On the Role of Parallel Trade on Manufacturers and Retailers Profits in the Pharmaceutical Sector

Pierre Dubois*

Toulouse School of Economics

Morten Sæthre†

Norwegian School of Economics

May 2016 ‡

Preliminary

Do not cite without permission

Abstract

We study how cross-national differences in regulated pharmaceutical prices within the EU lead to arbitrage decisions by pharmacy retailers through parallel imports of versions of drugs originally marketed in other countries by the same company. Such strategic decisions can affect the distribution of profits in markets for prescription drugs, including the profitability of innovating pharmaceutical companies even before patents expire and generic competition enters. Before patent expiry, parallel trade is the unique source of upstream competition from the perspective of a pharmacy retailer. We develop a structural model where retailers can alter the set of goods which the consumer can choose from, in response to differences in profitability across products. Our demand model with unobserved choice sets can be identified using individual consumers´ choices and supply side conditions for optimal choice sets decisions. Estimating our model with rich data on a pharmaceutical market featuring parallel imports, we find that retailer incentives play a significant role for the observed outcome. Our counterfactual simulations show that parallel imports of drugs has small effects on average consumer welfare, while it has large implications for the distribution of industry profits. In particular, retailers gain at the expense of pharmaceutical companies, while parallel traders only attain modest profits.

JEL codes: I11, L22

Key words: Parallel trade, pharmaceuticals, vertical contracts.

*e-mail: Pierre.Dubois@tse-fr.eu

†e-mail: Morten.Saethre@nhh.no

‡We are grateful to Bruno Jullien, Volker Nocke, Patrick Rey, Nicolas Schutz for insightful comments.

1 Introduction

The EU treaty made arbitrage trade of pharmaceutical drugs across countries fully legal within EU. Cross country price differences within the EU due to national price regulations have maintained substantive prices differences leading to an increase of parallel trade estimated at 5.5 billion euros in 2012 with highly heterogeneous national market shares than can be up to 25% in some countries. Cross country price differences can be as large as 300% and are due to regulatory caps or strict price setting government rules. These price differences are not surprisingly giving rise to parallel imports by high price countries form low price countries. There are however also large variations in pharmaceutical parallel import penetration across otherwise similar countries. As entry of parallel traders needs to be done through pharmaceutical retailers, it is likely that retailers and their incentives are very important intermediaries in the supply chain. The strategic role of profit maximizing pharmacies with respect to both drug manufacturers providing directly imported drugs and parallel traders providing parallel imports can thus be important in the organization of the pharmaceutical sector. Parallel imports are versions of drugs originally marketed in other countries by the same pharmaceutical company. They are potentially differentiated (by appearance and packaging) form the consumer point of view but are essentially the same products likely to be very substitutes. We thus study how the presence of parallel trade together affect incentives of pharmaceutical retailers and wholesale negotiations between pharmacies and upstream producers. These may have large implications for the distribution of surplus in the market.

In this paper, we thus develop a structural model to address questions concerning sales of parallel imported pharmaceutical drugs. Specifically, we address the question of incentives of retail pharmacies in facilitating sales of parallel imports amid regulatory policies towards parallel trade and price setting of pharmaceuticals. Our model allows us to explain how parallel imports can capture substantial market shares, even though savings afforded to consumers might be negligible or even non-existent. The mechanism of our model is that a retailer facing regulated prices might wish to restrict supply of less profitable products, which might reduce the number of consumers making their purchases with the retailer, as restrictions in the choice set generally decreases the expected utility of visiting the retailer. The problem of the retailer can thus be viewed as a variant of the classical trade-off between price-cost margin and quantity sold.

We estimate our model using a very rich data set on the Norwegian pharmaceutical market, where we are able to observe the purchases of individual consumers over time, the pharmacy chain at which a given purchase happened, and whether the specific drug dispensed was imported through the original producer (direct import) or by parallel traders. Since we also have data on the retail price the pharmacy obtains for all dispensed sales, as well as data on the wholesale prices paid by the pharmacy chain to the upstream firms for each product purchased —i.e. specific drug packages—we can calculate the gross margin obtained by the chain on each product, which affects retailers' incentives to dispense parallel imports. We find that the inclusion of retailer incentives in our model plays an important role in explaining consumer choices, and also that restrictions on supply are prevalent and have important implications for the outcomes in the market we study. Our counterfactual simulations imply that even though consumers prefer directly imported products, parallel imports have a very modest effect on consumer welfare. Further, we find that it allows the retail pharmacy chains to capture a much larger share of industry profits than would otherwise be the case, particularly at the expense of the original producer. This redistribution of profits is especially driven by the chains' ability to shift sales as a response to differences in profitability, which also allows the chains to capture larger shares of profits due to parallel imports than would be the case without this strategic instrument. Many industries rely on a downstream retailing sector to market goods. Vertical relationships between upstream producers and downstream retailers are determinant for market access of goods and consumer welfare. Actually, vertical relationships are not only determinant for price competition among substitute and differentiated goods, but intermediaries between producers and consumers—such as retailers—may also affect competition by engaging in other strategic actions affecting final consumers' demand. Such strategic actions can be very diverse and also affect upstream producers' behavior. Strategic behavior of retailers could include strategic choices regarding assortment of goods, introduction of store brands or private labels, advertising and promotion efforts. Equilibrium results of such structures can be analyzed thanks to game theoretic models and estimated through structural models. Typical sectors where retailers' behavior has attracted attention of economists are the food retailing industry with large supermarket chains and internet platforms.

Pharmacy retailing is another sector that has been less studied, though for which the growth in health care expenses among developed countries raises many questions on which policies allow containment of pharmaceutical drugs' costs while ensuring or improving access to patients. In Europe, most countries regulate prices of prescription drugs, though other aspects of competitive behavior, such as strategic choice of entry across different markets, matter substantially (Danzon et al., 2005). Within Europe, another element of competition is introduced by the possibility to trade drugs across countries. Parallel trade of drugs within Europe thus interacts with national regulatory pricing policies. With such parallel trade intermediaries, pharmaceutical retailers find an alternative upstream provider of drugs competing with national direct importers for the same branded or unbranded molecules. How pharmacists use those upstream providers strategically can rely on strategic contracting with upstream direct or parallel importers, and on pharmacists' action affecting final demand. One strategy can be based on strategic assortment choice of drugs, proposing one or two versions of the same drug which is either directly imported or imported by a parallel trader. This is similar to strategically choosing to stock out of some versions of drugs. Indeed, absent considerations of storage costs and logistics, retailers may find beneficial to stock out of a product to induce substitution towards other product, for which the retailer has higher margins. This behavior can be beneficial in other industries but it could especially be the case in tightly regulated markets where the price setting is constrained, such as is common in many European countries for pharmaceuticals.[1] What we have in mind is an endogenously set stock of each product, such that there is a chance that any given consumer entering the store will just face a subset of the products.

We will abstract from the decision on level of stocks and timing of consumer visit by letting the firm choose the probability of each product being available which will be independent across consumer visits for a suitable period of aggregation. The reason for pharmacy stores having variety is to attract consumers who obtain higher expected utility, while the reason for not having all products available at all times is to induce substitution towards more profitable products by proposing only the higher margin version of the drug. Within the European Economic Area, drugs are allowed to be bought in one member state and resold in another, as long as the product is sufficiently similar to one already authorized for sale in the destination member state[2]. This has made possible a large arbitrage trade between member states, where drugs are often bought in Southern European countries, such as Greece, Portugal and Spain and resold in Northern European countries, such as the UK, Netherlands and Sweden (Kyle, 2011).

To a large extent, the price differences in the European Economic Area can be attributed to differences in price regulation, and—where such arrangements are in place—the aggressiveness of each member state's authorities in negotiating with the manufacturers (Kyle, 2007).

Parallel traded drugs provide an interesting market in terms of studying retailer incentives. For many European countries, the prices of prescription drugs is regulated, while the consumers are covered by national health insurance, and will often face substantially lower costs than the full price. Often, this leads to very small or no price differences between direct and parallel imported drugs[3]. There are also large cross-country differences in the share of parallel import sales, which according to the findings of Kanavos et al. (2004) seems to have a clear link to the regulation governing margins at the pharmacy and domestic supply level. As an example, German pharmacies are subject to regulations fixing their margins and have to a low extent supplied parallel trade before the national insurance authorities set a minimum quota of 5%, while British pharmacies have no direct caps on their margins and have large shares of parallel import (Kanavos et al., 2004). Even though studies of aggregate prices have not found evidence of any significant reduction in price dispersion across EU countries (Kanavos et al., 2004; Kyle, 2011), Ganslandt and Maskus (2004) report that parallel imports might have led to a reduction in drug prices on the order of 12–19% for drug segments

[1] For details on regulation in different countries, see, e.g., Kanavos et al. (2008).

[2] These operations are done by firms specializing in parallel trade, and requires the necessary logistical capacity and facilities suitable for the repackaging of drugs, which is required for drugs where the imported package is in another language than the language of the destination country.

[3] Direct imported drugs are the ones supplied by the manufacturers or their marketing agencies.

subject to parallel imports entries in Sweden. Using a structural model of demand estimated using data on the German market for oral anti-diabetic drugs, Duso et al. (2014) evaluate the welfare impact of parallel imports. Their estimates imply that parallel imports have reduced the prices of on-patent drugs by 11%, but that the impact on consumer surplus is modest. In contrast to their approach, we explicitly model the vertical relationship between pharmaceutical manufacturers and pharmacy retail chains and show the strategic role of retailer incentives in the development of parallel imports market shares.

Using the same data from Norway, Brekke and Straume (2015) study the interaction between price cap regulation and parallel imports. Using data across a large number of drugs, they find evidence that original producers might benefit from tighter price regulation when there is competition from parallel trade. As we consider the market for a single, large drug, we do not touch upon this question. They also utilize a Nash bargaining model for the contracting between pharmacy chains and upstream producers to motivate their empirical analysis, though the channels they highlight are slightly different from ours. Novel features of our paper include the strategic decisions by the retailers on the drugs proposed to consumers, and structural estimation of the bargaining model.

There would be reason to suspect that some consumers would prefer the direct imported variety, even though the drugs are produced by the same company. The packaging will usually display the brand name of the parallel importer and the product can differ in visual appearance and inactive ingredients. This could be for instance tablets where the direct imported variety is round and white, while the parallel imported comes from a country where they are octagonal and red. This type of differentiation in appearance and specification across countries has been linked to attempts to reduce the scope for parallel trade (Kyle, 2011). In this sense, as we would suspect consumers to be either indifferent between direct and parallel import or skeptical towards parallel import and as the price they pay are usually the same, it seems necessary to consider the incentives of the retail side of the market to explain why parallel imports are sold and the reason for the observed cross-country differences in dispensing practice for parallel imports.

We estimate the full structural model of demand and vertical contracting on the supply chain.

We present the basic model in the following section. In Section 3, we present the market and data. In Section 4, we describe the empirical specification of our model and present the estimation results. In Section 5, we present the results from our counterfactual simulations, while Section 6 concludes.

2 The Norwegian Pharmaceutical Market and Parallel Imports

2.1 Overview

The supply side of the market for prescription drugs consists mainly of three large pharmacy retail chains, which are vertically integrated with each of their upstream wholesaler. The three largest chains, Apotek 1, Boots and Vitus, covers 85 % of all pharmacies, while public hospital pharmacies (6 %), a smaller retail chain (5 %), and independent pharmacies (4 %) make up the rest.

The Norwegian market for drugs is subject to a wide array of regulations. The Norwegian Medicines Agency, a governmental organization under the Ministry of Health and Care Services, is the main regulatory body for drug affairs, in charge of marketing authorization, drug classification, vigilance, price regulation, reimbursement regulation, and providing information on drugs to prescribers and the public.

2.2 Regulation

The Norwegian Medicines Agency, a governmental organization under the Ministry of Health and Care Services, is the main regulatory body for drug affairs, in charge of marketing authorization, drug classification, vigilance, price regulation, reimbursement regulation, and providing information on drugs to prescribers and the public.

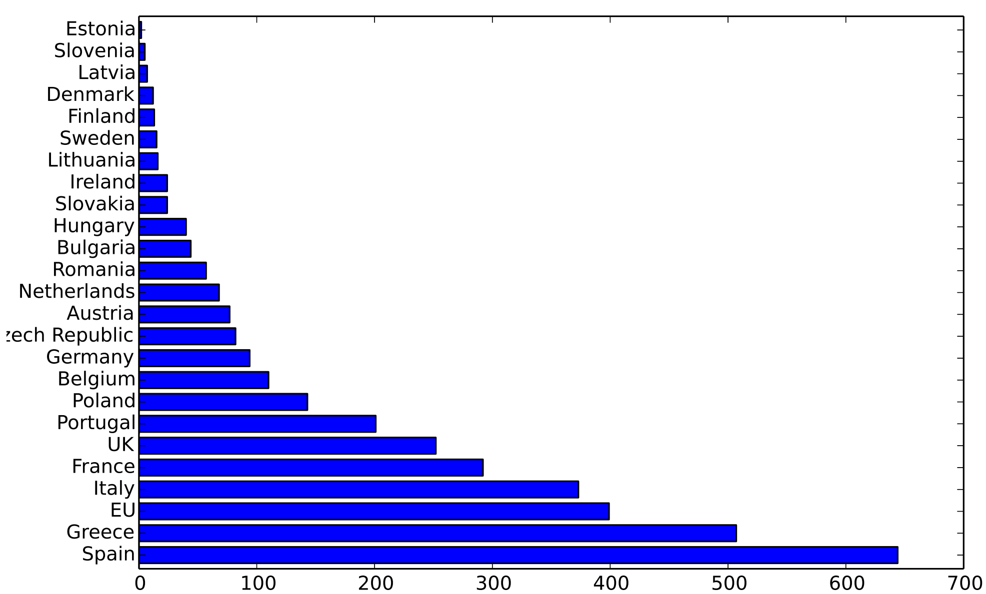
All drugs sold on the Norwegian market are subject to a price cap, set by the Norwegian Medicines Agency. In most cases, this price cap is set as the average of the three lowest among market prices in a fixed group of European comparison countries, consisting of Sweden, Finland, Denmark, Germany, United Kingdom, Netherlands, Austria, Belgium and Ireland. The price caps should normally not change more than once every year. Reconsideration of the price caps will be initiated by the Norwegian Medicines Agency, where selection is based on sales volume over the past 12 months. The price caps are set according to the active ingredient in the drug and amount of active ingredient (dosage). Per unit price caps (unit being defined by Defined Daily Dose (DDD) for drugs where it exists) should generally be equal within the category of a given dosage for a given active ingredient, although the Norwegian Medicines Agency imposes package sizes to have lower per-pack price in some cases.

