DISCUSSION PAPERS IN ECONOMICS

Working Paper No. 05-09

A Model of Parallel Imports of Pharmaceuticals
with Endogenous Price Controls*

Katherine M. Sauer+

University of Colorado at Boulder

November 2005

Abstract
United States policymakers are considering legislation which would allow parallel imports (PIs) of brand-name pharmaceuticals from Canada. I develop a model which explores the behavior of an original manufacturer in response to a policy permitting PI competition. The model suggests that a manufacturer will limit its supply to the exporting market. When the volume of PIs is small relative to the home market, the firm will accommodate a limited volume of competition. The price in the home market is decreasing as the volume of PIs increases. When the volume of PIs is large relative to the home market, the firm will deter PI competition completely through a severe supply limit. The price in the home market will

exhaustion of patented products is to give producers an incentive to invest in the research and development of new products. In the realm of pharmaceuticals, it is argued that reducing prices (through direct price regulation or the allowance of parallel imports) would lower profits and deter manufacturers from innovating new and better drugs.[7]

Arbitrage of a good is possible when either the retail price or retail margin (difference between wholesale and retail price) varies across markets. In the classic example, international price differences stem from retail price discrimination. A manufacturer will set the retail price on the basis of the local demand elasticity. Since demand elasticity varies across markets, so do retail prices. Thus an opportunity for arbitrage arises when the trade costs between the markets are smaller than the retail price difference. However, in reality the retail price is not usually directly controlled by the manufacturer. The manufacturer must set the wholesale price at a level that will result in the desired retail price after the distributor's mark-up is added. Arbitrage is a mechanism for exploiting the difference between the wholesale price in one market and retail price in another.

In the case of pharmaceuticals, international price differences also stem from government policies.[8] All OECD governments with the exception of the United States use some form of price controls on pharmaceuticals. Policies include direct price controls, profit controls, reference pricing, restrictions on prescribing and dispensing, and annual price cuts. Manufacturers are basically prohibited from charging a market-based price.

In Canada, prescription drug prices are re

and benefit structuring. They often negotiate rebates and other discounts for their pharmacy networks.

Even with the discounts and price negotiation, the US price is still usually higher than the Canadian price. Walgreens.com is the online outlet of a popular US pharmacy and CanadaDiscountRx.com is an online Canadian pharmacy catering to US patients. Table 1 displays the average retail and wholes price ranges from 72% to 170%.[16] The potential for arbitrage at both the retail price and retail margin exists.

Parallel trade in pharmaceuticals is permitted in the European Union. This is because in the EU, IPRs are subject to “community exhaustion”. The rights to control distribution end upon first sale of the item within the EU. Thus, parallel trade is allowed between member states, but not from outside of the Union. In the EU, member states are allowed to pursue individual national health policy objectives. Each country uses some form of price controls on pharmaceuticals as part of its health care spending containment.[17] Since the EU mandates the free movement of goods between members, national price controls coupled with income differences gives rise to an opportunity for arbitrage.

It costs roughly $800 million to bring a new prescription drug to market.[18] In the United States, firms are allowed to recoup some of this expense by charging monopoly prices while the drug is under patent. Perhaps as a result, consumers in the United States enjoy more newly-launched drug choices and more drug choices overall than consumers in other markets.[19] Europe is viewed as a less attractive location for pharmaceutical research and development investment when compared to the United States.[20] If the US permits low-priced imports, the implications may extend beyond current savings on drug spending.

---

[16] These margins may be overstated due to the wholesale and retail prices being collected a year apart.
[17] See Danzon (1997) for a description of pharmaceutical price regulation across countries.
[18] See DiMasi et. al. (2003).
[19] HHS Task Force on Drug Importation (2004).
[20] European Federation of Pharmaceutical Industries and Associations (2005).

III. Prior Literature

The economic literature has advanced four theories to explain why parallel imports arise: price discrimination, national price regulation, vertical price control, and free riding on authorized distributor services.[21] In each, parallel trade distorts the market in some way. Distortions may be resources wasted due to parallel trade activities, manufacturers no longer supplying to certain markets, inefficient vertical pricing, or free rider problems.

