Direct and indirect savings from parallel imports in Sweden*

David Granlund, Department of Economics, Umeå University, Sweden, david.granlund@umu.se

Abstract

Objective

The aim was: i) to quantify the direct and indirect savings from parallel imports in Sweden during a period when sellers were forbidden from giving discounts to pharmacies, and ii) to study if the effects of competition from parallel imports on list prices became smaller in absolute size when sellers were allowed to give discounts to pharmacies.

Methods

I analyzed the monthly prices for 3,068 products during 61 months when discounts were forbidden and for 2,059 products during 84 months when discounts were allowed. The price effects were estimated using dynamic models that rendered lagged numbers of competitors into valid and strong instruments for the current values.

Results

When discounts were forbidden, parallel imports had a market share of 16% and were on average 9% cheaper than locally sourced drugs, which yielded a direct saving of Swedish krona (SEK) 231 million per year. Also, parallel imports reduced the prices of products with the same substance by, on average, 6% in the long-term, which yielded indirect savings of SEK 421 million per year. In total, parallel imports reduced the cost for on-patent pharmaceuticals by 4%. When discounts were allowed, the average gap in list price between parallel imports and locally sourced products was reduced to 0.3%, and the list prices of locally sourced products were no longer significantly affected by competition from parallel imports.

Conclusion

When discounts were allowed, the savings of parallel imports through lower list prices were replaced by savings of pharmacies through secret discounts.

* I gratefully acknowledge a research grant from the Swedish Competition Authority [grant number 382/2018]. I am also grateful to IMS Sweden and Västerbotten County Council for supplying the data used in this article.

1. Introduction

In an attempt to practice third-degree price discrimination, producers might charge wholesalers in low-income countries less than they charge wholesalers in high-income countries. Parallel traders take advantage of these price differences by buying products intended for low-price countries and, without authorization from the patent holder, selling them to wholesalers in high-price countries. Parallel trade is allowed within the European Economic Area to help fulfill the objective of creating a single market.

This paper evaluates the savings from parallel imports in Sweden during a period when sellers of pharmaceuticals were forbidden from giving discounts to pharmacies. This implies that the official list prices were actual transaction prices, which enabled me to quantify the total savings. I also studied if the effects of competition from parallel imports on list prices became smaller in absolute size when sellers were allowed to give discounts to pharmacies.

With access to data on market shares for parallel imports and relative prices, direct savings can be easily calculated. This has been done previously, e.g., by West and Mahon [1], who reported direct savings of Swedish krona (SEK) 400 million per year for Sweden (measured in retail prices).

West and Mahon [1] also showed price plots and comparisons of price changes over 5–6 years, which indicated that parallel imports put downward pressure on prices. Estimating this effect is, however, difficult because parallel imports are more likely to be sold the higher the prices of the locally sourced products are, rendering the variable endogenous. To address this endogeneity problem, Ganslandt and Maskus [2] and Granlund and Köksal-Ayhan [3, 4] used exchange rates and the age of drugs as instruments for competition from parallel imports, and reported point estimates suggesting that competition from parallel imports reduced the prices of locally sourced drugs in Sweden by 12% to 21%. However, these instruments may affect the prices of locally sourced drugs in other ways than through the existence of parallel imports, which can create bias.

Vandoros and Kanavos [5] instead used instruments based on the number of policies promoting parallel imports and the distance between the source countries and the four destination countries they analyzed (Germany, Sweden, the Netherlands, the United Kingdom). They found no statistically significant price effect, but because of large standard errors, they could neither reject

the premise that the price effect was large.[1] Vandoros and Kanavos also analyzed the effect of the market share of parallel imports—following Kanavos and Costa-Font [6] and Kanavos and Vandoros [7]—but like the previous studies, they found no statistically significant price effects. The discussant to Kanavos and Costa-Font, Christian Gollier, mentioned as one potential explanation for the lack of significant estimate the possibility that "the local manufacturer actually matches the price of the importers by using hidden discounts to distributors rather than reducing the list price" [6, p. 793].

To overcome the problem with weak and potentially endogenous instruments, Granlund [8] used a dynamic model that allowed lags of competition variables to be used as instruments for their current values. This approach yielded sufficiently many strong instruments for also studying the causal effects on the intensive margins, i.e., how the number of parallel traders and the number of therapeutic competitors affect prices. Granlund [8] used part of the data used in the present paper: that for tablets and capsules sold in October 2002–October 2007. For this paper, I estimated similar price functions as Granlund [8], but also did so for the period of January 2011–December 2017, for all forms of administrations, and calculated the direct and indirect savings yielded by parallel imports.

2. Rules regarding parallel imports

All Swedish residents are covered by a mandatory and uniform pharmaceutical benefit scheme. Since October 2002, a substitution legislation requires pharmacy personnel to inform consumers if cheaper substitute products are available, unless the prescriber has vetoed substitution or if the pharmacist has reasons to believe that the patient would be adversely affected, e.g., because the low-cost alternative has a package that the patient would find difficult to open. For parallel imports, available substitutes are defined as those in stock at the pharmacy in question [9]. If consumers oppose the substitution, or choose to switch to a substitute other than the cheapest one available, they will be charged the entire incremental cost. The Swedish Medical Products Agency defines a

[1] Vandoros and Kanavos [5] did not discuss the magnitude of their estimates, but their point estimates for the indicator variable *pt* suggest that the presence of parallel imports reduced the prices of locally sourced product by 81% and 38%, respectively. These figures are calculated using the formula 100*[exp(β)-1].

product as a substitute if it has the same active substance, strength, and form of administration (e.g., pills or oral fluid) and nearly identical package size.[2]

Pharmaceutical producers and parallel traders are free to set their own prices, but to be included in the pharmaceutical benefits scheme, they must submit their prices to the Pharmaceutical Benefits Agency (PBA) 2 months prior. The PBA approves prices not exceeding the highest existing price of exchangeable products, which implies that parallel imports are allowed to be as highly priced as locally sourced products [11, 12]. Before July 2009, producers and parallel importers were not allowed to offer their products below these prices. That is, they were not allowed to give discounts to pharmacies.