In cases where the patient has a long term ailment, defined as demanding treatment for at least three months, and the drug under question has been judged to have sufficient effect compared to the costs, government reimbursement is available. The prescribing physician is responsible for deciding if the patient satisfies the criteria for treatment length, while the Norwegian Medicines Agency decides if a drug is efficient and cheap enough to be put on the list of drugs approved for reimbursement. When patients get reimbursed, they face a co-payment of 38 % of the total price, capped at 520 NOK in 2013 (approximately 65 EUR) per three months. The co-payments for drugs and health care spending are capped at 2040 NOK yearly in 2013 (approximately 260 EUR). For drugs that are on patent, the government will reimburse the full cost of the drug to the patient, net of co-payments. When the drug is off-patent and generic drugs have entered the market, the reimbursement rate will generally be reduced below the price ceiling according to governing regulation on the off-patent market. As we exclusively consider on-patent drugs (which are the main drugs parallel imported), we refrain from describing the regulation governing the off-patent market in more depth.

2.3 Parallel Trade

Parallel traded drugs will have to obtain a license for selling drugs in Norway from the Norwegian Medicines Agency, unless they already have obtained a license for sales in the European Economic Area through the centralized European Union procedure. Sales will be to one of the three large wholesalers, as only full-line wholesalers are allowed by law to supply pharmacies with drugs.[4] A license will be for a specific drug package, from a specific country, with the exception of licenses granted through the European Union procedure. In Figure 2.1, we see the number of licenses granted by the Norwegian Medicines Agency by source country, which is in line with the countries that are usually reported as major exporters of drugs in this fashion.

Figure 2.1: *Granted licenses in Norway for parallel import by country*

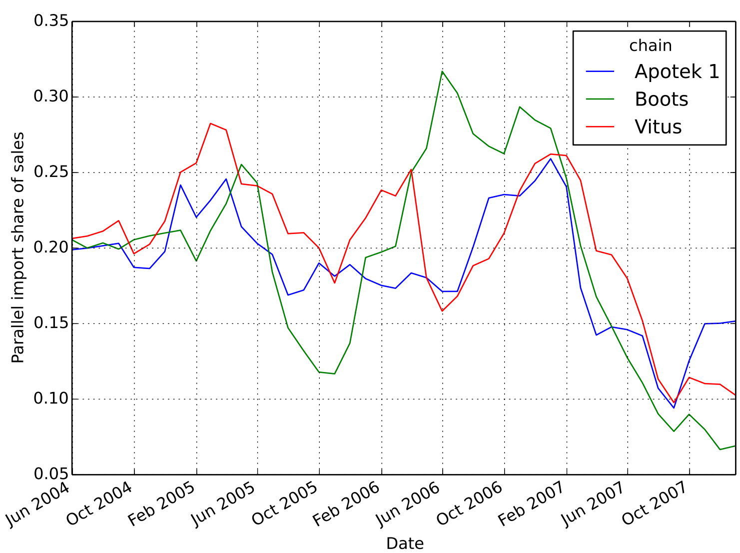


Source: *File provided to one of the authors by The Norwegian Medicines Agency.*

[4]Whether this matters is an open question, as the market is almost fully vertically integrated at the wholesaler-pharmacy level, although the full-line supply regulation could be an explanation for the concentration and vertical integration observed in the market.

We have access to highly detailed data on drug sales in Norway during 2004-2007, where we can identify whether the product is directly imported or parallel imported. Parallel trade happens most prominently in the on-patent period, and is virtually non-existent after generic entry. Cleaning the data for parallel imported products that obtain very low share of sales in their segment, we identify 50 active ingredients where parallel trade occurs in our sample period. All following analysis is done on drugs within this subsample of active ingredients. Monetary sizes are reported in nominal NOK (≈ 0.12 EUR / 0.16 USD in the period) unless noted otherwise.

Figure 2.2: *Parallel import share of sales in DDD by chain*

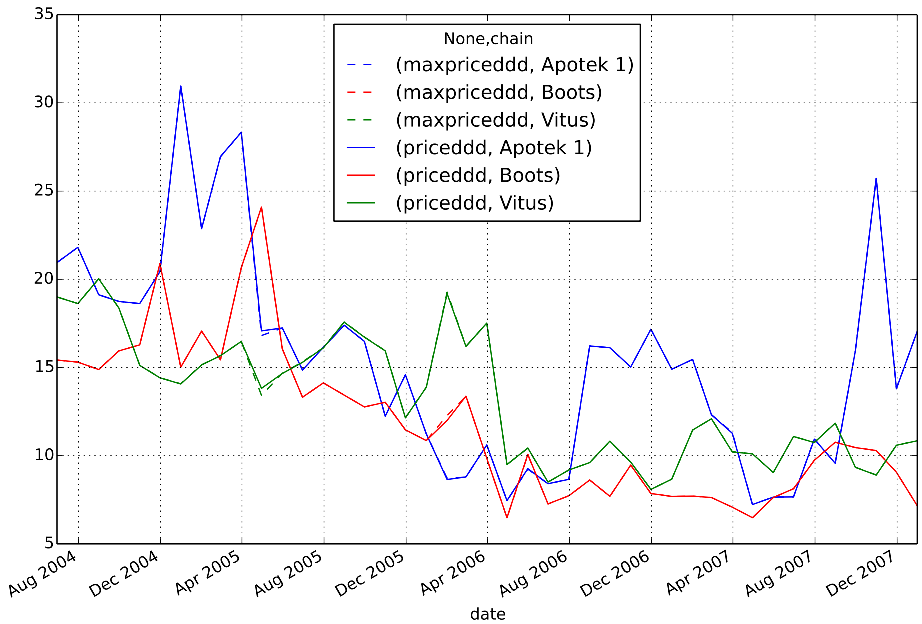


Note: *Only ATC codes featuring sales of parallel imports over the sample period are included.*

In Figure 2.2 we see the parallel import share of sales. It is interesting to note that there is large variation both between chains and over time. Figure 2.3 and 2.4 shows the volume weighted price and regulatory price cap in each chain for direct and parallel imported drugs respectively. In these figures, it is apparent that overall the price ceiling is binding for both categories. The differences between the chains and between direct and parallel imported drugs within chain in these figures are due to differences in the profile of drugs dispensed. Pharmacies could have differences in the type of packages they dispense, which could be due both to variation in the consumers they serve and differences in the size of the packages. For parallel imports, the differences are mostly due to differences in the size of packages dispensed compared to the direct imports within the same active ingredient and dosage category.

For products where we find both parallel and direct imports for a given active ingredient, dosage and package size within a pharmacy chain, we calculate the difference in price between parallel and direct

Figure 2.3: *Prices for direct imported drugs*



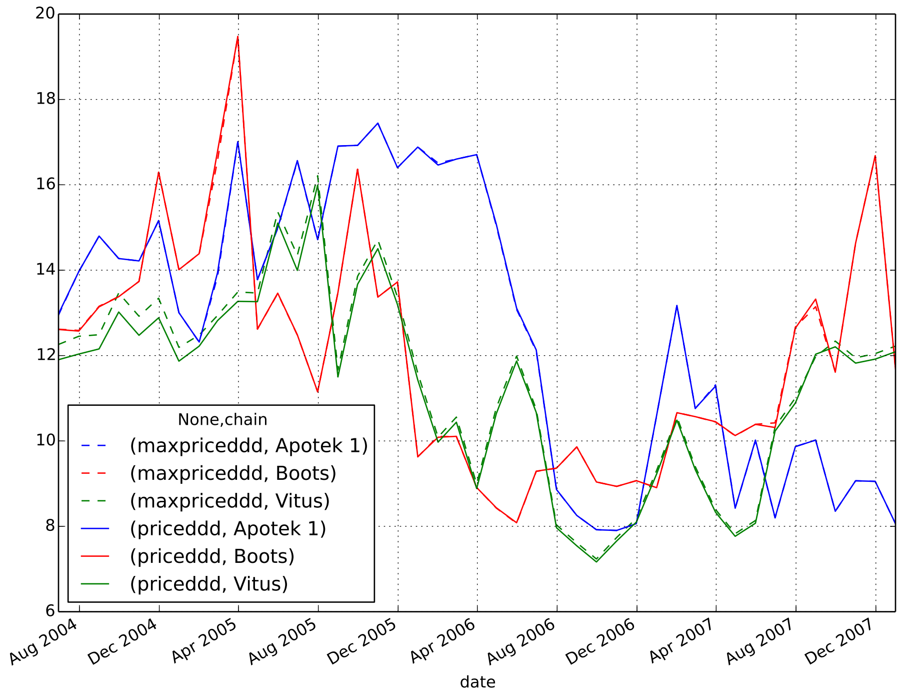
Price per DDD (priceddd) and price ceiling per DDD (maxpriceddd) by chain.

imported products.[5] The volume weighted averages of these price differences for each pharmacy chain are shown in Figure 2.5, where we see that prices are virtually the same for products that do not differ according to active ingredient, dosage and package size.

We also look at the differences in margin that the pharmacy chain obtains on comparable products, i.e. within categories defined by active ingredient, dosage and package size. Here, pharmacy chain margin is defined as the sales price in the pharmacy net of the price the pharmacy chain's integrated wholesaler pays to the supplier for obtaining the drug, where the supplier is either a marketing agency for the manufacturer, in the case of direct imports, or the parallel trading firm. These margin differences are shown in Figure 2.6. It might seem like there is some correlation with the parallel import share of sales in Figure 2.2, something that is confirmed by the significant chain-month level positive correlation between the parallel import shares and the margin difference between parallel and direct imports. This however cannot be interpreted as a causality effect but this is a first indication of pharmacy incentives mattering for the composition of drugs dispensed to consumers. The correlation is however far from perfect and also tells us that the margin difference is not the only thing driving the variation we see in the evolution of parallel import sales. We would suspect that there is also some strategic interaction going on between the importing firms and the pharmacy chains,

[5] Package size is defined as DDD per package, which for pills with the same active ingredient and dosage would be equivalent with pills per package.

Figure 2.4: *Prices for parallel imported drugs*

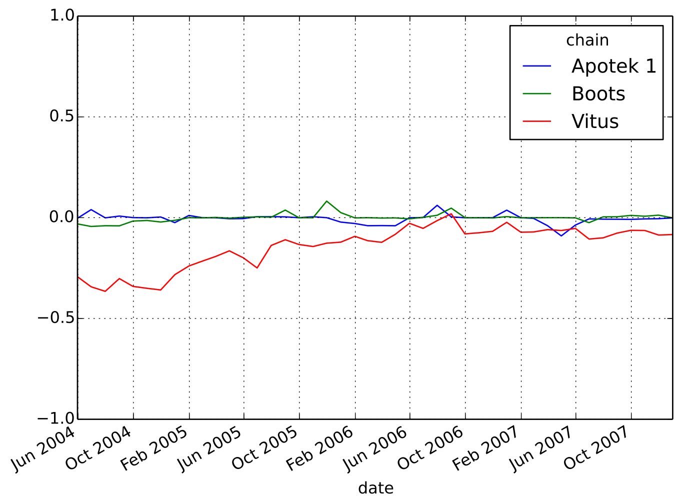


Price per DDD (priceddd) and price ceiling per DDD (maxpriceddd) by chain.

which will make the simple estimate biased if the benchmark is a causal estimate of the effect of margin difference on product sales, and also uninformative as input for a model for evaluating the market in terms of market power and policy evaluation.

In the Norwegian market in this period, there are five companies specializing in parallel trade with any noticeable activity, namely Cross Pharma, Euromedica, Farmagon, Orifarm and Paranova. The share of parallel import sales within each pharmacy chain for each of these companies are shown in Figure 2.7. There is some variation both between pharmacies and over time in terms of the relative presence for these companies. On each active ingredient, each pharmacy chain seems to stay with one parallel importer at a given time, although the identity of the parallel importer could be different across chains for the same drug, and also change over time for a drug within a chain. This could be an indication of harsh competition in the parallel import segment.

Figure 2.5: *Price difference between direct and parallel imports*



Prices differences are reported in NOK per DDD, price of parallel import minus price of direct import. Differences calculated on packages that are of the same ATC code, with the same amount of active ingredient and of comparable size.

3 A Structural Model of Demand and the Supply Chain

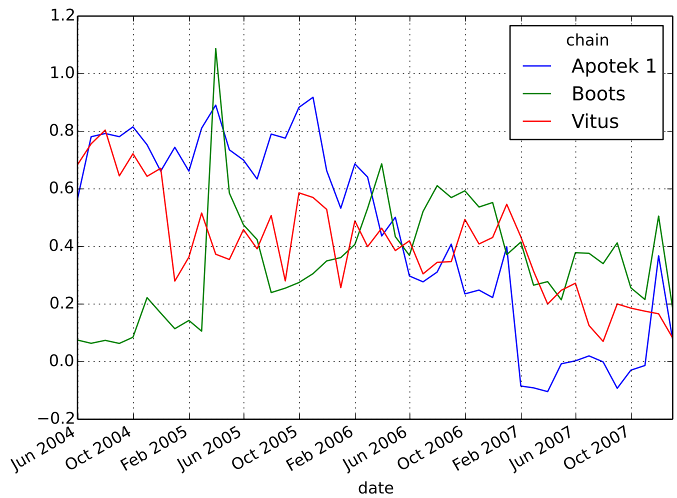
To explain the variation we see in the data, we aim at creating an estimable model that can be used for comparison and useful counterfactuals in terms of policy implications and cross-country differences.

3.1 Consumer Behavior and Demand for Parallel Trade Products

In this baseline model, we assume that the consumer has an exogenous need for a drug with a particular active ingredient and dosage and abstract from the issue of therapeutic choice decided by prescribers which, as we will show below, seems not be affected by the availability of parallel trade versions of active ingredients and thus exogenous to the fundamental mechanisms of our model.

The consumer makes a choice over which pharmacy chain c to visit, and - once in the pharmacy - a choice from the available choice set in the pharmacy. When the consumer chooses a pharmacy c, he does not know if parallel imported (PI) or direct imported (DI) versions of the drug will be available, but knows an expected availability when he chooses which pharmacy chain to visit. Because pharmacies potentially have higher margins on drugs that the consumer do not strictly prefer, it may be optimal not to propose the lower margin drug with certainty in order to induce consumers to buy the other option. It may also be optimal to propose consumers' preferred drug with a non zero probability in order to attract them. This phenomenon is confirmed by casual observation and the fact that pharmacists do consider this policy of non permanent

Figure 2.6: *Difference in product margin between direct and parallel imports*



Margin differences are reported in NOK per DDD, margin of parallel import minus margin of direct import. Differences calculated on packages that are of the same ATC code, with the same amount of active ingredient and of comparable size.

availability is also acknowledge in discussions with pharmacists [6]. We thus assume that consumers' know the probabilities of availability chosen by the pharmacy chains.

For a given active ingredient, the choice set at pharmacies can be $\{PI\}$, $\{DI\}$ or $B \equiv \{DI, PI\}$. As will be clear from inspection of the pharmacy chains profit maximization objective, it is never optimal for them to be completely stock out of both products[7]. We let the origin of the drug be indexed by $k \in \{0, 1\}$ where 0 denotes PI and 1 denotes DI.