Recall the classic example of retail price discrimination mentioned earlier. The retail price depends on the local demand elasticity and thus varies across markets. Malueg and Schwartz (1994) treat parallel trade as a mechanism for arbitraging away international price discrimination. A world with price discrimination (due to segmented markets) is contrasted with one of uniform pricing (due to the manufacturer's attempts to deter parallel trade). Depending on the degree of dispersion in demand, price discrimination may increase global welfare. When there is no dispersion in demand, welfare is the same with both discrimination and uniform pricing. As the dispersion becomes large, some of the smaller mark i s c c

a-2rke569 01 T437, vertical4

result in price differences across markets, thus creating an opportunity for arbitrage. The usual gains from trade (lower prices stemming from lower production costs) are not realized because the lower prices are a direct result of another market's artificially lowered prices. In addition, some resources are and Maskus (2004) present a theoretical model of the price-integrating impact of parallel imports and test it with data on pharmaceutical prices in the European Union. They model the actions of a manufacturer and parallel trader firms in a multi-stage game. The manufacturer produces a unique product and supplies it to the foreign market at a price exogenously set by the foreign government. The number of a foreign government over pharmaceutical prices in situations restricting and permitting parallel trade. The government's objective is to maximize foreign consumer surplus. The firm's objective is to maximize profits. Demand is assumed to be the same in each market, except for a scaling factor. The model also assumes no transportation costs so when imports are allowed, the home price falls to the foreign price. Knowing this, the manufacturer will bargain harder in the foreign price negotiations and so the uniform market price will be higher. As a result, the firm's profits will rise when parallel trade is permitted. The result holds true for both linear and constant elasticity demand. He also finds that the foreign price is decreasing in the size of the foreign market. The result that profits would increase with parallel trade is driven by the assumptions that the two markets have the same incomes, no trade costs, and that there exists unlimited potential for arbitrage.

III. Theoretical Model

Both the Ganslandt-Maskus and Pecorino frameworks are utilized to further explore the effects of parallel imports. As in Pecorino, the foreign price is endogenously determined by Nash bargaining. As in G-M, a multi-stage game is developed with positive trade costs and parallel imports arising endogenously. Additions to the models include the manufacturer's choice over the foreign supply v0iate ch15 0 TD0.000(a)-0As 33 0 3a, in manufacturers are permitted to limit supply to wholesalers.[23] Foreign demand uncertainty is incorporated because in reality, the firm only forecasts how much foreign consumers will want. If foreign demand were known with certainty, then the firm would simply always supply that amount and parallel trade would not be a threat.

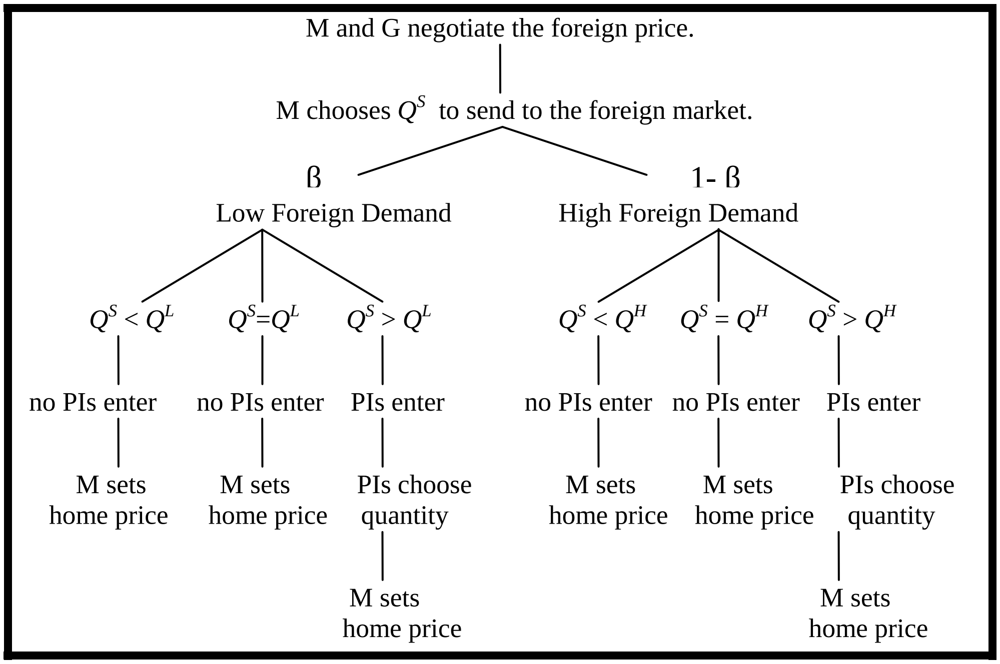
The actions of three economic actors (a home manufacturer, a foreign government, and a group of parallel traders) in two markets (a price-controlled market and a free market) are explored. Interaction between the agents takes place in a multi-stage game with the outcome found by solving backwards for a sub-game perfect Nash Equilibrium. Outcomes from situations permitting and prohibiting parallel trade are then compared.