3. Methods

3.1 Data

The study was based on two panel datasets obtained by merging datasets of pharmaceutical sales, compiled by IMS Sweden, with datasets provided by the Västerbotten county council, containing detailed information of each pharmaceutical product. An observation in the datasets represents a product with a certain active ingredient, strength, administrative form, and package size, supplied by a certain firm and sold in a certain month. The datasets cover all prescription drugs sold in Sweden during October 2002–October 2007 and during January 2011–December 2017. Data from November 2007–June 2009 were not used because prices during this period could have been by anticipation of the possibilities of giving discounts. Neither did I use data from July 2009–December 2010, as the business models related to discounts might still have been under rapid development under this period. Lacking information on patent expiration, I defined pharmaceuticals as off-patent starting the first time any generics with the same active ingredient were sold in Sweden. After excluding off-patent pharmaceuticals, the first dataset contained 132,008 observations of locally sourced product and 31,999 observations of parallel imported

[2] Packet sizes are allowed to vary slightly for non-narcotic groups; for example, substitution can be made from a 30-pill package to a package in the 28–32-pill range. For the 3% of observations classified as narcotic drugs, exactly the same packet size is required. Parallel imports usually have exactly the same package size as the locally sourced product. For locally sourced products that are exchangeable to at least on parallel imports, all parallel imported substitutes had exactly the same package size in 85% of cases in the first dataset and 79% in the second dataset.

products. During the second study period, the World Health Organization (WHO) had changed the Anatomical Therapeutic Chemical (ATC) code for many products. To prevent this affecting the controls for therapeutic competition, I excluded one-fourth of the products that occasionally during the study period belonged to a five-digit ATC group affected by these changes, resulting in 89,002 observations of locally sourced products and 55,748 observations of parallel imported products.[3]

3.2 Estimation of price effects and descriptive statistics

For several reasons, prices are not expected to adjust instantaneously to the new long-term equilibrium when market conditions change. One reason is possible price coordination between therapeutic alternatives, which can make companies limit price changes to reduce the risk of triggering price wars [12]. Another is the dynamic price cap on drugs in Sweden, which means that a drug whose price is raised so that it becomes more expensive than the most expensive substitute can be excluded from the pharmaceutical benefit scheme. A company that is uncertain about what the new optimal price is after it has received competition may, because of this price cap, find it wise to lower the price gradually, rather than to lower it more directly and then risk not being able to adjust the price if it is found that the price cut was unnecessarily large. For these reasons, I estimated dynamic models.

The preferred specification, which was estimated with two-stage least squares, is written as:

$$lnP_{it} = \theta lnP_{i,t-1} + \beta_1 D\_PiSubstance_{st} + \beta_2 D\_PiE_{it} + \beta_3 lnN\_PiSubstance_{st} + \beta_4 lnN\_PiE_{it} + \beta_5 D\_Th_{st} + \beta_6 D\_ThGen_{st} + \beta_7 lnN\_Th_{it} + \beta_8 lnN\_ThGen_{st} + \eta_t + \mu_i + \varepsilon_{it},$$

in which indices *i*, *s*, and *t* represent product, substance, and time in months, respectively. The dependent variable lnP_{it} is the natural logarithm of the wholesale price of the on-patent locally sourced product *i* in month *t*. The first lag of this variable, $lnP_{i,t-1}$, was included as an explanatory variable to make the model dynamic.

[3] Also, the WHO made changes during 2005–2007 WHO, but these only affected 7% of the observations, and the regression results are robust towards excluding these. As I wanted to calculate savings for the entire market for the first study period, these observations are retained in the analyses presented in this paper.

The variable $D\_PiSubstance_{st}$ is an indicator that takes the value of 1 if one or more parallel imported products with the same substance as product *i* were sold in Sweden in month *t*. D_PiE_{it} is also an indicator, but it only takes the value of 1 if at least one parallel imported product that is exchangeable with product *i* was sold in Sweden in month *t*. In accordance with the substitution rules, an exchangeable product was defined as a drug with the same active substance, form of administration, strength, and nearly identical package size.

The variable $lnN\_PiSubstance_{st}$ was defined as the natural logarithm of the number of parallel traders selling products with the same substance when this variable is strictly positive, and takes the value of 0 otherwise. The variable lnN_PiE_{it} has the corresponding definition for the number of parallel traders selling exchangeable products.[4] These definitions imply that, for example, the coefficient for $D\_PiSubstance_{st}$ describes the effect of the first parallel trader, whereas the coefficient for $lnN\_PiSubstance_{st}$ captures the effect of variations in strictly positive numbers of parallel traders.

The variable D_Th_{st} takes the value of 1 if at least one other firm sold a locally sourced product with the same five-digit ATC code in month *t*. If there was a generic version of at least one of these substances, also D_ThGen_{st} takes the value of 1. The variable N_Th_{it} (not included in the specification) was defined as the number of pharmaceutical substances with the same five-digit ATC code and with locally sourced drugs sold by other firms than firm *i* during month *t*. lnN_Th_{it} is the natural logarithm of N_Th_{it} for strictly positive values of this variable, and is 0 otherwise. Lastly, lnN_ThGen was defined as the natural logarithm of the number of therapeutic alternatives for which generic versions exist when this variable is strictly positive, and takes the value of 0 otherwise.