We denote θ^0_{ct} and θ^1_{ct} the probabilities that the choice sets are $\{PI\}$ or $\{DI\}$ respectively and thus $1 - \theta^0_{ct} - \theta^1_{ct}$ the probability that the choice set is $B = \{DI, PI\}$. We assume that the utility of consumer i is given by

$$u_{ikct} = \alpha_{ikct} + \epsilon_{ikct}$$

where α_{ikct} is the mean utility consumer i obtains from choosing the drug of origin k in pharmacy chain c in market t, and ϵ_{ikct} is an idiosyncratic i.i.d. random utility component, that we assume distributed independently across drugs and chains according to a Gumbel distribution. Note that the mean utility component could be a function of observable characteristics of drugs, chains, time, or the consumer.

[6]According to an industry source we've talked to, most customers do not object to substitution to parallel trade, though some consumers are concerned, and could even insist on obtaining the direct imported version. According to our source, the chain sets the standard policy for deliveries of stocks to the pharmacies, which the pharmacist can then alter or make additional orders for non-standard selections. The experience was that for several drugs, the standard policy included few or none of the direct imported version, such that the pharmacy easily could end up only having the parallel traded version available. Since obtaining an additional order would take at least one day, there was a worry that customers insisting on the direct imported version would rather go to a competing pharmacy.

[7]Note that this is absent of any considerations of storage costs.

Figure 2.7: *Composition of parallel importers*

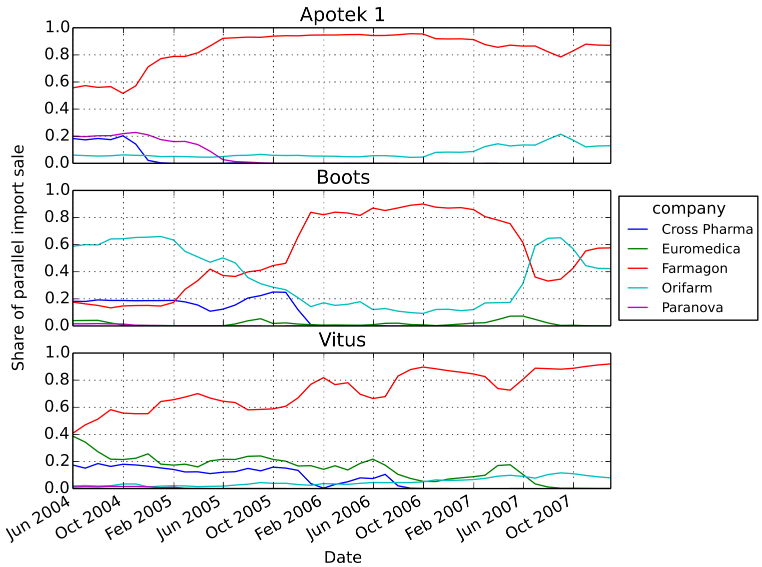


Figure 2.8: *Share of parallel import sales in DDD within pharmacy chain for each parallel importer (company name).*

Thus, within the pharmacy, the probability that consumer i chooses k conditional on choice of pharmacy chain c when both products are available in the pharmacy is given by

$$s_{ikt|c,B} = \frac{e^{\alpha_{ikct}}}{e^{\alpha_{i0ct}} + e^{\alpha_{i1ct}}} = \frac{1}{1 + e^{\alpha_{ikct} - \alpha_{ik'ct}}} \text{ with } k' = 1 - k$$

Assuming that the consumer always prefer the available drug than no drug, the choice probability of product k conditional on the choice of pharmacy c is

$$\underbrace{s_{ikt|c}}_{\substack{\text{choice probability of } k \\ \text{conditional on} \\ \text{going to chain } c}} = \underbrace{\theta_{ct}^{k}}_{\substack{\text{probability that} \\ \text{only } k \text{ is} \\ \text{available}}} + \underbrace{(1 - \theta_{ct}^{0} - \theta_{ct}^{1})}_{\substack{\text{probability that} \\ \text{both versions} \\ \text{are available}}} \underbrace{s_{ikt|c,B}}_{\substack{\text{choice probability of } k \\ \text{given both versions} \\ \text{are available}}},$$

that is, the probability of drug k being the only available plus the probability that drug k is chosen when both are available times the probability that both are available.

Then, the expected utility the consumer gets from any choice set S_{ct} in a pharmacy c is equal to the inclusive value of this choice set, denoted $I_{i|S_{ct}}$, which is given by the log-sum formula[8]

$$I_{i|S_{ct}} \equiv E\left(u_{ikct}|S_{ct}\right) = \ln\left(\sum\nolimits_{k\in S_{ct}} e^{\alpha_{ikct}}\right),$$

When k is the only drug available, the inclusive value is equal to α_{ikct}, i.e., the deterministic utility of drug k. Note that the inclusive value when both are available is always larger than any of the α_{ikct} alone.

The expected utility of visiting pharmacy chain c in market t is then given by the expected inclusive value, which we denote by I_{ict}, where the expectation is taken over the uncertain - from the consumer's point of view - choice set,

$$I_{ict} \equiv E_{S_{ct}}\left[I_{i|S_{ct}}\right] = \sum\nolimits_{k} \theta_{ct}^{k}\alpha_{ikct} + (1-\theta_{ct}^{0}-\theta_{ct}^{1})\ln\left(\sum\nolimits_{k} e^{\alpha_{ikct}}\right).$$

We assume that patient i chooses chain c in order to maximize his indirect utility $I_{ict}+\varepsilon_{ict}$, where ε_{ict} is extreme value Gumbel distributed independently across chains. The probability that consumer i visits chain c is then

$$s_{ict} = \frac{e^{I_{ict}}}{\sum_{c} e^{I_{ict}}}.$$

Then, denoting by $F(.)$ the cumulative distribution function of consumer preferences $\alpha_{it} \equiv (\alpha_{i01t}, .., \alpha_{iC1t}, \alpha_{i11t}, .., \alpha_{i1Ct})$, we can write the aggregate choice probability or market share of drug k sold by c at t as

$$s_{kct} = \int s_{ikct}dF(\alpha_{it}) = \int s_{ict}s_{ikt|c}dF(\alpha_{it}), \tag{3.1}$$

and the aggregate market share of drug k within the pharmacy chain c as

$$\begin{aligned} s_{kt|c} &= \int s_{ikt|c}dF(\alpha_{it}) \\ &= \int \left[\theta_{ct}^{k} + s_{ikt|c,B}(1-\theta_{ct}^{0}-\theta_{ct}^{1})\right] dF(\alpha_{it}) \\ &= \theta_{ct}^{k} + \left(1-\theta_{ct}^{0}-\theta_{ct}^{1}\right)\int s_{ikt|c,B}dF(\alpha_{it}). \end{aligned}$$

3.2 Pharmacy Chains Behavior

Let us now turn to the behavior of the pharmacy chains. The profits of chain c normalized by market size in time t are

$$\pi_{ct} = \sum\nolimits_{k\in\{0,1\}} \left(p_{kct} - w_{kct}\right) s_{kct},$$

[8]Note that we omit the means of all Gumbel distributed random utility terms, ϵ_{ijkt}, in the following. It is equal to the Euler-Mascheroni constant for all terms involving expectations of random utility terms, and will thus not affect choices.

where we will take $p_{kct} = \bar{p}_t$, that is, equal to a binding price ceiling chosen by the regulator for a given active ingredient-dosage combination, and where w_{kct} is the wholesale price of drug k in pharmacy c at t. We take the wholesale prices as given for now, and return to their determination when discussing the behavior of producers in the next section. We now denote by $m_{kct} \equiv \bar{p}_t - w_{kct}$ the product price-cost margin, where w_{kct} allows wholesale price discrimination across pharmacy chains.

Remark that we implicitly assume that both margins are positive, such that pharmacy chains accept both procurement channels. Necessary first order conditions for an interior solution for the θ's are

$$0 = \frac{\partial \pi_{ct}}{\partial \theta_{ct}^0} = \frac{\partial \pi_{ct}}{\partial \theta_{ct}^1}. \tag{3.2}$$

For θ_{ct}^0, this is

$$\begin{aligned} 0 &= \sum\nolimits_k m_{kct} \frac{\partial s_{kct}}{\partial \theta_{ct}^0} = \sum\nolimits_k m_{kct} \int \frac{\partial}{\partial \theta_{ct}^0} \left[s_{ict} s_{ikt|c} \right] dF(\alpha_{it}) \\ &= \int \sum\nolimits_k m_{kct} [\underbrace{\frac{\partial s_{ikt|c}}{\partial \theta_{ct}^0}}_{\substack{\text{change in} \\ \text{probability to} \\ \text{choose } k \text{ in chain } c}} \underbrace{s_{ict}}_{\substack{\text{probability} \\ \text{to choose} \\ \text{chain } c}} + \underbrace{s_{ikt|c}}_{\substack{\text{probability} \\ \text{to choose } k \\ \text{in chain } c}} \underbrace{\frac{\partial s_{ict}}{\partial \theta_{ct}^0}}_{\substack{\text{change in} \\ \text{probability to} \\ \text{choose chain } c}}] dF(\alpha_{it}), \end{aligned}$$

which has the interpretation that a marginal increase in θ_{ct}^0 will have two effects. The first term shows the potential increase in profit through higher sales of the more profitable good 0 and lower sales of the less profitable good 1, due to good 0 being more often the only option of the consumer; while the second term is the profit loss from a loss in market share, due to chain c's less attractive policy, from the consumer's point of view, that is, more often being stocked out of the other good.

As

$$\begin{aligned} \frac{\partial s_{ikt|c}}{\partial \theta_{ct}^{k'}} &= 1_{\{k=k'\}} - s_{ikt|c,B}, \text{ and} \\ \frac{\partial s_{ict}}{\partial \theta_{ct}^k} &= \left[\alpha_{ikct} - \ln \left(\sum\nolimits_k e^{\alpha_{ikct}} \right) \right] s_{ict}(1 - s_{ict}) \leq 0, \end{aligned}$$

using the fact that

$$\begin{aligned} &\frac{\partial s_{ikt|c}}{\partial \theta_{ct}^0} s_{ict} + s_{ikt|c} \frac{\partial s_{ict}}{\partial \theta_{ct}^0} \\ = \; &\left(1_{\{k=0\}} - s_{ikt|c,B} \right) s_{ict} + s_{ikt|c} \left[\alpha_{i0ct} - \ln \left(\sum\nolimits_k e^{\alpha_{ikct}} \right) \right] (1 - s_{ict}) s_{ict}, \end{aligned}$$

we obtain that the first order condition for optimal θ^0_{ct} implies

$$\frac{m_{0ct}}{m_{1ct}} = \frac{\int s_{i1t|c,B}s_{ict} + s_{i1t|c}\left[\ln\left(\sum_k e^{\alpha_{ikct}}\right) - \alpha_{i0ct}\right](1-s_{ict})s_{ict}dF(\alpha_{it})}{\int s_{i1t|c,B}s_{ict} - s_{i0t|c}\left[\ln\left(\sum_k e^{\alpha_{ikct}}\right) - \alpha_{i0ct}\right](1-s_{ict})s_{ict}dF(\alpha_{it})} \tag{3.3}$$

because $1 - s_{i0t|c,B} = s_{i1t|c,B}$ and $1 - s_{i0t|c} = s_{i1t|c}$.

Similarly, the first order condition with respect to θ^1_{ct} can be written

$$\frac{m_{1ct}}{m_{0ct}} = \frac{\int s_{i0t|c,B}s_{ict} + s_{0t|c}\left[\ln\left(\sum_k e^{\alpha_{ikct}}\right) - \alpha_{i1ct}\right](1-s_{ict})s_{ict}dF(\alpha_{it})}{\int s_{i0t|c,B}s_{ict} - s_{1t|c}\left[\ln\left(\sum_k e^{\alpha_{ikct}}\right) - \alpha_{i1ct}\right](1-s_{ict})s_{ict}dF(\alpha_{it})}. \tag{3.4}$$

We can see that only one of the first order condition will be satisfied. Indeed, as $1 - s_{i0t|c} = s_{i1t|c}$,

$$\begin{aligned}
& s_{i1t|c,B}s_{ict} + s_{i1t|c}\left[\ln\left(\sum\nolimits_k e^{\alpha_{ikct}}\right) - \alpha_{i0ct}\right](1-s_{ict})s_{ict} \\
= \; & s_{i1t|c,B}s_{ict} - s_{i0t|c}\left[\ln\left(\sum\nolimits_k e^{\alpha_{ikct}}\right) - \alpha_{i0ct}\right](1-s_{ict})s_{ict} \\
& + \left[\ln\left(\sum\nolimits_k e^{\alpha_{ikct}}\right) - \alpha_{i0ct}\right](1-s_{ict})s_{ict} \\
> \; & s_{i1t|c,B}s_{ict} - s_{i0t|c}\left[\ln\left(\sum\nolimits_k e^{\alpha_{ikct}}\right) - \alpha_{i0ct}\right](1-s_{ict})s_{ict},
\end{aligned}$$

and similarly

$$\begin{aligned}
& s_{i0t|c,B}s_{ict} + s_{0t|c}\left[\ln\left(\sum\nolimits_k e^{\alpha_{ikct}}\right) - \alpha_{i1ct}\right](1-s_{ict})s_{ict} \\
> \; & s_{i0t|c,B}s_{ict} - s_{1t|c}\left[\ln\left(\sum\nolimits_k e^{\alpha_{ikct}}\right) - \alpha_{i1ct}\right](1-s_{ict})s_{ict}.
\end{aligned}$$

Thus, Equation (3.3) cannot be true if $m_{1ct} > m_{0ct}$, and Equation (3.4) cannot be true if $m_{1ct} < m_{0ct}$.

Considering the case where $m_{1ct} < m_{0ct}$, there is then no interior solution for θ^1_{ct}, and thus we will have $\theta^1_{ct} = 0$, meaning that the pharmacy chain never proposes only the drug with lower margin. Then θ^0_{ct} is solution of Equation (3.3). The intuitive explanation is that when the chain increases the probability of only having the lower margin product available, profits are hurt both due to the opportunity cost of consumers who would otherwise have bought the high margin product when both were available, and the loss of market share due to offering on average less variety.

Assuming for now that the margins are larger for parallel imports (good 0) for all chains, we can set the probability of proposing direct imports alone, θ^1_{ct}, to zero for all c in the following exposition and, denote

$$\theta_{ct} \equiv 1 - \theta^0_{ct},$$

the probability that both goods are available in pharmacy chain c. We can now express the expected inclusive value as

$$I_{ict} = (1-\theta_{ct})\alpha_{i0ct} + \theta_{ct}\ln\left(\sum\nolimits_k e^{\alpha_{ikct}}\right) = \alpha_{i0ct} + \theta_{ct}\delta_{ict},$$

where

$$\delta_{ict} \equiv \ln\left(1+e^{\Delta\alpha_{ict}}\right),$$

where $\Delta\alpha_{ict} = \alpha_{i1ct} - \alpha_{i0ct}$. Thus, δ_{ict} is the incremental expected utility from having both drugs available to choose from, as opposed to only parallel import. Furthermore, let

$$\rho_{ict} \equiv s_{i1t|c,B},$$

that is, the probability that consumer i chooses the direct imported variety in chain c at t when both are available.[9] It will be helpful to note that $\delta_{ict} = -\ln(1-\rho_{ict})$, which has the natural interpretation that individual i's incremental utility from having both goods available, is increasing in the probability that she will choose the direct imported variety when both are available.