In the first stage, the manufacturer M and foreign government G negotiate the foreign price $\bar{p}$ in a Nash bargaining game. M's objective is to maximize expected profits while G's goal is to maximize foreign consumer surplus. Then M chooses the volume Q^S to send to the foreign market by maximizing expected profits. Next, the state of foreign demand (high or low) is revealed, after which, n symmetric parallel importing firms will enter the foreign market if there is a surplus volume above foreign consumer demand. Then, each parallel importing firm simultaneously ships a quantity q from the foreign market into the home market. In the final stage, the manufacturer sets the home price p. Figure 1 illustrates the sequence of events.

"GSK Acts to Prevent Illegal, Potentially Unsafe Imports of Prescription Drugs" www.gsk.com January 21, 2003, and "Pfizer to Restrict Drug Sales to US from Canada" *Reuters* August 6, 2003.

[23] Such is the 2000 ruling from The European Court of First Instance on *Bayer AG versus Commission of the European Communities* case T-41/96.

Figure 1: Timing of Interaction



Consider a manufacturer with a pharmaceutical patent that sells the drug in two markets, home and foreign. Demand at home is

$$D_H(p) = a - bp \qquad (1)$$

Since uncertainty exists with regard to foreign demand, demand abroad is

$$Q^L = h - g\overline{p} \quad \text{with probability } \beta \qquad (2)$$

and

$$Q^H = h - y\overline{p} \quad \text{with probability } 1 - \beta \qquad (3)$$

with 0 < β < 1 and g > y. Home and foreign demand are allowed to completely differ because in reality, nations not only differ in income levels but also in prescription drug use practices.[24] The manufacturer incurs marginal cost c in production; for simplicity in analysis c is set to zero.

[24] For example, in the United States, pharmaceutical therapies are used aggressively. The focus is on the effectiveness of treatment and there is a general tolerance for side effects. In Japan, even relatively benign side effects are viewed as intolerable and treatments containing low doses of multiple drugs and herbs are common. See Burroughs (2003).

met.[25] Each parallel importing firm will simultaneously ship a quantity of the product to the home country. Each of these firms will maximize profits, taking the quantities of all other firms as given

$$Max_q \Pi^{PI} = p(Q)q - (\bar{p} + t)q - C. \tag{9}$$

This number depends on costs and home market characteristics. Multiplying (11) by (13) yields the volume desired by parallel traders as a group,

$$Q(\overline{p}) = a - 2b(\overline{p} + t) - \sqrt{2bC}. \tag{14}$$

However, the actual volume of imports is constrained by the quantity that *M* originally sends. Since the government will ensure that foreign consumer demand is met before allowing parallel trade, the maximum volume of imports is the difference between foreign supply and low demand,

$$\overline{Q}(\overline{p}) = Q^S - Q^L . \tag{15}$$

For small startup costs, the volume available $\overline{Q}(\overline{p})$ is less than the volume desired by PI firms $Q(\overline{p})$ so (15) represents the actual volume of parallel trade.[26] Assume that startup costs are small.[27] Parallel trader firms will observe this maximum volume before making their decision to enter the market and thus fewer will enter than when there was no supply restriction. Re-solving (12) with $q(n) = \overline{Q}(\overline{p})/n$ yields

$$n(\overline{p}) = \frac{\overline{Q}}{C} \quad \frac{a - \overline{Q}}{2b} - (\overline{p} + t) \quad . \tag{16}$$