The eight competition variables were all instrumented with their first lags and with $lnQ_{s,t-3}$, which is the natural logarithm of the quantity sold of substance *s* in month *t*.[5] Producers have good information about the values of these instruments when, at the end of $t-2$, they set their prices

[4] In the first dataset, 382 observations were not used in the regressions, as lnN_PiE_{it} could not be defined because of missing information on package size for at least one product with the same substance, strength, and form of administration.

[5] The first dataset included information on the defined daily dozes for 83% of observations, which was then used to create the quantity variable. For the remaining observations, I used the sum over products of number of pills sold of each product multiplied by the strength of each pill. This yielded values that are proportional to daily doses, which is all that is needed for using variations in lnQ_{st} over time as a proxy for market growth.

for month *t*. For the first eight instruments, the reason is that the prices of all products that can be sold within the benefit scheme in month $t-1$ are announced in the first half of month $t-2$. Hence, producers can observe how many potential competitors they will have in month $t-1$ and can, based on this, predict the competition they will face in month *t*. Regarding $lnQ_{s,t-3}$, IMS/IQVIA had delivered sales data for month $t-3$ to its customers when prices from month *t* are set.

Lastly, month and product fixed effects (η_t and μ_i) were included in the specification, and the error terms were allowed to be correlated within substances. Descriptive statistics are presented in Table 1.

Table 1. Descriptive statistics.

	Jan. 2003–Oct. 2007		Jan. 2011–Dec. 2017			
	Mean	SD	Mean	SD	Min	Max
P_{it}	1462.75	4594.54	3608.85	11676.97	6.31	290,670.50
$D\_PiSubstance_{st}$	0.24	0.43	0.45	0.50	0	1
D_PiE_{it}	0.11	0.32	0.24	0.43	0	1
$N\_PiSubstance_{st}$	0.62	1.36	1.64	2.13	0	9
N_PiE_{it}	0.24	0.82	0.59	1.31	0	8
D_Th_{st}	0.82	0.39	0.82	0.38	0	1
D_ThGen_{st}	0.53	0.50	0.57	0.49	0	1
N_Th_{it}	2.93	2.50	3.31	4.28	0	28
N_ThGen_{st}	0.89	1.15	1.19	1.44	0	7
Q_{st} (in millions)	12.91	66.71	9.34	44.40	0.00	4500.00

Note: The number of observations is 132,008 for the first dataset and 89,002 for the second dataset.

4. Regression results

The estimation results for the preferred specification are presented in Table 2, while model checks and robustness analyses are presented in the Appendix. The results for the lag of the dependent variable ($lnP_{i,t-1}$) show that that prices reacted slowly to changes in competition. Taking one minus the coefficient for $lnP_{i,t-1}$ and multiplying with 100 reveals that only 4% of the long-term effects were realized immediately in the samples.

The coefficients for the eight competition variables show their short-term effects, whereas dividing them by one minus the coefficient for $lnP_{i,t-1}$ yields the long-term effects. To obtain the exact effect in percentage terms, the formula $100 * [\exp(B) - 1]$ should be applied, in which *B* is the coefficient estimate, or long-term estimate, of interest.

For the first study period, the estimates for $D\_PiSubstance_{st}$ and $lnP_{i,t-1}$ show that the first parallel trader selling products with the same active substance, but which were not exchangeable with product *i*, reduced the price of product *i* by 0.17% in the short-term and 3.9% [≈0.17/(1 - 0.9568)] in the long-term. If instead the parallel trader sold an exchangeable product, so that D_PiE_{it} also equaled 1, the price fell by an additional 2.7% in the long-term. Additional parallel importers only reduced the price if they sold exchangeable products, but in this case the effect was small as well; if the sellers of exchangeable product increased from one to three, the price was reduced by 2.2% in the long-term.

The differential $dlnP_i^*/dD\_PiSubstance_{st}^*$ shows the weighted average long-term effect of facing competition from at least one parallel importer selling the same substance.[6] Applying the formula $100 * [\exp(-0.0601) - 1]$, the effect equaled a price reduction by 5.83% for the first study period. As comparison, the raw (unweighted) average price reduction equaled 5.47%. The results are similar to those reported for tablets and capsules in [8]. For example, the long-term effect of a first parallel trader that also sold exchangeable products was estimated to be -7.0% in [8], whereas here it was -6.5%.

Column 3 of Table 2 shows that competition form parallel imports had no effect on the list prices of locally sourced products in the study period when discounts were allowed. Also, the weighted average effect of facing competition from at least one parallel importer selling the same substance was at the 5% level significantly smaller in the second study period compared with the first study period.

Figure 1 shows that similar results were obtained when indicator variables were used for the numbers of parallel importers instead.

[6] The differential $dlnP_i^*/dD\_PiSubstance_{st}^*$ was defined as $(\beta_1 + m_2\beta_2 + m_3\beta_3 + m_4\beta_4)/(1 - \theta)$, in which m_2, m_3, and m_4 are the within-sample weighted means of $PiCompClose_{it}$, $lnPiFirmsS_{st}$, and $lnPiFirmsClose_{st}$, respectively, when $D\_PiSubstance_{st} = 1$.

Table 2. Estimation results for lnP_{it}.

