domestically.

However, the price chosen by the Southern government could be higher than marginal cost, and actually will be when royalties in accordance with the TRIPs system are due and effectively paid. This is explored later in the paper.

III – Extensions

As in Grossman and Lai (2008), we first extend the baseline model to allow the South to also innovate, then to the case where the North also sets a binding price control, and finally to a setting with many countries in the South.
V – Conclusion

The main intention of our work is to analyse countries’ decision on the implementation of binding price controls for medicines, which are of particular interest for the less innovative developing world, as such policy might dull incentives to innovate but provides broader consumer access to drugs. Such a priori domestic regulatory policy has an increasingly international character in the post-TRIPs arena, the best-suited setup for analyzing it being a strategic North-South model. This is due to the need of incorporating into the analysis the relevant features of an ever more globalized pharmaceutical market: parallel trade and compulsory licensing.

Grossman and Lai (2008) performed a similar exercise, without however considering the possibility of compulsory licensing, as they explicitly focused on the relation between USA and Canada, or Northern Europe and Southern Europe. We extend their model to deal with genuine North-South interaction, where an innovative North is concerned about its firms’ profits directly and through their impact on innovation, while the non-innovative South is only concerned about its consumers and thus mainly cares about access to medicines (at affordable prices) and only secondarily about the impact of its policy on innovation.

The key driving force in Grossman and Lai (2008) was the South’s fear of not being served if imposing too low a price cap on northern drugs in the presence of parallel trade (international IPR exhaustion). In the present paper, we allow the developing South to make use of compulsory licenses, and it turns out that this counterbalances the power of such threat by northern firms.

Grossman and Lai (2008) have shown that parallel trade (and the attached threat) increased R&D and thus innovation and welfare in the North, at the expense of welfare in the South, as compared to the situation where there is no need for interactive decision (national IPR exhaustion, no parallel trade). Compulsory licensing ends up being a way for the South to reverse back to a situation close to the one with no parallel trade, restoring most of its welfare. At the same time as this complements the main results in Grossman and Lai (2008), this challenges their generality.

In the particular case where the Southern government can choose to pay a price (of its choice) higher than marginal cost when it issues compulsory licenses, results are identical to the ones under national exhaustion of IPRs (no parallel trade), except for the (small) fixed cost incurred by the South for implementing compulsory licensing. This will be the case when royalties due in accordance with the TRIPs system are effectively paid, and the South has full
discretion in establishing their amount (but can beforehand commit to that). We argue this is the present situation, where it is not clear for the Dispute Settlement Procedure (DSP) what is the “adequate remuneration in the circumstances of each case, taking into account the economic value of the authorization” the right holder should receive if a compulsory license is issued, nor the procedure itself is enforced in cases of compulsory licensing.

The most straight policy recommendation from the present paper is that governments in the developing world (the South) should first of all be conscious of their possibility of using compulsory licensing and use it as a credible threat in negotiating regulated price controls. Grossman and Lai (2008) advocated that the pharmaceutical lobby (and more broadly, the North) should move towards allowing parallel trade in a way to increase profits of innovative firms and innovation at the same time and thus welfare in the North. The South would be hurt by such decision in their setting. Its only (and legal!) remedy is the one analyzed in our paper: compulsory licensing.

Welfare in the North (and profits and innovation) is not decreased by compulsory licensing. Actually, the outcome of compulsory licensing is that the North achieves the same welfare (and profits and innovation) as when there is no parallel trade and the South is strictly worse (due to bearing the (fixed) cost of issuing licenses). However, the South through compulsory licensing recovers partially what it would have lost in the presence of parallel trade (alone) and the threat of not being served.

The indirect recommendation would be to lower the (fixed) costs associated with using compulsory licensing. As allowing parallel trade is a Northern decision (and benefits the North at the expense of the South), issuing compulsory licensing should be a sovereign Southern decision, and therefore should be facilitated and not retaliated. Any such retaliation should be punished through the WTO legal procedures.

Furthermore, as the model shows, the most efficient situation is a national exhaustion regime. If North and South could cooperatively search for the best outcome, this would be to bar compulsory license and parallel trade at a time. (Actually, a bar at parallel trade is enough, as our Proposition 1 ensures that compulsory licensing is never used in national exhaustion.) Innovation and northern welfare would be the same as in a regime with the former and the latter, and southern welfare would be increased (by an amount equal to the fixed cost wasted in issuing compulsory licenses). International drug price discrimination is shown to be a Pareto superior outcome!
Appendix A

In this Appendix A we comment the main results of the baseline model in GL.

*Condition 1* is a technical assumption on the concavity of the first-order condition of the Southern government’s maximization problem. It is assumed to hold, implying that \( m_1 < m_2 \) in Figure 1. Although necessary to the results, it is shown to hold for broad classes of demand functions commonly used: constant elasticity and linear demands.

*Proposition 1* states that the price cap regulated by the South is larger under international exhaustion than national exhaustion, for all possible ratios between market sizes.

*Lemma 1* shows that northern firms obtain higher aggregate profits when there is international exhaustion of IPRs in North (and the South sets its price control at \( p^* \)) than under national exhaustion and \( p^*_S > c \) (i.e. for \( m < m_1 \)). In other words, profits are higher when parallel trade is allowed even if the market in the South is large enough to induce the South to impose a price cap strictly above marginal cost even in the absence of parallel trade.

For \( m_1 < m < m_2 \) firms also strictly prefer the uniform pricing induced by parallel trade, as \( p^*_S > \tilde{p} \) and \( \tilde{p} \) is the price that would make them indifferent. For \( m \geq m_2 \) northern firms are indifferent between the two IPR exhaustion regimes, which is also implied by the definition of \( \tilde{p} \). This all can be seen in Figure 1.

This implies that the value of a patent under \( ie \) is no less its value under \( ne \) (and actually strictly larger for \( m < m_2 \)), which implies that innovation is larger under \( ie \) than under \( ne \), i.e. innovation is larger when parallel trade is allowed (given Condition 1 holds).

*Proposition 2* then shows that the North will prefer choosing a regime of international exhaustion to a regime where it bars parallel trade, as profits of its firms and innovation are both at least as large in such regime, and strictly larger for \( m < m_2 \). Finally, not surprisingly, *Proposition 3* shows that such decision harms welfare in the South (as compared to \( ne \)).

Grossman and Lai (2008) provide no result on overall welfare.
References