Stage 2: Manufacturer's Choice of Foreign Supply

Now, *M* chooses how much to ship to the foreign market. Intuition suggests that the firm would not choose to send a volume less than Q^L because parallel trade

[26] $C < \frac{\left[a - 2b(\overline{p} + t) - Q^S + Q^L\right]^2}{2b}$

[27] According to IMS Health (2004), in March 2003, there were 99 internet pharmacies catering to US patients and by January 2004 there were 214. Since the number doubled in less than a year, it is likely that startup costs for parallel traders of pharmaceuticals would indeed be small. This assumption seems justified.

competition can not occur (due to government mandate) even if it is permitted. Thus, are similar in size or when the home market is very small.

To determine the volume of foreign supply, *M* maximizes expected profits. To make sure markets clear in times of shortage, assume that *M* incurs a penalty for undersupplying the foreign market. This penalty could plausibly be the additional transaction costs to the firm of fulfilling the shortage out of inventories. Let the penalty be equal to the shortage amount times some per unit cost *k*. That is, $P = \left(Q^H - Q^S\right)k$.

The relevant profit function over the range $Q^L < Q^S <$

parallel trade is

$$\frac{a^2}{4b} - \frac{\left(a - Q^S + Q^L\right)^2}{4b} \tag{19}$$

and the loss from having to fulfill a shortage is

$$k\left(Q^H - Q^S\right). \tag{20}$$

If the firm supplies the low amount, the expected loss is

$$E[loss]^L = (1-\beta)k(g-y)\overline{p} \tag{21}$$

and when supplying high,

$$E[loss]^H = \beta \;\; \frac{-\overline{p}^2(g-y)^2}{4b} + \frac{2a\overline{p}(g-y)}{4b} \quad . \tag{22}$$

Comparing the two expected losses, we see that (22) is less than (21) when

$$k > \frac{[2a - \overline{p}(g-y)]}{4b} . \frac{\beta}{(1-\beta)} \, . \tag{23}$$

When the marginal cost of fulfilling a shortage is large, the firm will accommodate parallel trade competition. The same is true when the volume of PIs is small relative to the home market. When the cost of fulfilling a shortage is small and when the volume of PIs is large relative to the home market, the firm would choose to strictly limit supply. The firm will always limit supply to some degree. The extent of the restriction will determine if the firm accommodates some or deters all PI competition.

Case 1: Accommodate PIs

The firm will accommodate PI competition when *k* is large or when the volume of PIs is relatively small. This means the firm will send $Q^{S*} = Q^H$. The volume of parallel trade and home price respectively are

$$\overline{Q}(\overline{p}) = Q^H - Q^L = \overline{p}(g - y) \tag{24}$$

and $$p(\overline{p}) = \frac{a - \overline{p}(g - y)}{2b} \tag{25}$$

when foreign demand is low. The home price falls with the volume of PIs. Under high demand there is no excess supply so there are no PIs and the home price is not different from the segmented market price.

Case 2: Deter PIs

The firm will choose to deter PIs through a strict supply limit when the marginal cost of fulfilling a shortage is small and the volume of PIs is large relative to the size of the home. The firm sends $Q^{S*} = Q^L$. Because of this supply limit, there is no competition from PIs and so the home price is not reduced.

Table 2 summarizes. Notice that the firm will only choose to deter parallel imports completely when it is feasible (i.e. the marginal cost of fulfilling a shortage is small) and when the home price stands to be substantially eroded by competition (i.e. home is small relative to the PI volume). Accommodating competition will be the strategy either when the cost of fulfilling a shortage is large or when the volume of PIs is relatively small.