	Discounts forbidden Oct. 2002–Oct. 2007	*Discounts allowed* Jan. 2011–Dec. 2017
$lnP_{i,t-1}$	0.9568***	0.9573***
	(0.0055)	(0.0066)
$D\_PiSubstance_{st}$	-0.0017***	-0.0007
	(0.0006)	(0.0007)
D_PiE_{it}	-0.0012*	-0.0006
	(0.0007)	(0.0006)
$lnN\_PiSubstance_{st}$	0.0001	-0.0002
	(0.0004)	(0.0006)
lnN_PiE_{it}	-0.0010*	0.0001
	(0.0006)	(0.0007)
D_Th_{st}	-0.0001	0.0003
	(0.0008)	(0.0015)
D_ThGen_{st}	-0.0007	0.0012
	(0.0006)	(0.0009)
lnN_Th_{it}	-0.0012	-0.0044**
	(0.0012)	(0.0021)
lnN_ThGen_{st}	0.0010	-0.0043**
	(0.0007)	(0.0018)
$dlnP_i^*/dD\_PiSubstance_{st}^*$	-0.0601***	-0.0237
	(0.0151)	(0.0154)
Observations	119,945	79,659
R^2	0.9183	0.9130
K-P rk LM	72.9704	57.5830
K-P rk LM, p-value	0.0000	0.0000
Hansen J, p-value	0.1293	0.3320

Note: The specifications include product-specific fixed effects and 58 and 81 indicator variables for months, respectively. In the first-stage regressions, data from October 2002–October 2007 and Jan. 2011–Dec. 2017 were used. K-P rk LM refers to the Kleibergen-Paap rk LM statistic, which indicates the strength of the instruments. The null hypothesis in the K-P test is that the model is under-identified. The null hypothesis for the Hansen J test is that the instruments are valid, i.e., uncorrelated with the error term. Standard errors, robust to correlations within substances, are given in parentheses. ***, **, and * indicate that the coefficient is statistically significant different from zero on the 1%, 5% and 10% significance levels, respectively. The estimation results for the indicator variables for months and for the first-stage regression are available on request.

Regarding the results for therapeutic competition, there were no statistically significant price effects in the first study period. For the second study period, the estimates for D_Th_{st} and D_ThGen_{st} imply that there was no significant price effect of a first therapeutic competitor, but the estimates for lnN_Th_{it} and lnN_ThGen_{st} indicate that prices fell with additional competitors.