We can now rewrite

$$s_{i1t|c}\left(\theta_{ct}\right) = \theta_{ct}\rho_{ict},$$
$$s_{i0t|c}\left(\theta_{ct}\right) = 1-\theta_{ct}\rho_{ict},$$

and

$$s_{ct}\left(\boldsymbol{\theta}_t\right) = \int \frac{e^{\alpha_{i0ct}+\theta_{ct}\delta_{ict}}}{\sum_c e^{\alpha_{i0ct}+\theta_{ct}\delta_{ict}}} dF(\alpha_{it}) = \int s_{ict}(\boldsymbol{\theta}_t) dF(\alpha_{it}),$$

where $\boldsymbol{\theta}_t \equiv (\theta_{0t}, \cdots, \theta_{Ct})'$ is the vector consisting of the probability that both goods are available for each chain.

The profit maximization problem for each chain c at t is now given by the program

$$\max_{0\leq\theta_{ct}\leq 1} \pi_{ct},$$

[9] Keep in mind that ρ could be a function of characteristics of the drugs through its dependence on the deterministic utility components α_{ikct}

where the optimality conditions are

$$\frac{\partial \pi_{ct}}{\partial \theta_{ct}} \begin{cases} \leq 0 & \text{if } \theta_{ct} = 0, \\ = 0 & \text{if } 0 < \theta_{ct} < 1, \\ \geq 0 & \text{if } \theta_{ct} = 1. \end{cases} \tag{3.5}$$

The derivative of profits with respect to the probability that both goods are available is

$$\frac{\partial \pi_{ct}}{\partial \theta_{ct}} = m_{0ct} \frac{\partial s_{0ct}}{\partial \theta_{ct}} + m_{1ct} \frac{\partial s_{1ct}}{\partial \theta_{ct}}, \tag{3.6}$$

where the derivatives of shares with respect to θ_{ct} are

$$\frac{\partial s_{0ct}}{\partial \theta_{ct}} = \int \big(-\rho_{ict} s_{ict} + (1 - \theta_{ct}\rho_{ict})\delta_{ict} s_{ict}(1 - s_{ict})\big) dF(\alpha_{it}), \text{ and}$$
$$\frac{\partial s_{1ct}}{\partial \theta_{ct}} = \int \big(\rho_{ict} s_{ict} + \theta_{ct}\rho_{ict}\delta_{ict} s_{ict}(1 - s_{ict})\big) dF(\alpha_{it}).$$

From these expressions, we see that there are basically two effects from increasing the probability that both products are available. To give a better sense of how the model works, we will discuss these effects first from the point of view of an individual i. The first effect is a change in the conditional choice probability of the product—that is, the choice probability given that the individual has chosen pharmacy chain c—weighted by the probability s_{ict} that chain c is chosen by individual i in the first place. This is *negative* for parallel imports, as it reduces the number of times where it is the only product available, while it is *positive* for the direct import, as it reduces the number of times that it is stocked out. The second effect is a change in the probability of choosing chain c, weighted by individual i's conditional probability of choosing the product. This effect is *positive* for both products, since the incremental expected utility of having both drugs available, δ_{ict}, is positive for all individuals, i.e., more individuals will choose chain c when the variety is greater. The *aggregate* effect then depends on the distribution of individual tastes in the population. As an example, let us consider a *decrease* in θ_{ct} to induce more consumers to choose the parallel imported variety. This will have a larger impact on the relative shares of the goods within pharmacy chain c, when consumers have a strong preference for the direct imported variety on average, and even more so if this correlates positively with the probability of choosing chain c in the population. However, if people on average have a strong preference for the direct imported variety, the incremental utility δ_{ict} will tend to be large, thus implying a stronger substitution away from chain c. This negative aggregate effect will be weaker if people have strong preferences for a specific pharmacy, such that s_{ict} tends to be either very high or very low, and also if there is a positive correlation between the taste for direct imports and chain c. From this, we can see that the distribution of tastes in the population will be central in the decision of a pharmacy chain on how much to constrain supply.

From Equation (3.6), together with the previous discussion, we can see that an increase in m_{0ct}—the margin on parallel imports—will lead to decrease in θ_{ct}, as $\frac{\partial s_{0ct}}{\partial \theta_{ct}} < 0$, while the opposite holds for an increase in m_{1ct}. We can also see that only the relative margin matters for the decision of the pharmacy chain, though the relative margin will depend on both the wholesale prices and the price ceiling.

We assume that each pharmacy chain c sets θ_{ct} to maximize its profits conditional on the wholesale prices it faces, while taking the choices of all other pharmacy chains as given. The equilibrium in each market t will then be given by the vector $\boldsymbol{\theta}^*_t(\mathbf{w}_{0t}, \mathbf{w}_{1t})$, which consists of the elements $\theta^*_{ct}(\mathbf{w}_{0t}, \mathbf{w}_{1t})$, that is, the vector of equilibrium θ in t as a function of the wholesale prices of direct and parallel import in the market. Note that we have suppressed the dependence on $\overline{p}_t$ for convenience. The vector $\boldsymbol{\theta}^*_t(\mathbf{w}_{0t}, \mathbf{w}_{1t})$ is defined such that Equation (3.5) is satisfied simultaneously for all pharmacy chains at t.

3.3 Upstream Producers and Importers

We now model the upstream behavior of both originator producers and of parallel importers. We assume that upstream firms and pharmacy chains bargain over wholesale prices, leading to the *Nash-in-Nash bargaining* model, first proposed by Horn and Wolinsky (1988). As documented by Brekke and Straume (2015), the assumption that they bargain over a piece-rate price can be defended on the grounds of the prohibition against side-payments in contracts between producers and wholesalers in the Norwegian pharmaceutical market.

3.3.1 Producer Behavior

The total sales of the originator producer of a drug in a given market (country) come from two channels: the direct import channel of its product (good 1) to all chains c in that market, and the parallel imports of the same patented active ingredient (good 0) by all chains c. Here, we hypothesize a fully rational producer, internalizing the sales in a given market induced by parallel trade with other countries.

Thus, letting $\boldsymbol{\theta}^*_t \equiv \boldsymbol{\theta}^*_t(\mathbf{w}_{0t}, \mathbf{w}_{1t})$ in the following for notational brevity, the profits of the producer are given by

$$\Pi_t(\mathbf{w}_{1t}) = \sum\nolimits_c \left[(w_{1ct} - c_t) s_{1ct}(\boldsymbol{\theta}^*_t) + (p^I_{1t} - c_t) s_{0ct}(\boldsymbol{\theta}^*_t) \right],$$

where all w_{1ct} are the wholesale prices charged for direct imported drugs to chain c at time t, w_{0ct} the wholesale prices for parallel imports, c_t the marginal cost of production, and p^I_{1t} the producer price in the source country of the parallel importer.

We assume that in each pairwise negotiation with the pharmacy chains, the producer and pharmacy chain c sets wholesale prices to maximize the Nash-product

$$(\Pi_t - \Pi_{-c,t})^{b_{1c}}(\pi_{ct} - \pi_{-1,ct})^{1-b_{1c}}, \tag{3.7}$$

where b_{1c} is the bargaining weight of the producer when negotiating with chain c, $\Pi_{-c,t}$ is the producer's profit in absence of an agreement with pharmacy chain c, and $\pi_{-1,ct}$ is likewise pharmacy chain c's profit in absence of an agreement with the producer. We make the assumption that all contracts remain the same if another negotiation fails, that each bargaining pair observes the wholesale prices of parallel imports to each pharmacy chain $\mathbf{w}_{0t} = (w_{01t}, w_{02t}, \cdots, w_{0Ct})$. These assumptions are commonplace in the literature estimating structural bargaining models (see e.g., Gowrisankaran et al. (2015) and Crawford and Yurukoglu (2012), Ho and Lee (2015)). The first order condition for a solution to Equation (3.7), is

$$b_{1c}\frac{\partial \Pi_t / \partial w_{1ct}}{\Pi_t - \Pi_{-c,t}} + (1 - b_{1c})\frac{\partial \pi_{ct} / \partial w_{1ct}}{\pi_{ct} - \pi_{-1,ct}} = 0, \tag{3.8}$$

In maximizing the Nash-product, there will be an effect on the producer's profit from how changes in wholesale prices affect the equilibrium $\boldsymbol{\theta}_t^*(\mathbf{w}_{0t}, \mathbf{w}_{1t})$ in the next stage of the game.

Denote the net value of agreement for the direct importer as $\Delta_{1c}\Pi_t \equiv \Pi_t - \Pi_{-c,t}$ and chain c as $\Delta_{1c}\pi_{ct} \equiv \pi_{ct} - \pi_{-1,ct}$. The derivative of the direct importer's profits with respect to the wholesale price is

$$\begin{aligned}
\frac{\partial \Pi_t}{\partial w_{1ct}} &= s_{1ct}\left(\boldsymbol{\theta}_t^*\right) + \sum\nolimits_{\tilde{c}}\left[(w_{1\tilde{c}t} - c_t)\frac{\partial s_{1\tilde{c}t}}{\partial w_{1ct}}\left(\boldsymbol{\theta}_t^*\right) + (p_{1t}^I - c_t)\frac{\partial s_{0\tilde{c}t}}{\partial w_{1ct}}\left(\boldsymbol{\theta}_t^*\right)\right] \\
&= s_{1ct}\left(\boldsymbol{\theta}_t^*\right) + \sum\nolimits_{\tilde{c}}\left[w_{1\tilde{c}t}\frac{\partial s_{1\tilde{c}t}}{\partial w_{1ct}}\left(\boldsymbol{\theta}_t^*\right) + p_{1t}^I\frac{\partial s_{0\tilde{c}t}}{\partial w_{1ct}}\left(\boldsymbol{\theta}_t^*\right)\right] \\
&= s_{1ct}\left(\boldsymbol{\theta}_t^*\right) + \boldsymbol{w}_{1t}'\frac{\partial \boldsymbol{s}_{1t}}{\partial w_{1ct}}\left(\boldsymbol{\theta}_t^*\right) + p_{1t}^I\frac{\partial \boldsymbol{s}_{0t}}{\partial w_{1ct}}\left(\boldsymbol{\theta}_t^*\right),
\end{aligned}$$

where $\boldsymbol{w}_{1t}$ is the (column) vector of direct import wholesale prices to the pharmacy chains, and $\boldsymbol{s}_{kt}$ is the (column) vector of market shares of variant k (parallel or direct import) in the pharmacy chains and where the second equality follows from the assumption of inelastic aggregate demand, such that $\sum_{\tilde{c}}(\frac{\partial s_{1\tilde{c}t}}{\partial w_{1ct}} + \frac{\partial s_{0\tilde{c}t}}{\partial w_{1ct}}) = 0$.

Consequently, the gradient of the direct importers profits with respect to the vector of wholesale prices $\boldsymbol{w}_{1t}$ is

$$\frac{\partial \Pi_t}{\partial \boldsymbol{w}_{1t}} = \boldsymbol{s}_{1t} + \frac{\partial \boldsymbol{s}_{1t}'}{\partial \boldsymbol{w}_{1t}}\boldsymbol{w}_{1t} + \frac{\partial \boldsymbol{s}_{0t}'}{\partial \boldsymbol{w}_{1t}}p_{1t}^I,$$

where we have dropped the dependence on $\boldsymbol{\theta}_t^*$ for notational brevity.

Note that the derivatives of market shares with respect to wholesale prices follows from the chain rule and the implicit function theorem governing the change in equilibrium $\boldsymbol{\theta}_t$ when wholesale prices change, i.e.,

$$\begin{aligned}\frac{\partial \boldsymbol{s}_{kt}}{\partial \boldsymbol{w}'_{1t}} &= \left(\frac{\partial \boldsymbol{s}_{kt}}{\partial \boldsymbol{\theta}'_t}\bigg|_{\boldsymbol{\theta}_t=\boldsymbol{\theta}^*_t} \right) \frac{\partial \boldsymbol{\theta}^*_t}{\partial \boldsymbol{w}'_{1t}} \\ &= \left(\frac{\partial \boldsymbol{s}_{kt}}{\partial \boldsymbol{\theta}'_t}\bigg|_{\boldsymbol{\theta}_t=\boldsymbol{\theta}^*_t} \right) \left(-\frac{\partial \mathbf{F}_{\boldsymbol{\theta},t}}{\partial \boldsymbol{\theta}'_t}\bigg|_{\boldsymbol{\theta}_t=\boldsymbol{\theta}^*_t} \right)^{-1} \frac{\partial \mathbf{F}_{\boldsymbol{\theta},t}}{\partial \mathbf{w}'_{1t}}\bigg|_{\boldsymbol{\theta}_t=\boldsymbol{\theta}^*_t},\end{aligned}$$

which shows that the change in a given market share, s_{kct}, caused by the change in a given wholesale price, $w_{1\tilde{c}t}$, will depend on the change in the full vector of θ's following from the change in the Nash equilibrium in the competition between chains.

The derivative of chain c's profits with respect to the wholesale price w_{1ct} is

$$\begin{aligned}\frac{\partial \pi_{ct}}{\partial w_{1ct}} &= -s_{1ct} + (\bar{p}_t - w_{1ct}) \frac{\partial s_{1ct}}{\partial w_{1ct}} + (\bar{p}_t - w_{0ct}) \frac{\partial s_{0ct}}{\partial w_{1ct}} \\ &= -s_{1ct} + \left[m_{1ct} \left(\frac{\partial s_{1ct}}{\partial \boldsymbol{\theta}'_t}\bigg|_{\boldsymbol{\theta}_t=\boldsymbol{\theta}^*_t} \right) + m_{0ct} \left(\frac{\partial s_{0ct}}{\partial \boldsymbol{\theta}'_t}\bigg|_{\boldsymbol{\theta}_t=\boldsymbol{\theta}^*_t} \right) \right] \frac{\partial \boldsymbol{\theta}^*_t}{\partial w_{1ct}},\end{aligned}$$

where the element of the vector in the square brackets which correspond to the derivative of chain c's profit with respect to θ_c will be zero from the envelope theorem. If some elements of $\boldsymbol{\theta}^*_t$ is not interior, the corresponding elements in the last vector in the expression will be zero.

We can then rewrite Equation (3.8) governing the solution to the bargaining between the direct importer and chain c as

$$s_{1ct} + \boldsymbol{w}'_{1t} \frac{\partial \boldsymbol{s}_{1t}}{\partial w_{1ct}} + p^I_{1t} \frac{\partial \boldsymbol{s}_{0t}}{\partial w_{1ct}} = \frac{1 - b_{1c}}{b_{1c}} \frac{\Delta_{1c}\Pi_t}{\Delta_{1c}\pi_{ct}} \left(s_{1ct} - m_{1ct} \frac{\partial s_{1ct}}{\partial w_{1ct}} - m_{0ct} \frac{\partial s_{0ct}}{\partial w_{1ct}} \right), \tag{3.9}$$

which shows that the direct importer considers the change in all the shares in the market through the change in equilibrium θ , while the pharmacy chain only considers the change in their own shares. The expression in parentheses on the right hand side is the (negative of) loss in profits to chain c from a change in the direct import wholesale price, which will depend on how much is lost in direct import sale from the marginal change in equilibrium $\boldsymbol{\theta}^*_t$, and how much is gained in parallel import sale. The larger the relative bargaining power, $\frac{1-b_{1c}}{b_{1c}}$, of the chain is, and the larger the relative net value of agreement for the direct importer, $\Delta_{1c}\Pi_t/\Delta_{1c}\pi_{ct}$, the larger weight will be given to the change in profits for the pharmacy chain from a change in the wholesale price.