Table 2: The Firm's Decision to Accommodate or Deter Parallel Trade

Size of the Marginal Cost to Fulfill a Shortage *(k)*	
k is large	*k is small y .TD0.001.82. 17o5.6012.0059 0 c29524 0.924 sc*

Stage 1: Foreign Price Negotiation

The government and manufacturer enter in to a Nash bargaining game to negotiate the foreign price. The government's objective is to maximize expected consumer surplus while the firm's goal is to maximize expected profits. If no agreement is reached, the firm does not sell in the foreign market so foreign consumer surplus would be zero and the firm would receive home profits only ($a^2/4b$). Thus the negotiated foreign price is the solution to

$$Max_{\bar{p}} \left[E[CS(\bar{p}^*)] - 0\right]^{\lambda} \left[E[\Pi(\bar{p}^*)] - \frac{a^2}{4b}\right]^{1-\lambda} \qquad (26)$$

where λ is the foreign government's degree of bargaining power. The (rearranged) First Order Condition is

$$\frac{\lambda E[CS(\bar{p}*)]'}{E[CS(\bar{p}*)]} = -\frac{(1-\lambda)[E[\Pi(\bar{p}*)]]'}{E[\Pi(\bar{p}*)] - \frac{a^2}{4b}} \qquad (27)$$

where the prime denotes a derivative with respect to foreign price. The left hand side of the equation is essentially the percent change in foreign consumer surplus due to a change in foreign price while the right hand side is the firm's percent change in profits. When λ=1 and the government has all of the bargaining power, the foreign price will be set as low as possible. In other words, the firm will be held to marginal cost pricing. When the firm has all of the leverage (λ=0) the foreign price will be the one which maximizes the firm's expected profits.

Benchmark: No Parallel Imports

As a benchmark, consider the case when parallel imports are not permitted. The firm will fully supply to the foreign market and will not face any competition in the home market. Thus the government's objective is to maximize

$$E[CS(\overline{p})] - 0 = \beta \frac{(h - g\overline{p})^2}{2g} + (1 - \beta)\frac{(h - y\overline{p})^2}{2y} \tag{28}$$

by choice of foreign price and the firm will maximize

$$E[\Pi(\overline{p})] - \frac{a^2}{4b} = \beta \ \ \frac{a^2}{4b} + (h - g\overline{p})\overline{p} \ \ + (1 - \beta) \ \ \frac{a^2}{4b} + (h - y\overline{p})\overline{p} \ \ - \frac{a^2}{4b}. \tag{29}$$

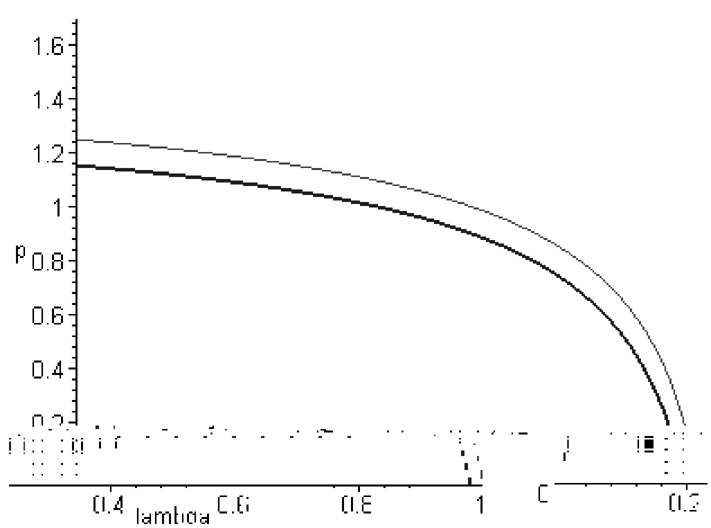
First order conditions are taken and substituted into (27). The resulting expression can

benchmark.