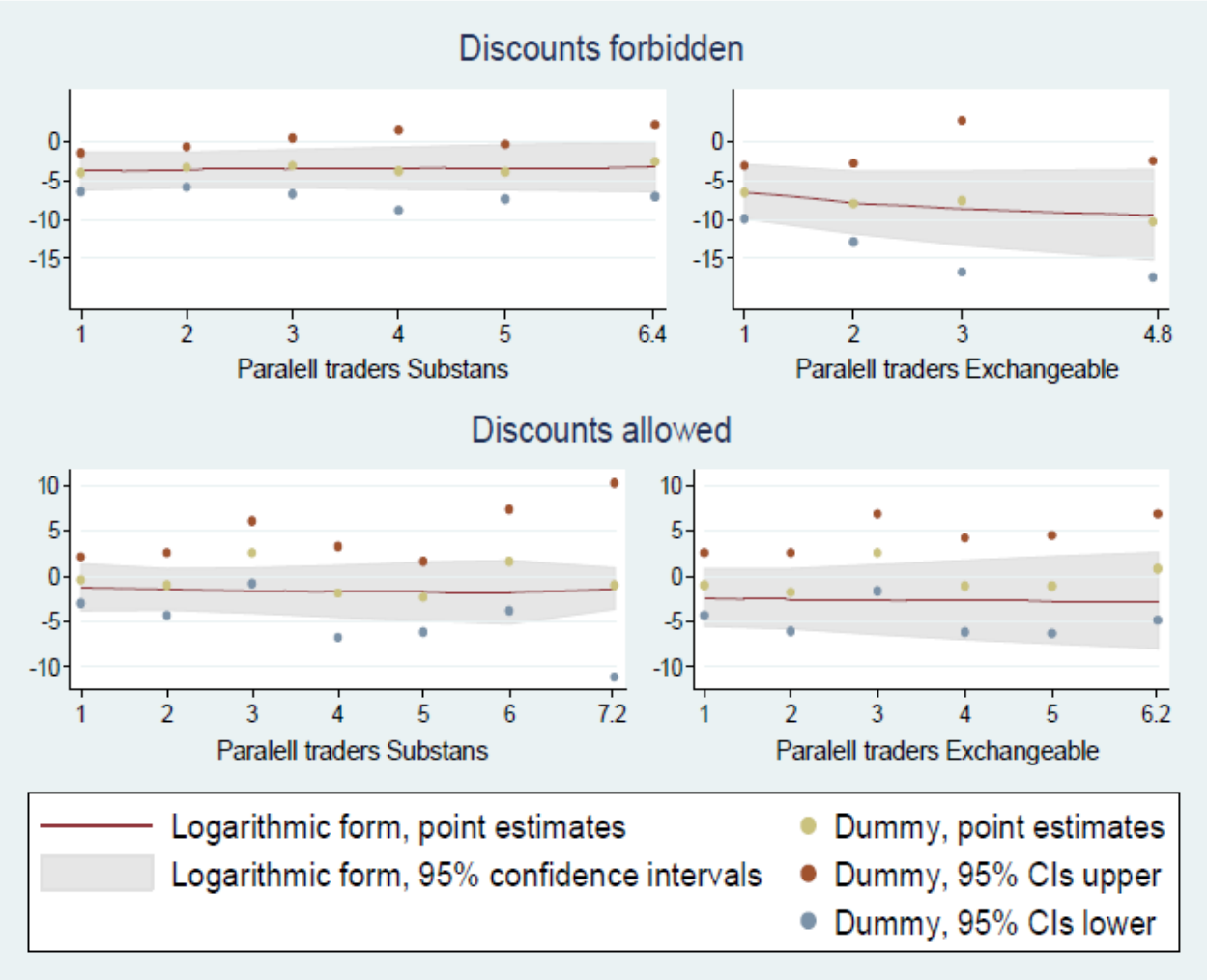

Figure 1. Estimated long-term price effects in percentages of the number of parallel traders selling products with the same active substance and exchangeable products, respectively; comparison of logarithmic-form and flexible-form estimates. The effects in the left panels are plotted holding N_PiE_{it} at zero, while the effects in the right panels are plotted holding $N\_PiSubstance_{st}$ equal to N_PiE_{it}. The smooth lines are the long-term effects predicted from the preferred specification of $D\_PiSubstance_{st}$ and $lnN\_PiSubstance_{st}$ (left panels) and of $D\_PiSubstance_{st}$, D_PiE_{it}, $lnN\_PiSubstance_{st}$, and lnN_PiE_{it} (right panels). The gray area shows the associated 95% confidence intervals. PiDummy shows the long-term effects of indicator variables for the numbers of $N\_PiSubstance_{st}$ (left panels) and for the numbers of $N\_PiSubstance_{st}$ and N_PiE_{it} (right panels), and PiDummyCL and PiDummyCU show the lower and upper bounds of the associated 95% confidence intervals. These estimates come from an IV specification including indicator variables for the numbers of parallel importers. However, groups with few observations were grouped together to avoid indicators that take the value of 1 for less than 1% of the observations. The estimates for these merged groups are plotted at the average value of $N\_PiSubstance_{st}$ and N_PiE_{it} in each merged group, respectively.

5. Total savings of parallel imports when discounts were forbidden

The pharmaceutical costs in the absence of parallel imports were calculated by multiplying the sold quantity of all products—locally sourced as well as parallel imports—with the price that locally sourced products would have had in the absence of parallel imports.[7] The saving is the difference between this cost and the actual cost. The saving was calculated using data from January 2003–October 2007, i.e., for the period used in the second stage of the IV regressions, and was divided into one direct and two indirect parts; one for locally sourced products and one for parallel imported products.

The direct saving consists of the sum over all parallel imported products of the number of packages sold multiplied by the price difference between these and their locally sourced counterparts. Parallel imports were on average 9% cheaper than locally sourced products, yielding an average annual direct saving of SEK 231 million.[8] After discounts were allowed, the gap in list prices between parallel imports and locally sourced drugs were on average only 0.3%.[9]

The indirect savings for locally sourced products were calculated as the total sales value of locally sourced products for which $D\_PiSubstance_{st} = 1$, multiplied by 0.0619, which shows in decimal form how much more expensive the products were estimated to have been if they had not faced competition from parallel imports.[10] This saving was estimated to average SEK 260 million per year. The indirect savings for parallel imported products were calculated correspondingly, except that I used the sales values that would have existed if these products were sold at the same price as

[7] This means that the total quantity (including both parallel imports and locally sourced drugs) are assumed to be unaffected by the price. If this assumption does not hold, the saving should only be interpreted as an estimate of how much more the pharmaceutical quantities sold would have cost without parallel imports. If instead a price elasticity of 0.2% is assumed (as in [13]), parallel imports would have been estimated to have reduced the expenditures by 80% of the figures reported in this paper, but in this case parallel imports would also have resulted in consumer surplus from additional quantities used. As reported by Kanavos and Costa-Font [6], estimates on the demand elasticity for prescription pharmaceuticals range from close to 0 to -0.33.

[8] For 17% of the observations of parallel imported products, locally sourced products with the same active substance, form of administration, strength, and package size were not available in the same month. For these observations, the relative price has been assumed to equal the weighted average relative price of parallel imported products in this month. As weight, I used the product of the number of packages sold of the parallel imported product and the price of the locally sourced product, i.e., the sales values the parallel imports would have had if they had been sold at the same price as the locally sourced products.

[9] Allowing for correlation within substances, the standard error for the price difference is only 0.1%, implying that 0.3% is statistically significantly different from zero.

[10] The value 0.0619 is calculated as C/(1 - C), in which C = 0.0583 is the weighted average long-term reduction in prices resulting from competition from at least one parallel imported product, reported in the previous section. The formula reflects that a price cut of 5.83% must be followed by a price increase of 6.19% to be fully off-set.

the locally sourced products. This yielded an estimated average indirect savings for the parallel imports of SEK 161 million per year.

The savings are illustrated in Figure 2 and summarized in Table 3. The rectangle of Figure 2 illustrates how large the annual expenditure on patent prescription pharmaceutical was estimated to have been without competition from parallel imports. For locally sourced products that did not face competition from parallel imports, no savings arose. The savings for the other two categories are illustrated in the upper right corner of the figure. All in all, the estimated annual savings generated by parallel imports of pharmaceuticals before discounts were allowed were SEK 652 million (≈ 260 + 161 + 231). This amounts to 4% of the SEK 16,619 million that on-patent prescription pharmaceuticals were predicted to have cost without parallel imports.