Note that $\Delta_{1c}\Pi_t = \Pi_t - \Pi_{-c,t}$ and $\Delta_{1c}\pi_{ct} = \pi_{ct} - \pi_{-1,ct}$ are the net values of an agreement for the producer and chain c respectively. Letting $s_{jrt\backslash 1c}$ denote the share of chain r's product j in t when direct imports are not available at chain c, we can express the net value for the producer, suppressing argument

$\boldsymbol{\theta}^*$ in shares, as

$$\begin{aligned}\Delta_{1c}\Pi_t &= \sum\nolimits_{\tilde{c}}[(w_{1\tilde{c}t} - c_t)s_{1\tilde{c}t} + (p^I_{1t} - c_t)s_{0\tilde{c}t}] \\ &\quad - \sum\nolimits_{\tilde{c}}[(w_{1\tilde{c}t} - c_t)s_{1\tilde{c}t\backslash 1c} + (p^I_{1t} - c_t)s_{0\tilde{c}t\backslash 1c}] \\ &= \sum\nolimits_{\tilde{c}}\left(w_{1\tilde{c}t}\Delta_{1c}s_{1\tilde{c}t} + p^I_{1t}\Delta_{1c}s_{0\tilde{c}t}\right),\end{aligned}$$

because $s_{jct\backslash 1c} = 0$, and denoting $\Delta_{1c}s_{j\tilde{c}t} \equiv s_{j\tilde{c}t} - s_{j\tilde{c}t\backslash 1c}$, that is, the difference in share of product j in chain $\tilde{c}$ between the case of agreement and disagreement in the negotiations between the producer and chain c. Since aggregate demand is assumed to be constant, particularly since consumer prices are fixed at $\bar{p}_t$, such that market shares sum to one both in the case of agreement and disagreement, the cost of production is immaterial to the change in producer profit. Moreover, because of price regulation, we have that the producer takes as given the price (p^I_{1t}) obtained on sales in source countries for parallel imports.

Similarly, the net value for the chain is

$$\Delta_{1c}\pi_{ct} = (\bar{p}_t - w_{1ct})s_{1ct} + (\bar{p}_t - w_{0ct})\Delta_{1c}s_{0ct},$$

where the first term simply evaluates to $(\bar{p}_t - w_{1ct})s_{1ct}$. When the shape of demand is identified, it is possible to calculate the differences in shares, $\Delta_{1c}s_{j\tilde{c}t}$, using the estimated demand system. Note that this will also depend on the change in optimal θ for each pharmacy chain.

In order to obtain how wholesale prices affect the equilibrium $\boldsymbol{\theta}^*_t(\mathbf{w}_{0t}, \mathbf{w}_{1t})$, let $\mathbf{F}_{\boldsymbol{\theta},t}$ denote the vector of derivatives of pharmacy chain profit with respect to direct import availability at time t, i.e.

$$\mathbf{F}_{\boldsymbol{\theta},t} \equiv \left(\tfrac{\partial \pi_{1t}}{\partial \theta_{1t}}, \tfrac{\partial \pi_{2t}}{\partial \theta_{2t}}, \cdots, \tfrac{\partial \pi_{Ct}}{\partial \theta_{Ct}}\right)'$$

The derivative of $\boldsymbol{\theta}^*_t(\mathbf{w}_{0t}, \mathbf{w}_{1t})$ with respect to its argument can be found by the implicit function theorem. Implicit differentiation of the system of first order conditions $\mathbf{F}_{\boldsymbol{\theta},t} = \mathbf{0}$ yields

$$\left.\frac{\partial \mathbf{F}_{\boldsymbol{\theta},t}}{\partial \boldsymbol{\theta}'_t}\right|_{\boldsymbol{\theta}_t = \boldsymbol{\theta}^*_t} \mathrm{d}\boldsymbol{\theta}_t + \left.\frac{\partial \mathbf{F}_{\boldsymbol{\theta},t}}{\partial \mathbf{w}'_{1t}}\right|_{\boldsymbol{\theta}_t = \boldsymbol{\theta}^*_t} \mathrm{d}\mathbf{w}_{1t} = \mathbf{0}.$$

The Jacobian of $\boldsymbol{\theta}^*_t$ with respect to $\mathbf{w}_{1t}$ is then

$$\frac{\partial \boldsymbol{\theta}^*_t}{\partial \mathbf{w}'_{1t}} = -\left(\left.\frac{\partial \mathbf{F}_{\boldsymbol{\theta},t}}{\partial \boldsymbol{\theta}'_t}\right|_{\boldsymbol{\theta}_t = \boldsymbol{\theta}^*_t}\right)^{-1} \left.\frac{\partial \mathbf{F}_{\boldsymbol{\theta},t}}{\partial \mathbf{w}'_{1t}}\right|_{\boldsymbol{\theta}_t = \boldsymbol{\theta}^*_t}.$$

Recalling that $\frac{\partial \pi_{ct}}{\partial \theta_{ct}} = (\bar{p}_t - w_{0ct})\frac{\partial s_{0ct}}{\partial \theta_{ct}}(\boldsymbol{\theta}_t) + (\bar{p}_t - w_{1ct})\frac{\partial s_{1ct}}{\partial \theta_{ct}}(\boldsymbol{\theta}_t)$, we have that

$$\frac{\partial \mathbf{F}_{\boldsymbol{\theta},t}}{\partial \mathbf{w}'_{1t}} = -\begin{pmatrix} \frac{\partial s_{11t}}{\partial \theta_{1t}} & 0 & \cdots & 0 \\ 0 & \frac{\partial s_{12t}}{\partial \theta_{2t}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\partial s_{1Ct}}{\partial \theta_{Ct}} \end{pmatrix},$$

while

$$\frac{\partial \mathbf{F}_{\boldsymbol{\theta},t}}{\partial \boldsymbol{\theta}'_t} = \begin{pmatrix} \sum_k m_{k1t}\frac{\partial^2 s_{k1t}}{\partial \theta_{1t}^2} & \sum_k m_{k1t}\frac{\partial^2 s_{k1t}}{\partial \theta_{1t}\partial \theta_{2t}} & \cdots & \sum_k m_{k1t}\frac{\partial^2 s_{k1t}}{\partial \theta_{1t}\partial \theta_{Ct}} \\ \sum_k m_{k2t}\frac{\partial^2 s_{k2t}}{\partial \theta_{2t}\partial \theta_{1t}} & \sum_k m_{k2t}\frac{\partial^2 s_{k2t}}{\partial \theta_{2t}^2} & \cdots & \sum_k m_{k2t}\frac{\partial^2 s_{k2t}}{\partial \theta_{2t}\partial \theta_{Ct}} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_k m_{kCt}\frac{\partial^2 s_{kCt}}{\partial \theta_{Ct}\partial \theta_{1t}} & \sum_k m_{kCt}\frac{\partial^2 s_{kCt}}{\partial \theta_{Ct}\partial \theta_{2t}} & \cdots & \sum_k m_{kCt}\frac{\partial^2 s_{kCt}}{\partial \theta_{Ct}^2} \end{pmatrix}$$

Note that in the case where the producer has all the bargaining weight, that is, $b_{1c} = 1$, Equation (3.8) reduces to the first order condition for an optimal take-it-or-leave-it contract on w_{1ct} for the producer, while in the case of $b_{1c} = 0$, it can be rewritten as the condition for an optimal contract proposed by the chain.

Once the demand shape is identified, together with the optimal behavior of pharmacy chains, the system (3.9) has one equation per molecule-pharmacy chain-period, with in principle one unknown parameter b_{1c}. The system also depends on the exogenous opportunity cost of drugs for the producer –the price earned in the foreign country p^I_{1t}– but not the producer cost c_t. The reason is that, with the assumption of an exogenous, fixed market size, the producer will anyway expend the cost of supplying these drugs. More specifically, this assumption means that $\sum_c (s_{0ct} + s_{1ct}) = 1$, with the implication that the derivatives sum to zero. In such a market, and given the expression for producer profits above, the relevant opportunity cost for the producer is the price it obtains in source countries. If p^I_{1t} is known, the system of equations (3.9) allows us to identify the bargaining weight of each pharmacy chain.

When the condition in Equation (3.9) holds for all c, we have a Nash-in-Nash solution for the bargaining between direct importer and each of the pharmacy chain. The full Nash-in-Nash solution is obtained when we also consider the conditions for bargaining between the parallel importer and each of the pharmacy chains as described below.

3.3.2 Parallel Importers Behavior

We now consider the parallel importer's profits from its total sales of a drug in the importing market. Of course, parallel importers can only make a profit if the imported drug price in the source country is lower than the maximum retail price, that is if $p_{0ct}^I < \bar{p}_t$. This profit is given by

$$\Pi_t^{PI} = \sum\nolimits_c (w_{0ct} - p_{0ct}^I) s_{0ct}(\boldsymbol{\theta}_t^*),$$

where w_{0ct} is the wholesale price paid for parallel imported drugs by chain c at time t, and p_{0ct}^I is the price that the importer has to pay for the drug in the source country, which we allow to vary across chains c for full generality because each chain may require different source countries. The wholesale price that parallel importers get from the pharmacy chains must be in the interval $\left[p_{0ct}^I, \bar{p}_t\right]$.

We assume that the parallel importer bargains over the wholesale price with each pharmacy chain c, where they take as given the negotiated wholesale prices of originator products to each pharmacy chain $\mathbf{w}_{1t} = (w_{11t}, w_{12t}, \cdots, w_{1Ct})$. When bargaining over the wholesale prices charged to the chains, $\mathbf{w}_{0t}$, the parallel importer will also take into account how changes in these prices will affect the equilibrium $\boldsymbol{\theta}_t^*$. Similarly to Equation (3.8), the first order conditions for the solution to the Nash bargaining between each pharmacy chain c and the parallel importer is

$$b_{0c} \frac{\partial \Pi_t^{PI} / \partial w_{0ct}}{\Pi_t^{PI} - \underline{\Pi}_{ct}^{PI}} + (1 - b_{0c}) \frac{\partial \pi_{ct} / \partial w_{0ct}}{\pi_{ct} - \underline{\pi}_{0ct}} = 0, \tag{3.10}$$

which can be rewritten, largely following the same approach as above,

$$s_{0ct} + \left(\boldsymbol{w}_{0t} - \boldsymbol{p}_{0t}^I\right)' \frac{\partial \boldsymbol{s}_{0t}}{\partial w_{0ct}} = \frac{1 - b_{0c}}{b_{0c}} \frac{\Delta_{0c} \Pi_t^{PI}}{\Delta_{0c} \pi_{ct}} \left(s_{0ct} - m_{1ct} \frac{\partial s_{1ct}}{\partial w_{0ct}} - m_{0ct} \frac{\partial s_{0ct}}{\partial w_{0ct}} \right), \tag{3.11}$$

which follows from noting that the derivative of parallel importer profits with respect to wholesale price w_{0ct} is

$$\frac{\partial \Pi_t^{PI}}{\partial w_{0ct}} = s_{0ct} + \left(\boldsymbol{w}_{0t} - \boldsymbol{p}_{0t}^I\right)' \frac{\partial \boldsymbol{s}_{0t}}{\partial w_{0ct}}$$

Again, since wholesale prices are observed, one can use these optimality conditions to identify the parallel importers bargaining parameters b_{0c}, provided we observe or can model the prices at which imports are paid from the source country $\mathbf{p}_{0t}^I$.

4 Market Data, Econometrics and Empirical Results

4.1 Data and Descriptive Statistics

We estimate our model on the Norwegian market for Atorvastatin, which is a member of the statins drug class, used for lowering blood cholesterol. It is marketed by Pfizer under the trade name *Lipitor*. The patent expired towards the end of 2011, so it is on patent the whole period of our data from 2004 to 2007. The drug comes in four distinct strengths in the Norwegian market: Tablets with 10, 20, 40 and 80 milligram of the active ingredient. The prescription will decide which of these strengths the consumer can obtain at the pharmacy, though the pharmacy can freely propose substitution to a parallel imported alternative given that it has the same strength. Atorvastatin was used by roughly 140,000 individuals in 2004 and 2005, but the number of users dropped to about 100,000 in 2006 and 85,000 in 2007.[10] The explanation for this can largely be attributed to a change in the regulation of statin prescriptions introduced in June 2005. The regulation required that Simvastatin was to be prescribed for all new cases requiring statin treatment, while present users were to be put on treatment with Simvastatin within a year, unless medical considerations dictate otherwise.[11] The motivation for the regulation was to reduce expenditure for the Norwegian National Insurance Administration.

We combine data from several sources: Transaction data from the *Norwegian Directorate of Health* covering all purchases of reimbursable drugs by individuals in Norway, wholesaler registry data from the *Norwegian Institute of Public Health* on monthly wholesale prices of drug wholesalers in Norway, data on price regulation, substitutability and parallel marketing licenses from the *Norwegian Medicines Agency*, and data on aggregate wholesale prices in several countries from *IMS Health*. We thus have data on all purchases of Atorvastatin in Norway for the period 2004–2007, which amounts to about 1,4 million transactions. The transactions are performed by around 170,000 individuals, where a pseudo-ID for each individual allows us to use information on repeated choices. The demographic information on individuals is otherwise limited to age and gender. For each transaction, we know the price the pharmacy charges for the drug, the co-payment paid, the specific pharmacy at which the transaction happened, the number of packages bought, and the specific drug package.[12]

The supply side of the market for prescription drugs consists mainly of three large pharmacy retail chains, which are vertically integrated with each of their upstream wholesaler. The three largest chains, Apotek 1, Boots and Vitus, cover 85 % of all pharmacies, while public hospital pharmacies (6 %), a smaller retail chain (5 %), and independent pharmacies (4 %) make up the rest.[13]

Table 4.1 shows the yearly size of the Atorvastatin market in Norway in millions of *Defined Daily Doses* (DDD), segmented by the amount of active ingredient.[14] We have also calculated the share of DDD which are parallel imports within each segment. We see that for 10 and 20 mg., parallel imports were not present

[10]The population of Norway was roughly 4.6 million in this period.

[11]More details on this regulatory change can be found in Sakshaug et al. (2007).

[12]An example of a specific drug package is *Lipitor* with 40 mg of the active ingredient, containing 98 tablets, and imported by Farmagon from France.

[13]The shares are calculated from our own data and checked against data obtained from the Norwegian Medicines Agency. The numbers correspond exactly to official statistics reported by the Norwegian Medicines agency and the Norwegian Pharmacy Association.

[14]Our definition of the market includes direct purchases in pharmacies by individuals exclusively. Though there might be some usage of Atorvastatin in hospitals—for instance as part of statin treatment after heart attacks—the numbers in our data are virtually identical to official statistics (gathered to monitor drug utilization in Norway) on aggregate usage of Atorvastatin, which makes us conclude that this would represent a negligible share of sales.