Case 1: PIs are Permitted and the Firm Accommodates PIs

When the potential volume of PIs is small relative to the home market or when the marginal cost of fulfilling a shortage is large, the firm will accommodate PIs. This means the firm will send a quantity equal to high foreign demand knowing that competition may result. The firm's objective is therefore to maximize

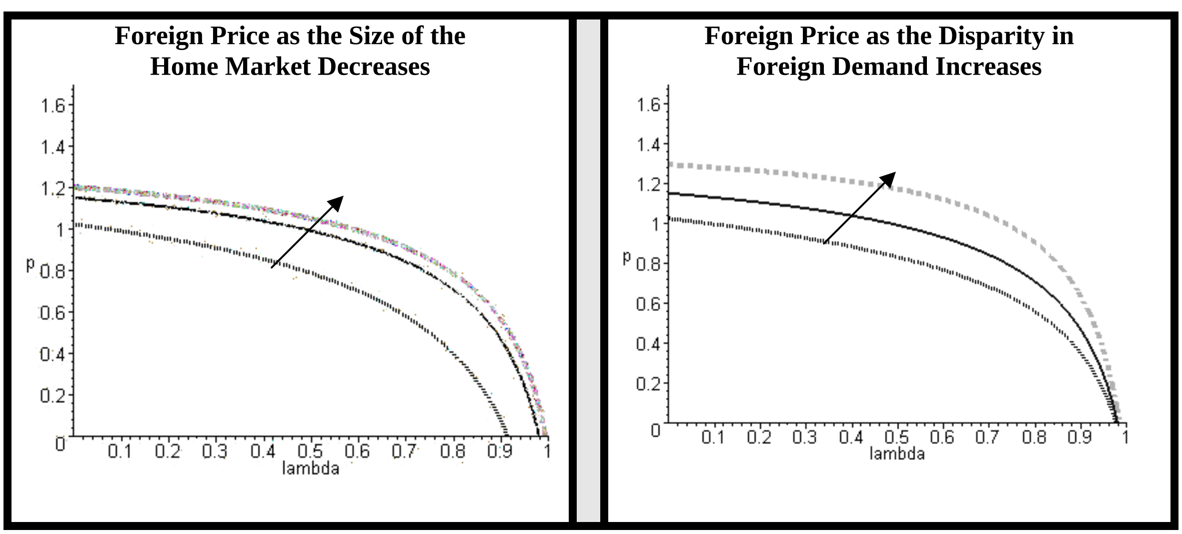
$$- \quad \_\_ \quad \frac{(\quad -(\quad))}{\quad} \; (\quad -)$$



This result seems counterintuitive. Intuition suggests that when the firm faces competition, it would bargain harder for a higher foreign price to offset losses. Such was the result from the Pecorino model. Recall from his model that the potential for competition is unlimited and the home price converges exactly to the foreign price. In such a case, it makes sense that the firm would bargain more aggressively. However in the present scenario of limited arbitrage, the home price will not converge to the foreign price. The firm would not need to bargain as aggressively since profits will not be as severely affected.

When accommodating competition, the relative size of the volume of PIs to the home market is key. As the size of the home market decreases relative to the volume of PIs, the firm will bargain harder for a higher price to offset its losses. Similarly as the dispersion in foreign demand increases, so does the potential volume of PIs. The firm will bargain for a higher foreign price to offset the reduction in home price from competition. Figure 6 illustrates.

Figure 6: Foreign Price as the Relative Volume of PIs Increases



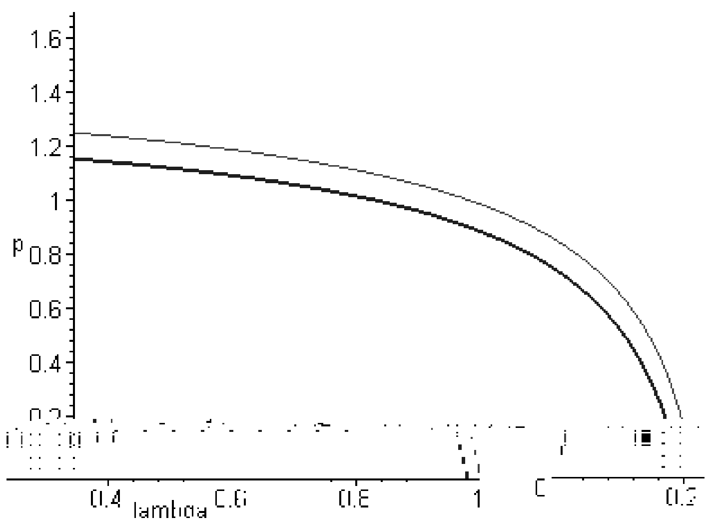
Parameter values are h=10, g=5, y=3, =0.5, and b=5 as the home market size "a" is set to 20, 10, and 8.