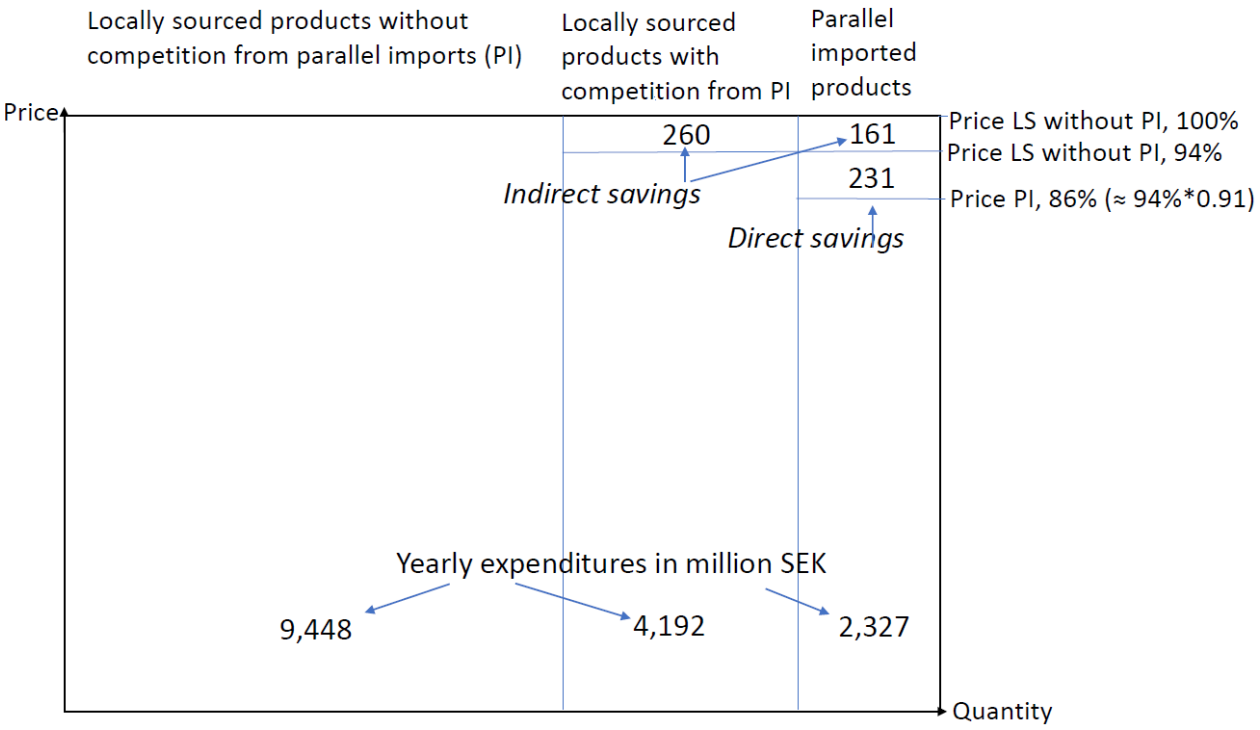

Figure 2. Illustration of average yearly savings for January 2003–October 2007. LS refers to locally sourced products; PI refers to parallel imported products. The amounts are measured in million SEK of pharmacies' purchase prices and are expressed in year 2017 prices.

In Table 3, Column 2 reports the point estimates, and Column 3 reports the 95% confidence intervals (95% CI) that only reflect uncertainty in the estimated price effects of competition from parallel imports. The last column reports the CIs from a probabilistic sensitivity analysis (PSA)

that also accounted for variability in competition from parallel imports, market shares, and relative prices. The CIs are further described and discussed in the Appendix.

Table 3. Predicted savings with CIs, in millions SEK.

	Point estimate	95% CI estimation uncertainty	95% CI PSA
Direct savings	231		113–375
Indirect LS	260	130–394	134–424
Indirect PI	161	81–245	81–271
Indirect LS+PI	421	211–638	223–679
Direct+indirect PI	393	312–476	238–577
Total savings	652	442–869	391–968

Note: The asymmetry in the CIs in Column 3 is explained by the concavity of C/(100 - C) (described in footnote 10), which is only partly offset by the convexity of C = 100*[exp(B)-1]. For the PSA CIs, the randomness of the Monte Carlo simulations is also a source of asymmetry.

6. Discussion and conclusions

When discounts were forbidden, parallel imports had a market share of 16% and were on average 9% cheaper than locally sourced drugs, which directly reduced the cost for on-patent pharmaceuticals by 1.4%. Parallel imports also led to price competition, and in total, parallel imports reduced the cost of on-patent pharmaceuticals by 4%. When discounts were allowed, parallel imports raised their list prices to nearly the same level as locally sourced products, and the significant effect of competition from parallel imports vanished. The results highlight the importance of using actual transaction prices when studying the savings caused by parallel imports.

A likely explanation for why no significant price effect of competition from parallel imports was observed when discounts were allowed is that producers then chose to use discounts, rather than lower list prices, to incentivize pharmacies to dispense locally sourced products. The advantages of using discounts includes that they can be quickly reverted and do not affect the maximum prices producers are allowed to charge in countries that use the Swedish price as an external reference price.

When discounts were allowed, the savings of parallel imports through lower list prices were replaced by savings of pharmacies through secret discounts. The discounts given by parallel traders

have been estimated to be about SEK 470 million per year,[11] and this is within the CI of the savings in the pre-reform period caused by parallel imports having lower prices than their locally sourced counterparts would have had if their prices were not lowered due to competition from parallel imports (point estimate, SEK 393 million). No estimates exist regarding the discounts given by sellers of locally sourced products, but it is conceivable that allowing discounts had small effects on the total savings caused by parallel imports, but instead of the savings being directly acquired by consumers and the pharmaceutical benefit scheme, they now go to the pharmacies.

Appendix

Discussion about the confidence intervals reported in Table 3

Table 3 in the paper lists the point estimates of the savings together with 95% confidence intervals. The confidence intervals that only reflect the uncertainty in the estimated price effects of facing competition from parallel imports are relevant if one knows: i) the sales values for locally sourced products facing competition from parallel imports, ii) the extent of the competition these products face, and iii) market shares and relative prices of parallel imports. However, beforehand also these variables are unknown because they depend on, among else, the decision of parallel traders and prescribers, pharmacies policies and patient preferences. Therefore, Column 4 presents confidence intervals from a probabilistic sensitivity analysis (PSA) that also accounts for these sources of uncertainty.

The PSA was performed by making 10,000 independent draws from the distribution of the estimates for the lag of the dependent variable and for the variables describing the extent of competition from parallel imports, and of one of the 61 months in the data. For each month drawn and the following eleven months, direct savings, averages values (weighted with sales) of the four variables describing competition from parallel imports, and markets share of parallel imports were calculated. Together with the draws from the distribution of the estimates, these were used to calculate estimates of yearly savings. To equalize the expected number of times data from each

[11] Own calculations based on [14-16].

month were used, I treated time in a circular manner meaning that, e.g., if the 61st month was drawn, also data from the first 11 month were used to calculate 12-months values.