Table 4.1: *Market size in million DDD, share of parallel imports, and price to consumers (Price) and wholesale prices (Wholesale) in NOK/DDD*

		2004	2005	2006	2007
	DDD (mill.)	16.36	15.10	9.13	4.61
	1.00	1.00	1.00	0.87	
10 mg	Share parallel	0.00	0.00	0.00	0.13
	Price	8.78	8.84	8.39	8.43
	Wholesale direct	6.21	6.20	5.86	5.86
	Wholesale parallel	-	-	-	4.42
	DDD (mill.)	34.15	34.99	22.14	12.07
	1.00	1.00	1.00	0.90	
20 mg	Share parallel	0.00	0.00	0.00	0.10
	Price	6.62	6.67	6.37	6.42
	Wholesale direct	4.74	4.74	4.52	4.53
	Wholesale parallel	-	-	-	3.15
	DDD (mill.)	23.78	31.22	26.42	29.32
	0.21	0.52	0.93	0.83	
40 mg	Share parallel	0.79	0.48	0.07	0.17
	Price	4.16	4.21	3.82	3.90
	Wholesale direct	3.00	3.01	2.71	2.76
	Wholesale parallel	2.91	2.93	2.87	2.03
	DDD (mill.)	12.03	20.12	27.38	35.69
	0.07	0.14	0.04	0.37	
80 mg	Share parallel	0.93	0.86	0.96	0.63
	Price	2.15	2.23	1.98	1.97
	Wholesale direct	1.55	1.60	1.40	1.39
	Wholesale parallel	1.52	1.50	1.38	1.35

before 2007. For 40 and 80 mg., parallel imports often cover a substantial share of the market, covering around 90% of the 80 mg. segment in 2004-2006. The reason for the differences in parallel import shares are likely due to a combination of differences in parallel export opportunities, differences in profitability across parallel import locations and differences in the relative price in the source country and Norway. Due to the lack of data on parallel trader behavior outside of Norway and the difficulty of determining the reason for why parallel imports are absent in some markets, we will focus solely on the markets where they are present for the part of our estimation regarding upstream producer and importer behavior.[15] There is also substantial variation between some of these years. The market size for 10 and 20 mg. decreases substantially over the sample period, while it stays at roughly the same level for 40 mg. and increases substantially for 80 mg. It seems likely that the large changes in the number of consumers underlying these figures will have an impact on the distribution of preferences in the market. We will allow the average taste for each

[15] As we are mostly interested in the behavior of retailers, and how this feeds into the behavior of the upstream firms, this is not a large problem. In the markets where parallel imports are absent, there is no scope for strategic behavior by the pharmacy chains. We do, however, include this part of the sample when estimating the consumer choice model, as it gives us more variation in the choice sets of the consumers over time to identify preference heterogeneity.

available drug to change across segments and time, which we return to when discussing the specification of the consumer choice model. The price to consumers reflects the regulatory price ceiling set by the Norwegian Medicines Agency, as all packages—both parallel and direct imports—are consistently priced at the price ceiling. From the wholesale prices, we see that the aggregate margin is larger for parallel imports in almost all cases, except for 40 mg. in 2006, which also corresponds to a large drop in sales of parallel imports (see Figure 4.1 below).

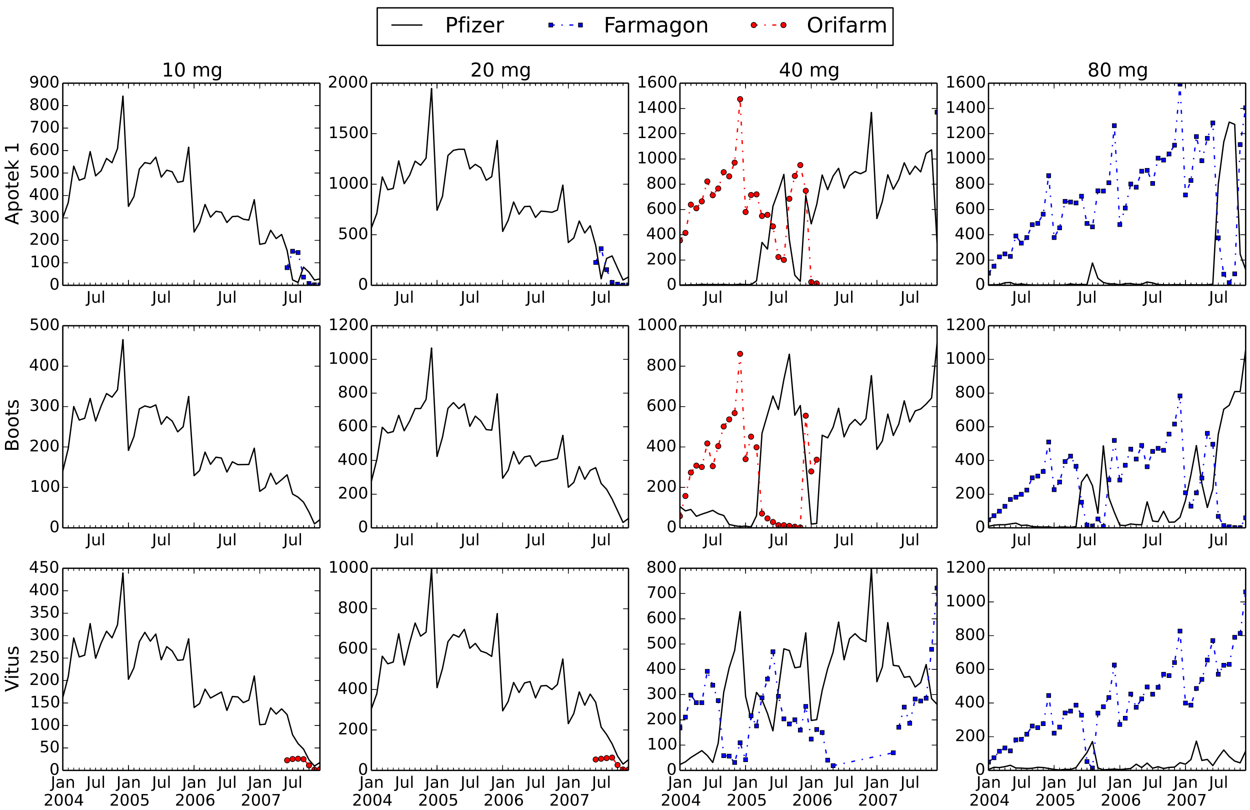
Table 4.2: *Drug packages and parallel import licensing*

Dose	Company	#Tablets	Source country	License Year
10 mg	Pfizer	100	-	-
	Farmagon	100	Czech Rep.	2006
	Orifarm	100	Poland	2006
20 mg	Pfizer	100	-	-
	Farmagon	100	Czech Rep.	2006
	Orifarm	100	Poland	2006
40 mg	Pfizer	100	-	-
	Farmagon	98	UK, France	2002, 2004
	Farmagon	100	Poland, Czech Rep.	2004, 2006
	Orifarm	98	UK	2002
80 mg	Pfizer	100	-	-
	Farmagon	98	UK, France	2002, 2004
	Farmagon	100	Czech Rep.	2006

Table 4.2 shows information about the specific packages sold in the Norwegian market for Atorvastatin in the sample period. The active upstream firms are Pfizer, Farmagon and Orifarm, where Pfizer holds the patent and is responsible for the direct imports, and Farmagon and Orifarm are parallel importers. The parallel importers have licenses to import from the United Kingdom, France, Czech Republic and Poland, where typically the licenses from the Eastern European countries are acquired in 2006. The underlying sales patterns in the data shows that the packages imported from Eastern Europe are only sold in 2007. Where several source countries are listed, the packages imported from the different countries are given the same ID in the national drug classification system, which means that they are identical. The parallel import process is such that the drugs will be repackaged by the parallel importer to be in accordance with nation specific guidelines on package labels, language and warnings, so the correct interpretation is that the packages are indiscernible after repackaging. In several of the cases, the parallel importers have license to import the package from two countries. In Figure 4.1, we show sales of each upstream company by month in 1000 DDD, displayed separately for each segment (amount of active ingredient) and pharmacy chain. Inspection of the data underlying this figure shows that parallel imports are exclusively from the Eastern European countries for 10 and 20 mg., where parallel imports enter in the second half of 2007. For 40 mg., parallel imports

are from the Western European countries until 2007, when it switches to Eastern European imports after a large drop in parallel imports in 2006. For 80 mg., parallel imports are exclusively from Western Europe until 2007, when about 10% of the parallel traded drugs are imported from Eastern Europe.

Figure 4.1: *Monthly sales in 1000 DDD of direct imports (Pfizer) and parallel imports (Farmagon or Orifarm) for each chain and dosage (mg)*

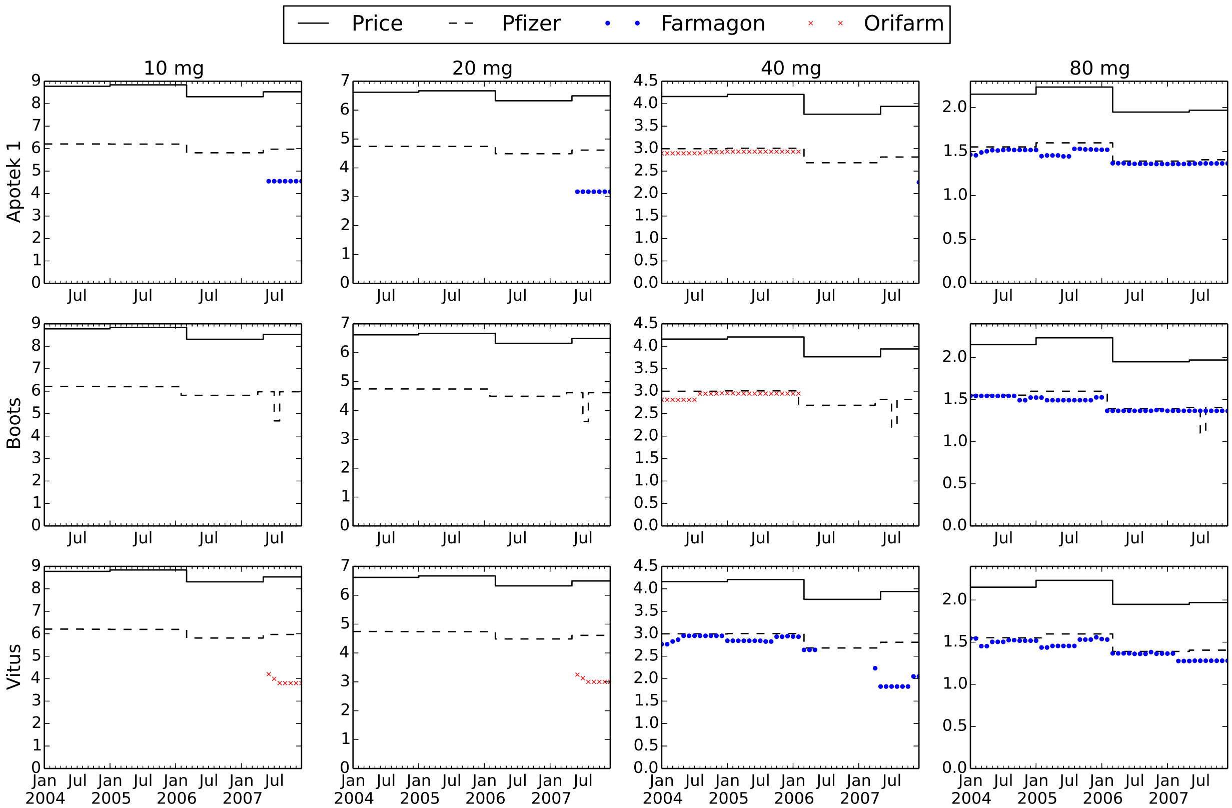


In Figure 4.2, we show consumer prices and wholesale prices for each upstream company by month, separately for each segment and pharmacy chain. Note that the consumer price is entirely decided by the price cap, which is binding for both the direct and parallel imported varieties. The wholesale prices of parallel importers are consistently lower than the direct import wholesale price—except in one instance—and often substantially so.

4.2 Econometric Identification and Estimation

Our structural model of demand and supply can be estimated using data on consumer choices between parallel trade and directly imported versions of a drug and data on the pharmacy retail chain margins or wholesale prices. Section 3.1 developed a consumer discrete choice model where consumers choose between pharmacy chains and direct versus parallel imported drugs. Our random utility model resembles a classic random coefficients logit model but in which random utilities depend on pharmacies strategic random as-

Figure 4.2: *Monthly price to consumers (Price), wholesale price of direct imports(Pfizer) and wholesale price of parallel imports (Farmagon or Orifarm) in NOK/DDD for each chain and dosage (mg)*



sortment choices of parallel trade versus direct imported drugs that are unobserved by the econometrician. Thus, the estimation of parameters governing individual choice also includes estimation of the assortment set probabilities in any given pharmacy chain that is possible using pharmacy chains optimal choice conditions together with the demand model. In a second step, we use the estimated parameters and choice set probabilities to identify the opportunity and marginal costs for both direct and parallel importers as well as bargaining parameters thanks to the vertical chain bargaining model.

4.2.1 Consumer and chain behavior

Observing individual choices, we first estimate the discrete choice demand model described in section 3.1 using the individual choice probability

$$s_{ijct} = s_{ict} s_{ijt|c} \tag{4.1}$$

where

$$
\begin{aligned}
s_{ict} &= \frac{e^{\alpha_{i0ct}+\theta_{ct}\delta_{ict}}}{\sum_{\tilde{c}} e^{\alpha_{i0\tilde{c}t}+\theta_{\tilde{c}t}\delta_{i\tilde{c}t}}} \quad , \quad s_{i1t|c} = \theta_{ct}\rho_{ict} \quad , \quad s_{i0t|c} = 1-\theta_{ct}\rho_{ict} \\
\text{with } \delta_{ict} &= \ln(1+\alpha_{i1ct}-\alpha_{i0ct}) \quad \text{and} \quad \rho_{ict} = \frac{e^{\alpha_{i1ct}}}{e^{\alpha_{i0ct}}+e^{\alpha_{i1ct}}}
\end{aligned}
$$

where α_{ijct} is individual i's mean utility from product type j bought at pharmacy chain c in market t. In particular, we specify individual utility as

$$\alpha_{ijct} = \alpha_{jct} + \nu_{ijct} \tag{4.2}$$

where α_{jct} is the average utility in market t for product j at chain c, common to all individuals and thus capturing any market fixed effects for each product and ν_{ijct} is the individual deviation from the mean utility for that good, representing the heterogeneity of consumers tastes. In practice, we will specify ν_{ijct} such that $\nu_{ijct} = \sigma_j\nu_i^j + \sigma_c\nu_i^c + \gamma_j\,\mathbf{1}_{\{age_i > age_{0.5}\}} + \delta_c^{g_i} + \sigma_c^{g_i}\nu_i^c$ where ν_i^k is individual i's taste characteristic for product characteristic k, here taken to be either product type j or a specific chain c. We assume that these taste characteristics are standard normal distributed, such that σ_k is a parameter measuring the scale of individual utility in deviations from the mean. Demographics enter by an indicator for whether individual i is above the median age in the sample (denoted $age_{0.5}$), which is allowed to affect the relative taste for product type j with utility γ_j. As only the difference across alternatives of this individual characteristic effect is identified, we normalize $\gamma_1 = 0$ and identify the relative effect of age on mean utility for parallel imports with γ_0. Finally, $\delta_c^{g_i}$ is a chain specific utility term and $\sigma_c^{g_i}$ is a chain specific utility dispersion term, conditional on individual i's group $g_i \in \mathcal{G}$, where $\mathcal{G}$ is the set of groups in the population. The group of individual i is unobserved, and is thus treated as a latent class during estimation. Note that $\sigma_c^{g_i}$ is interacted with individual i's unobserved taste characteristic for chain c, ν_i^c, which allows each group in $\mathcal{G}$ to have a different scale for the unobserved chain specific taste characteristics.