Parameter values are h=10, g=5, =0.5, b=5 a=10 as "y" is set to 4.5, 3, and 2.

Case 2: PIs are permitted but the firm deters PIs

When the potential volume of PIs is large relative to the home market and when the marginal cost of fulfilling a shortage is small, the firm will choose to limit supply to the level of low foreign demand. In doing so, the firm makes parallel trade impossible. While the firm will not suffer losses from competition, it will face the cost of fulfilling a shortage with probability 1 – . The firm's objective is therefore to maximize

E0.500p3 0 TD(g)T8033 0375 gTD(b)Tj42.2513 25.6 m398 5395 144k g y p $\frac{a}{b}$ $\frac{a}{b}$ $\frac{a}{b}$ $\frac{a}{b}$



In addition, as *k* increases the firm would prefer a lower price to offset the rising marginal cost of fulfilling a shortage. Figure 8 illustrates that the foreign price decreases as *k* increases.

Figure 8:
Foreign Price as the Marginal Cost of Fulfilling a Shortage Increases

What about the home price? When the firm deters PIs through a strict supply limit, no parallel trade occurs in equilibrium so the home price is unaffected. When the [redacted]

Expected profits are much lower when the firm chooses to accommodate PIs. This is because the firm not only receives a lower price in the foreign market, but also receives a lower price at home when competition occurs. When the firm deters competition, profits are lower due to the lower foreign price as well as the marginal cost of fulfilling a shortage. However, since the firm will choose to deter only when the marginal cost of fulfilling a shortage is small, it follows that the resulting reduction in profits would be small.

These results again contradict Pecorino's findings. Recall that in his model the firm bargains harder for a higher foreign price when PIs are permitted. Subsequently, he finds that profits will rise when the PIs ar

US consumers may not experience lower prices as such a policy would intend.

V. Conclusion

In this paper I developed a model of parallel imports with endogenous price controls. A monopoly manufacturer will choose to limit its foreign supply to lessen the quantity available for parallel trade competition. The relative size of the home market to potential volume of imports emerges as major determinant in a manufacturer's choice of how strictly to limit supply. When the potential volume of PIs is relatively small, the manufacturer will accommodate parallel trade by sending a volume equal to high foreign demand. When the potential volume of PIs is relatively large, the manufacturer will deter competition by means of a strict supply limit. Home consumers can enjoy lower prices only when the manufacturer accommodates parallel trade. The manufacturer will have reduced profits when PIs are permitted. For situations similar to the US-Canada scenario, the model predicts that the manufacturer will accommodate parallel trade.

VI. References

Chard, J.S. and C.J. Mellor (1989). "Intellectual Property Rights and Parallel Imports". *World Economy*. 12: 69-83.

Burroughs, Valentine J. (2003). "The Importance of Individualizing Prescribing Among Genetically and Culturally Diverse Groups". *Group Practice Journal.* 52.

Danzon, Patricia (1997). Price Regulation for Pharmaceuticals: Global vs. National Interests. Washington, DC: The American Enterprise Institute Press.

Danzon, Patricia (1999). *"Price Comparisons for Pharmaceuticals: A Review of US and Cross National Studies"* The Wharton School. University of Pennsylvania. http://knowledge.wharton.upenn.edu/PDFs/154.pdf.

United States Department of Commerce- International Trade Administration (2004) *"Pharmaceutical Price Controls in OECD Countries"*.

Vernon, John A. (2005). "Examining the Link Between Price Regulation, Importation,