Column 3 of Table 3 shows the additive property of the confidence intervals that only accounts for the estimation uncertainty. For example, the sum of the confidence interval for *Indirect LS* and *Indirect PI* equals the confidence interval for *Indirect LS+PI*, except from rounding effects, and the widths of the confidence intervals for *Total Savings* and *Indirect LS+PI* are equal because *Direct savings* is not affected by the estimation uncertainty. This property does not hold for the confidence intervals from the probabilistic sensitivity analysis. The reason is that uncertainties in some variables canceled out at least partially. For example, variation in the market shares of parallel imports affects the PSA confidence intervals for both *Indirect LS* and *Indirect PI*, but do not affect *Indirect LS+PI*.

Model checks and robustness analyses for the price effects of facing competition from parallel imports

Because non-stationarity can result in spurious results, I tested for this using the ADF version of a Fisher-type test, as implemented in xtunitroot in STATA. This test allows for unbalanced panels with gaps within panels and rejected non-stationarity of lnP_{it} on the 1% level for both samples. I included panel means and time trends and tested for non-stationarity both with one lag and without lags.

The instruments used in the two-stage least squares estimations were made valid by including the lag of the dependent variable. This lag controls for previous price changes, which makes the error term dependent only on the current price shock and price shocks are hard to predict. However, if the error-terms are serially correlated, firms might be able to predict them. I have therefore tested for serial correlation up to the third order using the test proposed by Cumby and Huizinga [17]. The null hypothesis of no serial correlation of the second order was rejected on the five percent significance level for both estimations presented in the paper, and for the second specification also the hypothesis of no third-order serial correlation was rejected. Still, the Hansen J test (reported in Table 1 in the article) does not reject that the instruments are valid, suggesting the serial correlation of the second or higher order must be small. The first lags of the competition variables would be

invalid instrument if firms, when they in month $t-3$ decides whether or not to have a price for month $t-1$, can predict the error term for month t. Because prices for month $t-2$ is announced in month $t-3$, firms can partly do this if they have information about serial correlation of the second, or higher, order, but serial correlation of the first order is not problematic in this respect. Estimations of a generalized linear model accounting for serial correlation up to order three confirm that the serial correlations of order two and three are small even though three of them are statistically different from zero. In the first (second) dataset the estimated correlations are 0.03 (0.06) between ε_{it} and $\varepsilon_{i,t-2}$, and 0.01 (-0.05) between ε_{it} and $\varepsilon_{i,t-3}$.

With a correlation of 0.06 between ε_{it} and $\varepsilon_{i,t-2}$, firms could (if they had estimated models like those presented here) predict 6% of the variation in ε_{it} when they take decisions that affect the value of the first lags of the competition variables. This could causa a small bias. For example, based on Monte Carlo simulations, Keele and Kelly [18] report biases of less than 1% for both the short- and long-term effects when the correlation coefficient is 0.10. The bias would be smaller if the second lags of the competition variables were used as instruments instead of the first lags, because ε_{it} is less correlated with $\varepsilon_{i,t-3}$ than with $\varepsilon_{i,t-2}$. Table A1 therefore report estimation results obtained when using second lags of the competition variables as instruments. These results are similar as those reported in Table 2, which indicates that the choice of instruments have small effects on the results and that the possible bias is indeed small.

Simultaneously including fixed effects and a lag of the dependent variable can give bias. Fortunately, this bias in the estimator for the coefficient of the lagged dependent variable (θ) is decreasing in the number of time periods. According to Nickell [19], the limit of the bias for the parameter θ as N approaches infinity can be approximated by $-(1+\theta)/(\mathrm{T}-1)$, in which N and T are the number of fixed effects and time periods. In addition, for $\theta = 0.9$, and when 90% (95%) of the total variance are due to the fixed effects, Nerlove [20] finds a bias that is just 40% (26%) of the bias suggested by the approximation written above. For the two samples, the fixed effects explain 93% and 89% of the total variation, and the average of time periods a product is included in the analyses are 42 and 46, respectively. With values of θ of about 0.96, this bias would about -0.016 and -0.017, assuming that the bias is 40% and 33% of $-(1+\theta)/(\mathrm{T}-1)$, respectively. Because of the small magnitudes of the expected biases and since I was not able to instrument the

explanatory variables when using first difference transformation, I presented results from estimations in which I have not accounted for this small bias.

Table A1. Robustness checks, estimation results for lnP_{it}

	Discounts forbidden	*Discounts allowed*
	IV 2	IV 2
$lnP_{i,t-1}$	0.9568***	0.9475***
	(0.0055)	(0.0118)
$D\_PiSubstance_{st}$	-0.0019***	-0.0005
	(0.0007)	(0.0008)
D_PiE_{it}	-0.0012	-0.0008
	(0.0008)	(0.0007)
$lnN\_PiSubstance_{st}$	0.0000	-0.0003
	(0.0004)	(0.0006)
lnN_PiE_{it}	-0.0012*	-0.0000
	(0.0007)	(0.0007)
D_Th_{st}	-0.0002	0.0003
	(0.0010)	(0.0015)
D_ThGen_{st}	-0.0012**	0.0017*
	(0.0006)	(0.0010)
lnN_Th_{it}	-0.0003	-0.0063**
	(0.0013)	(0.0026)
lnN_ThGen_{st}	0.0010	-0.0061**
	(0.0008)	(0.0028)
$dlnP_i^*/dD\_PiSubstance_{st}^*$	-0.0671***	-0.0244
	(0.0165)	(0.0160)
Observations	119,058	78,820
R^2	0.9181	0.9156
K-P rk LM	69.1940	64.504
K-P rk LM, p-val.	0.0000	0.0000
Hansen J, p-value	0.1496	0.3050

Note: The specifications include product-specific fixed effects and 58 and 81 indicator variables for months, respectively. In the first-stage regressions, data from October 2002–October 2007 and Jan. 2011–Dec. 2017 were used. K-P rk LM refers to the Kleibergen-Paap rk LM statistic, which indicates the strength of the instruments. The null hypothesis in the K-P test is that the model is under-identified. The null hypothesis for the Hansen J test is that the instruments are valid, i.e., uncorrelated with the error term. Standard errors, robust to correlations within substances, are given in parentheses. ***, **, and * indicate that the coefficient is statistically significant different from zero on the 1%, 5% and 10% significance levels, respectively. The estimation results for the indicator variables for months and for the first-stage regression are available on request.