Additional restrictions on the parameters are that σ_j is the same for both direct and parallel imports, and $\sigma_c^{g_i}$ is the same for all chains, such that $\sigma_0 = \sigma_1 = \sigma_J$ and $\sigma_c^{g_i} = \sigma_C^{g_i}$, for all c, though we do allow the baseline scale of individual taste for chain c, σ_c, to differ across chains.[16] We allow there to be four latent classes, where one is arbitrarily chosen as the base group, $g = 0$ with $\delta_c^0 = 0$ and $\sigma_C^0 = 0$. Each group, g, has a population share τ_g, assumed to be the same across markets, which is introduced into the likelihood as a parameter to be estimated.

Denoting $\beta = (\sigma_J, \sigma_C, \gamma_0, \tau_g)$ the full vector of parameters governing heterogenous preferences, for some given mean preference parameters α_{0ct}, α_{1ct} and θ_{ct}, one can estimate β by maximum likelihood using the

[16]Initial attempts at different specifications did not seem to indicate large gains from more complex parameterizations.

likelihood of individual i's choice sequence $\mathcal{P}_i$

$$L_i(\beta;\alpha_{0ct},\alpha_{1ct},\theta_{ct}) = \sum\nolimits_{g\in\mathcal{G}} \tau_g \int \left(\prod\nolimits_{p\in\mathcal{P}_i} s_{i,j(p),c(p),t(p)}(\nu_i)\right) dF(\nu_i|\beta), \tag{4.3}$$

where $\mathcal{P}_i$ is the set of purchase events in which consumer i is involved, $j(p)$, $c(p)$, $t(p)$ denote respectively consumer i's choice of product and chain under purchase event p and the market in which purchase event p happens, such that $s_{i,j(p),c(p),t(p)}(\nu_i)$ is the individual i choice probability conditional on his unobserved heterogeneity ν_i and where $F(\nu_i|\beta)$ is the cumulative distribution function of ν_i.

Then, the mean parameters α_{0ct}, α_{1ct} and θ_{ct} are also identified adding the pharmacy chain optimality equilibrium conditions for θ_{ct} and pharmacy chains aggregate market shares conditions for all c, t

$$\theta^*_{ct} = \arg\max_{0\leq\theta_{ct}\leq 1} \pi_{ct}(m_{0ct}, m_{1ct}, \boldsymbol{\theta}^*_t, \alpha_{0ct}, \alpha_{1ct}) \tag{4.4}$$

$$\hat{s}_{jct} = s_{jct}(\boldsymbol{\theta}_t, \alpha_{0ct}, \alpha_{1ct}, \beta) \tag{4.5}$$

where $\hat{s}_{jct}$ are observed pharmacy chains aggregate market shares and where pharmacy chain profits depend on observed margins m_{0ct} and m_{1ct} as

$$\pi_{ct}(m_{0ct}, m_{1ct}, \boldsymbol{\theta}_t, \alpha_{0ct}, \alpha_{1ct}, \beta) = m_{0ct}s_{0ct}(\boldsymbol{\theta}_t, \alpha_{0ct}, \alpha_{1ct}, \beta) + m_{1ct}s_{1ct}(\boldsymbol{\theta}_t, \alpha_{0ct}, \alpha_{1ct}, \beta)$$

with aggregate shares given by[17]

$$s_{jct}(\boldsymbol{\theta}_t, \alpha_{0ct}, \alpha_{1ct}, \beta) = \sum\nolimits_i s_{ijct} = \sum\nolimits_i \sum\nolimits_{g\in\mathcal{G}} \tau_g \int s_{ijt|c}(\nu_i)s_{ict}(\nu_i)dF(\nu_i|\beta)$$

The pharmacy chains' incentives equation (4.4) can be described by the first order condition given in Equation (3.5) such that

$$\left.\frac{\partial\pi_{ct}}{\partial\theta_{ct}}\right|_{\boldsymbol{\theta}_t=\boldsymbol{\theta}^*_t} \begin{cases} = 0 & \text{if } 0 < \theta^*_{ct} < 1 \\ \leq 0 & \text{if } \theta^*_{ct} = 0 \\ \geq 0 & \text{if } \theta^*_{ct} = 1 \end{cases}, \tag{4.6}$$

Even if there are lot of parameters as we have $(\alpha_{0ct}, \alpha_{1ct}, \theta_{ct})$ for each chain-market combination, utilizing the fact that these parameters are common across consumers within each chain-market, they can be solved for by a simpler root-finding algorithm, conditional on the parameter vector β. The intuition is that, within

[17] For the sake of parsimony, we allow latent classes to govern only chain specific utility but not other utility parameters, which implies that the unobserved group of the individual only enters the probability of the individual choosing a given chain, s_{ict}.

each market t, these parameters can be set such that observed aggregate shares are equal to predicted aggregate shares both within and across chains, and such that the equilibrium conditions hold for $\frac{\partial \pi_{ct}}{\partial \theta_{ct}}$.

The identification conditions of our problem depends on some conditions explained below.

First, for a given vector (θ, β), we know from Berry et al. (1995) that one can solve for all $\alpha_{0ct}, \alpha_{1ct}$ such that (4.5) is true for all j, c, t. This means that we can uniquely define $\alpha_{0ct}(\boldsymbol{\theta}_t, \beta)$, $\alpha_{1ct}(\boldsymbol{\theta}_t, \beta)$ that are continuous in all θ_{ct}.

Second, for any $\alpha_{0ct}, \alpha_{1ct}$ we assume that there exists a unique Nash equilibrium in $\boldsymbol{\theta}_t$ of (4.4). As for each pharmacy chain c, the profit function π_{ct} is continuous in all θ_{ct}, the best response of each chain is well defined and we only require that conditions of single crossing of best response functions be satisfied. We will assume this is the case and can be verified empirically. Thus we can define $\theta_{ct}(\boldsymbol{\alpha}_{0t}, \boldsymbol{\alpha}_{1t}, \beta) \in [0,1]$ that solves the maximization (4.4) and are continuous in all $\alpha_{0ct}, \alpha_{1ct}$ because $\pi_{ct}(\boldsymbol{\theta}_t, \alpha_{0ct}, \alpha_{1ct}, \beta)$ is continuous in all θ_{ct} that belong to [0,1].

Then, under technical conditions that the image of $[0,1]^C$ by $\boldsymbol{\theta}_t(\boldsymbol{\alpha}_{0t}(., \beta), \boldsymbol{\alpha}_{1t}(., \beta), \beta)$ is $[0,1]^C$, we can use Brouwer's fixed point theorem, and obtain that there is a vector $\boldsymbol{\theta}_t$ that is solution of

$$\boldsymbol{\theta}_t(\boldsymbol{\alpha}_{0t}(\boldsymbol{\theta}_t, \beta), \boldsymbol{\alpha}_{1t}(\boldsymbol{\theta}_t, \beta), \beta) = \boldsymbol{\theta}_t \tag{4.7}$$

This proves that there is a vector $(\boldsymbol{\alpha}_{0t}(\hat{\boldsymbol{s}}_t, \boldsymbol{m}_{0t}, \boldsymbol{m}_{1t}, \beta), \boldsymbol{\alpha}_{1t}(\hat{\boldsymbol{s}}_t, \boldsymbol{m}_{0t}, \boldsymbol{m}_{1t}, \beta), \boldsymbol{\theta}_t(\hat{\boldsymbol{s}}_t, \boldsymbol{m}_{0t}, \boldsymbol{m}_{1t}, \beta))$ solution of (4.4) and (4.5). At this step, we can search for the possibility of multiple solutions over the support of θ.

We will then assume that the following likelihood function

$$L_i(\beta; \hat{\boldsymbol{s}}_t, \boldsymbol{m}_{0t}, \boldsymbol{m}_{1t}) = L_i(\beta; \boldsymbol{\alpha}_{0t}(\hat{\boldsymbol{s}}_t, \boldsymbol{m}_{0t}, \boldsymbol{m}_{1t}, \beta), \boldsymbol{\alpha}_{1t}(\hat{\boldsymbol{s}}_t, \boldsymbol{m}_{0t}, \boldsymbol{m}_{1t}, \beta), \boldsymbol{\theta}_t(\hat{\boldsymbol{s}}_t, \boldsymbol{m}_{0t}, \boldsymbol{m}_{1t}, \beta)) \tag{4.8}$$

has a unique maximum in β. The estimation routine then becomes a nested fixed point algorithm, where we solve for the parameters $\alpha_{0t}(\beta)$, $\alpha_{1t}(\beta)$ and $\theta_t(\beta)$ conditional on the current value of β in the inner loop, while searching for the parameter vector β that maximizes the log likelihood in the outer loop.

4.2.2 Identifying Bargaining

We now use the vertical structure competition game developed in section 3.3 to identify the supply side parameters of the model. The objective is to identify all the bargaining parameters b_{0c} and b_{1c} respectively for the parallel importer and the producer negotiation with each pharmacy chain c. The marginal cost of production c_t are not identified because retail prices are regulated ($\bar{p}_t$) so that the total margin, $\bar{p}_t - c_t$, is given and does not affect wholesale prices, except by imposing implicit bounds conditions for non negative profits of producers that we assume satisfied for all drugs present on the market.

The optimality conditions of the bargaining game (3.9) and (3.11) relate demand and bargaining parameters to the marginal opportunity costs of drugs for the parallel importer ($\mathbf{p}_{0t}^I$) and the producer (p_{1t}^I).

Remark that all p_{0ct}^I $(c = 1, .., C)$ and p_{1t}^I can be different because of the costs related to packaging and extra logistics when importing from source countries. We assume that parallel importers' costs ($\mathbf{p}_{0t}^I = (p_{01t}^I, .., p_{0Ct}^I)$) and the opportunity costs of the producer (p_{1t}^I) are a function of observables X_t such as the wholesale prices in the source countries and company fixed effects for the producer or parallel importer, as well as interactions with source country prices[18]. With $\mathbf{p}_{0t}^I$ and p_{1t}^I from the optimal bargaining equations (3.8) and (3.10), stacked in the vector $\mathbf{p}_t^I = (\mathbf{p}_{0t}^I, p_{1t}^I)$ for each market t, we specify

$$p_t^I(\mathbf{b}) = X_t\eta + \epsilon_t, \tag{4.9}$$

where $\mathbf{b}$ is the vector of bargaining parameters $\mathbf{b} = (b_{00}, \cdots, b_{0C}, b_{10}, \cdots, b_{1C})$.

Then, we assume that we observe instrumental variables Z_t such that $E[\epsilon_t|Z_t] = 0$ and then identify the parameter vector $(\eta, \mathbf{b})$ using the moment condition $E[\epsilon(\eta, \mathbf{b})|Z] = 0$. Our instrumental variables Z_t include variables X_t, as well as the price ceiling $\bar{p}_t$, indicators for pharmacy chain identity, and interactions. The specific moment conditions we use are the sample means $E[Z'\epsilon(\eta, \mathbf{b})] = 0$, such that our GMM estimator is

$$(\hat{\eta}, \hat{\mathbf{b}}) = \arg\min_{\eta, \mathbf{b}} \epsilon(\eta, \mathbf{b})' Z W Z' \epsilon(\eta, \mathbf{b}),$$

where W is a weighting matrix for the moments.

The intuition for identification of the bargaining parameters, in light of the instrument set, is that pharmacy chain identity should be informative about the overall bargaining strength of the chain, while being plausibly uncorrelated with unobserved determinants of costs related to parallel trade. We thus preclude the possibility that sorting of parallel importers across pharmacy chains is related to the costs of parallel trade.[19] In addition, the price ceiling affects sales revenues of a given product with differential impact on the total value of agreement in the different pharmacy chains. The price ceiling can impact the relative net value of agreement between the upstream firm and pharmacy chain due to differences in the response of demand and other chains' strategies (θ_{ct}) in the event of a disagreement. Thus, we believe that the interactions between pharmacy chain indicators and the price ceiling will help identifying the bargaining parameters, due to the equilibrium effect of changes in net values of agreement being dependent on the bargaining parameters.

[18] We use the wholesale prices of the source countries France, UK but also in Germany, Italy, Spain, Turkey, France, UK and US, that will be informative about the price at which parallel traders acquire the drugs and that the direct importer earns on parallel trade.

[19] The costs here interpretable as both the total costs of parallel traders, e.g., procurement and handling, and opportunity costs of the direct importer, e.g., sales value in source country and differences in import costs between Norway and the source country.

The necessary assumption for the price ceiling—and thus the interactions with pharmacy chain indicators—to be valid instruments, is that the price ceiling is uncorrelated with ϵ_t, conditional on the wholesale prices in other countries included in X_t. It is possible that the price ceiling—being a function of prices in several other countries, as described in Section 2.2—is correlated with the unobserved determinants of parallel trade costs. However, the UK is the only source country in our sample that is also in the reference countries for regulatory price ceilings, we believe this to be less of a concern and perform robustness check with respect to this. Most prices in countries in X_t should help capturing general movements in trade costs, exchange rates and relative prices between different locations.

4.3 Empirical Results

As our data contain a very large amount of choices, we draw a random sample of 50,000 individuals from the full sample of about 170,000 for estimating the individual choice model. The maximum likelihood estimates of the individual heterogeneity parameters are shown in Column *Full model* of Tables 4.3 and 4.4. Table 4.3 shows the baseline parameters, common across all unobserved groups (and specific to reference group 0), while Table 4.4 shows the parameters for the unobserved groups with differing values from the baseline.[20] With these estimates, over 80% of the 95 markets that feature parallel imports have rationing to some extent, where 127 of the 345 chain-market combinations with parallel imports are estimated to feature rationing. Given that the chain is estimated to perform rationing, it is often substantial, with an average θ of about 0.3, i.e., a 70% probability of being stocked out of the less profitable product. We find that the largest chain, *Apotek 1*, sets $\theta < 1$ about twice as often as the other two chains and has an average θ of 0.4, while it is about 0.8 for the other chains.