References

1. West P, Mahon J. Benefits to payers and patients from parallel trade. York: York Health Economics Consortium; 2003.
2. Ganslandt M, Maskus KE. Parallel imports and the pricing of pharmaceutical products: Evidence from the European Union, Journal of Health Economics, 2004;23(5);1035–1057.
3. Granlund D, Köksal-Ayhan MY. Parallel imports and mandatory substitution reform–A kick or a muff for price competition in pharmaceuticals? The European Journal of Health Economics, 2015;16;969–983.
4. Granlund D, Köksal-Ayhan MY. EU enlargement, parallel trade and price competition in pharmaceuticals: Has the price competition increased? The B.E. Journal of Economic Analysis and Policy, 2016;16;1069–1092.
5. Vandoros S, Kanavos P. Parallel Trade and Pharmaceutical Prices: A Game-theoretic Approach and Empirical Evidence from the European Union. The World Economy, 2014;*37*(6);856–880.
6. Kanavos P, Costa-Font J. Pharmaceutical parallel trade in Europe: Stakeholder and competition effects, Economic Policy, 2005;20(44), 751–798.
7. Kanavos P, Vandoros S. Competition in prescription drug markets: Is parallel trade the answer? Managerial and Decision Economics, 2010;31(5), 325–338.
8. Granlund, D (2020) Price effects of competition from parallel imports and therapeutic alternatives–Using dynamic models to estimate the causal effect on the extensive and intensive margins, manuscript Umeå University.
9. Dental and Pharmaceutical Benefits Agency. TLVFS 2009:4, Tandvårds- och läkemedelsförmånsverkets föreskrifter och allmänna råd om prissättning av utbytbara läkemedel och utbyte av läkemedel m.m. 2009. [Dental- and Pharmaceutical Benefit Agency's regulations and general advices about price setting of exchangeable pharmaceuticals and substitution of pharmaceuticals etc.] (in Swedish) http://www.tlv.se. Accessed 2 April 2 2012.
10. Pharmaceutical Benefits Agency. LFNFS 2003:1, Läkemedelsförmånsnämndens föreskrifter om ansökan och beslut hos Läkemedelsförmånsnämnden. 2003. [The Pharmaceutical Benefits Board's regulations about applications and decisions at the Pharmaceutical Benefits Board] (in Swedish) https://www.tlv.se/download/18.73f061881604dfc0a6ca916/1513168391008/LFNFS-2003-1.pdf. Accessed 10 June 2020.

11. Pharmaceutical Benefits Agency. LFNAR 2006:1, Läkemedelsförmånsnämndens allmänna råd om grunder för prishöjningar på läkemedel. 2006. [Guidelines Concerning Price Increases of Pharmaceuticals from the Pharmaceutical Benefits Board]. (in Swedish) www.tlv.se/download/18.7e3d365215ec82458645b2/1510316403476/LAG-lfnar-2006-1.pdf. Accessed 10 June 2020.
12. Bolotova Y, Connor JM, Miller DJ. The impact of collusion on price behavior: Empirical results from two recent cases. International Journal of Industrial Organization, 2008;26(6),1290–1307.
13. Linnosmaa I, Karhunen T, Vohlonen I. Parallel Importation of Pharmaceuticals in Finland. Pharmaceutical Development and Regulation, (2003);*1*(1);67–74.
14. Dental- and Pharmaceutical Benefit Agency. 2014/2015 års översyn av apotekens handelsmarginal – slutrapport. 2015. [The years 2014/2015's review of the pharmacy trade margin – final report.] (in Swedish) https://www.tlv.se/download/18.1d85645215ec7de2846bf339/1510316371481/oversyn_apotekens_handel_marginal_del_5_slutrapport.pdf. Accessed 10 June 2020.
15. Dental- and Pharmaceutical Benefit Agency. 2017 års uppföljning av apoteksmarknadens utveckling. 2018. [The year 2017's follow-up of pharmacy market development.] (in Swedish) https://www.tlv.se/download/18.13a4b04161527df3769591a/1517836689989/uppfoljning_av_apoteksmarknadens_%20utveckling_2017.pdf. Accessed 10 June 2020.
16. Dental- and Pharmaceutical Benefit Agency. 2018 års uppföljning av apoteksmarknadens utveckling. 2018. [The year 2018's follow-up of pharmacy market development] (in Swedish) https://www.tlv.se/download/18.192533fa166f516fb27da81b/1542378329327/uppfoljning_av_apoteksmarknadens_utveckling_2018.pdf. Accessed 10 June 2020.
17. Cumby RE, Huizinga J. Testing the autocorrelation structure of disturbances in ordinary least squares and instrumental variables regressions, Econometrica, 1992;60(1);185–195.
18. Keele L, Kelly NJ. Dynamic models for dynamic theories: The ins and outs of lagged dependent variables, Political analysis, 2006;14(2);186–205.
19. Nickell S. Biases in dynamic models with fixed effects, Econometrica: Journal of the Econometric Society, 1981;49(6);1417–1426.
20. Nerlove M. Experimental evidence on the estimation of dynamic economic: Relations from a time series of cross-sections, Economic Studies Quarterly, 1967;18(3);42–74.