We also estimate the model excluding the possibility that pharmacies affect consumers by changing availability for comparison (imposing $\theta = 1$). The results of this specification are shown in Column *Reduced* of Table 4.3 and 4.4. The difference in likelihood of over 15,000 log points tells us that our proposed model has a substantially better fit than the alternative where θ is restricted to be equal to one. It should be noted that θ are not free parameters, but are set according to restrictions from pharmacies optimizing behavior and the additional data afforded by observing wholesale prices. The results here imply that the extra information and the way the model puts it to use has a significant contribution towards explaining the choices we see in our data.

From the estimates of parameters governing preferences according to unobserved, discrete groups in the population in Table 4.4, there are two striking features: The first is that the statistical and economic significance of these parameters imply that the specification is appropriate, compared to a more usual mixed

[20] Note that the group given by $g = 0$ is defined by having $\delta_c^g = 0, : \forall c$ and $\sigma_C^g = 0$, such that the preferences of this group is represented by the baseline utility parameters.

parameters logit specification with a single distribution on each coefficient.[21] The second is a pattern where each group has a stronger relative preference for each of the pharmacy chains. This seems reasonable, as one would suspect that many unobserved factors, such as travel distance or chain store preference, would contribute to exactly such a pattern.

Table 4.3: *Choice model estimates: utility parameters common across all consumers*

	Full model	Reduced
σ_J	0.50	0.53
	(0.01)	(0.01)
γ_1	0.00	-0.23
	(0.02)	(0.02)
σ_1	13.07	6.94
	(0.19)	(0.29)
σ_2	14.45	15.56
	(0.18)	(0.41)
σ_3	5.38	6.93
	(0.13)	(0.29)
$\ln \mathcal{L}(\hat{\beta})$	-243,244	-258,477
N	50,000	

Standard errors in parenthesis.

The results from the estimation of the bargaining parameters, following our discussion in Section 4.2.2, are shown in Table 4.5. Keep in mind that the bargaining parameters are the bargaining weights of the upstream firms. From these estimates, we can see that, perhaps surprisingly, the parallel importers wield a larger bargaining weight on average, compared to the originator. The difference in bargaining weights between the direct importer and parallel importers could partly reflect the fact that the original producer likely also makes profits on sales of parallel imports, albeit in other countries. In addition, if the originator would come to a disagreement with a given chain, some of the lost sales will be captured by increased sales of its direct imported variety in other chains, while the rest will be captured by parallel imports, both of which will generate profits for the direct importer. For the parallel importer, all sales going to the direct imported product is irrevocably lost profits. Remark that $0 < b < 1$ is not imposed by our estimation.

4.4 Prescription behavior

One worry for the identification of our model is that doctors will change what they prescribe if pharmacies induce consumers to consume parallel traded Lipitor more frequently. An example of what we have in mind is that consumers might oppose getting parallel traded drugs, thereby making their doctor prescribe them

[21] Note that the specification here is a finite mixture of normal distributions. The economic significance is based on comparisons of behavioral implications under a simpler distributional specification not reported here.

Table 4.4: *Choice model estimates: Shares and utility parameters of unobserved groups*

	Full model			Reduced		
	$g = 1$	$g = 2$	$g = 3$	$g = 1$	$g = 2$	$g = 3$
$\tau_g(\boldsymbol{\eta})$	0.33	0.28	0.18	0.51	0.12	0.09
	(0.005)	(0.007)	(0.004)	(0.010)	(0.006)	(0.008)
η_g	0.5	0.34	-0.13	1.61	0.2	-0.12
	(0.05)	(0.06)	(0.04)	(0.06)	(0.11)	(0.07)
δ_1^g	9.07	-0.08	-0.42	-1.08	5.55	2.27
	(0.16)	(0.19)	(0.16)	(0.22)	(0.29)	(0.23)
δ_2^g	0.16	6.51	-0.1	-1.64	10.71	0.26
	(0.19)	(0.00)	(0.13)	(0.28)	(0.40)	(0.42)
δ_3^g	2.94	-0.11	6.17	-4.01	1.14	3.55
	(0.38)	(0.49)	(0.31)	(0.24)	(0.36)	(0.23)
σ_C^g	5.16	1.84	64.82	2.35	3.06	0.76
	(0.32)	(0.26)	(0.79)	(0.35)	(0.59)	(0.42)

Standard errors in parenthesis.

Standard errors of τ_g calculated by the delta method.

Group 0 constitutes 0.21 for *Full model* and 0.28 for *Reduced*.

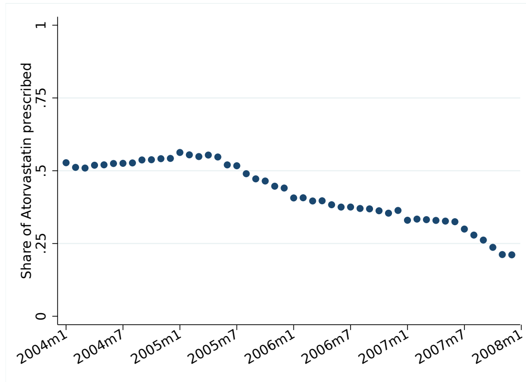
Table 4.5: *GMM estimation of upstream firms bargaining parameters*

	Direct import	Parallel import
Chain 1	0.47	0.83
	(0.02)	(0.02)
Chain 2	0.06	0.64
	(0.01)	(0.09)
Chain 3	0.77	0.57
	(0.12)	(0.11)

other types of statins for which there does not exist parallel traded alternatives. Over the sample period, there has been an increase in the share of statin prescriptions going to Simvastatin due to new guidelines for Statin prescriptions from the Norwegian Medicines Agency. This increase has gone together with a similar decrease in the share of statin prescriptions going to Atorvastatin (Lipitor), as shown in Figure 4.3. We regard this decrease as a function of the change in policy for statin prescription, induced by the government who implemented a lower reimbursement rates on Simvastatin than Atorvastatin, and not necessarily related to the preferences of consumers or doctors for directly imported versus parallel trade.

We want to investigate potential endogeneity issues arising from doctors responding to pharmacies strategies for selling parallel traded Lipitor by changing what statin they prescribe. To do this, we use data on the prescription behavior of individual doctors, where we can look at the share of statin prescriptions going to Atorvastatin together with the behavior of the pharmacies that the doctor's patients are exposed to. This is feasible due to availability of information linking the doctor to the prescription used by a patient for each

Figure 4.3: *Doctors' prescription of Atorvastatin as share of total statin prescription*



transaction at each given pharmacy. Since we do not directly observe the behavior of pharmacies, we use information on availability of parallel imports, and the ratio of margin for parallel and direct imports at a given pharmacy. The availability gives a sense of whether the doctors patients potentially was faced with a choice of parallel trade, while the margin can be thought of as an instrument for the pharmacies decision to induce choice of parallel trade. To operationalize this, we calculate the weighted sum of availability and margin ratio for each doctor, where the measure for each pharmacy is weighted by the share of the doctor's patients going to this pharmacy. More precisely, for doctor j in month t

$$available_{jt} = \frac{1}{N_{jt}} \sum_{i=1}^{N_{jt}} \mathbf{1}_{\{parallel_{it}\}},$$

where N_{jt} is the number of patients for doctor j in month t, and $\mathbf{1}_{\{parallel_{it}\}}$ is an indicator for whether patient i went to a pharmacy offering parallel traded Lipitor in month t. Similarly

$$ratio_{jt} = \frac{1}{N_{jt}} \sum_{i=1}^{N_{jt}} \frac{m_{0it}}{m_{1it}},$$

where $\frac{m_{0it}}{m_{1it}}$ is the ratio of margins for parallel (0) and direct (1) imported Lipitor at the pharmacy visited by patient i in month t. If parallel trade is not available at the pharmacy visited by consumer i, this ratio is set to 0. The summary statistics over all doctors and months of our sample is shown in Table 4.6. Overall, we see that doctors prescribe Lipitor in 43% of the cases where a statin was prescribed, while parallel trade is available for 25% of the patients. The margin ratio is 0.32 on average, which includes the zero margin ratio for the 75% of patients which did not have parallel trade available at the pharmacy they visited. The number of unique doctors in our sample is 14,051, who are observed for a maximum of 48 months between January 1 2004 and December 31 2008.

In Table 4.7, we show the results of regressions of Atorvastatin share of statin prescriptions on weighted margin ratios and parallel trade availability. The observation unit is doctor-month. Column (1) shows the

Table 4.6: *Summary statistics for doctors' prescription of Atorvastatin*

	Mean	Std. Dev.	Min	Max
Share of Lipitor prescribed	0.43	0.28	0.01	1.00
Margin ratio	0.32	0.54	0.00	5.02
Parallel offered	0.25	0.34	0.00	1.00
N	258,281			

Table 4.7: *Share of Atorvastatin prescribed by doctor as function of the margin ratio between direct and parallel imports, and availability of parallel imports faced by the doctor's patients*

	(1)	(2)	(3)	(4)	(5)	(6)
Margin ratio	-0.052***	-0.000	0.003	-0.036***	0.005**	0.007***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Parallel offered	-0.018**	-0.010	-0.013*	-0.047***	-0.022***	-0.023***
	(0.006)	(0.006)	(0.006)	(0.004)	(0.003)	(0.003)
Time trend		x			x	
Time FE			x			x
Dr. FE				x	x	x
N	258,281	258,281	258,281	258,281	258,281	258,281
R^2	0.01	0.11	0.13	0.02	0.18	0.20

Standard errors clustered by doctor

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

pure OLS, giving a large negative coefficient on margin and availability, though this is driven by the overall downward trend in Atorvastatin prescriptions, together with a tendency for both the margin ratio and the availability of parallel trade to increase over time. This is confirmed by the coefficient on margin ratio going to a quite precisely estimated zero in Columns (2) and (3), where we add a linear time trend and time-fixed effects respectively. When we add doctor-fixed effects together with a time trend or time-fixed effects in Columns (5) and (6), we obtain a positive coefficient on the margin ratio and a negative coefficient on availability, both statistically significant. However, considering the size of the coefficients, none of them are economically significant. The coefficient on the margin ratio tells us that the effect of an increase of roughly two standard deviations (from Table 4.6), the Atorvastatin share of statin prescriptions will increase by roughly one half percentage point. Similarly for availability, an increase in availability from none to full would yield a decrease in Atorvastatin prescriptions by 2.2 percentage points. Considering that the average availability is 25%, this implies that very large changes in pharmacies behavior is related to relatively small changes in the prescription behavior of doctors in our sample. We thus conclude that we should not be concerned by a potential identification problem due to doctors changing molecule prescriptions in response to pharmacies incentives to sell parallel traded Lipitor more frequently.

5 Counterfactual Simulations

With our estimated model, it is possible to answer interesting questions based on several counterfactuals. The first regards of parallel trade on market equilibrium, firms profits and consumer welfare. It is based on comparing the current situation with the counterfactual equilibrium obtained absent parallel trade. The second regards the impact of pharmacy chains' strategic behavior on the vertical chain profits and final consumers. In order to evaluate such effect, we consider the counterfactual framework where a possible regulation of pharmacies would restrict their possibility to use availability of direct versus parallel imports and thus to strategically affect consumers' choices.

5.1 The Impact of Parallel Trade

Using our structural model, we can simulate the counterfactual situation where producers and pharmacy chains would not be able to use parallel imports. In such a case, it is clear that pharmacy chains would then propose only the directly imported version of drugs and as retail prices would still be regulated and equal to the price ceilings, it is easy to identify the effect on demand of banning parallel trade and thus identify consumer welfare. Concerning pharmacy chains bargaining with producers in the absence of parallel imports, we simply need solve the counterfactual equilibrium in absence of parallel trade.

In such as case, a consumer chooses chain c with counterfactual probability s^*_{ict} which is equal to choice probability of the directly imported drug in chain c, s^*_{i1ct}, that is

$$s^*_{ict} = s^*_{i1ct} = \frac{e^{\alpha_{i1ct}}}{\sum_{\tilde{c}} e^{\alpha_{i1\tilde{c}t}}}$$

and the aggregate counterfactual market share of chain c is

$$s^*_{1ct} = \int s^*_{i1ct} dF(\nu_i).$$

Then, the producer profit without parallel imports is given by

$$\Pi_t(\mathbf{w}_{1t}) = \sum\nolimits_c \left[(w_{1ct} - c_t)s^*_{1ct}\right],$$

while the pharmacy chains profits are

$$\pi_{ct} = (\bar{p}_t - w_{1ct})s^*_{1ct}$$

As in each pairwise negotiation with the pharmacy chains, the producer and pharmacy chain c sets wholesale prices w_{1ct} to maximize the Nash-product

$$(\Pi_t - \Pi_{-c,t})^{b_{1c}}(\pi_{ct} - \pi_{-1,ct})^{1-b_{1c}}, \tag{5.1}$$

where now

$$\begin{aligned}\Pi_t - \Pi_{-c,t} &= \sum\nolimits_{\tilde{c}} [(w_{1\tilde{c}t} - c_t)s^*_{1\tilde{c}t}] - \sum\nolimits_{\tilde{c}\backslash c}(w_{1\tilde{c}t} - c_t)s^*_{1\tilde{c}t\backslash 1c} \\ &= w_{1ct}\Delta_{1c}s^*_{1ct} + \sum\nolimits_{\tilde{c}\backslash c} w_{1\tilde{c}t}\Delta_{1c}s^*_{1\tilde{c}t}\end{aligned}$$

where $\Delta_{1c}s^*_{1\tilde{c}t} \equiv s^*_{1\tilde{c}t} - s^*_{1\tilde{c}t\backslash 1c} < 0$ for all $\tilde{c} \neq c$ because pharmacy chains are substitute and $\Delta_{1c}s^*_{1ct} = s^*_{1ct} > 0$ because $s^*_{1ct\backslash 1c} = 0$. As $\sum_{\tilde{c}} \Delta_{1c}s^*_{1\tilde{c}t} = 0$, we have $\Delta_{1c}s^*_{1ct} = -\sum_{\tilde{c}\backslash c} \Delta_{1c}s^*_{1\tilde{c}t} > 0$, thus if w_{1ct} is strictly smaller than all other wholesale prices $w_{1\tilde{c}t}$, it will imply that $\Pi_t - \Pi_{-c,t} < 0$. As these conditions must hold for all c, it must be that all wholesale prices are equal and then $\Pi_t - \Pi_{-c,t} = 0$ and the wholesale price is indeterminate. Pharmacy chains being unable to access parallel imports, they obtain zero profit in case of disagreement with the producer ($\pi_{-1,ct} = 0$) and thus

$$\pi_{ct} - \pi_{-1,ct} = (\bar{p}_t - w_{1ct})s^*_{1ct}$$

We will assume that wholesale prices fixed by manufacturers are then equal to the regulated retail price such that $w_{1ct} = \bar{p}_t$ and then we also have $\pi_{ct} - \pi_{-1,ct} = 0$.

Concerning consumer surplus, it is equal to the traditional logit log-sum under the current equilibrium

$$E\left[U_{it}\right] = \int E\left[\max_{j,c}\{u_{ijct}\}\middle| \nu_i\right] \mathrm{d}F(\nu_i) = \int \ln\left(\sum\nolimits_c e^{\alpha_{i0ct}+\theta_{ct}\delta_{ict}}\right) \mathrm{d}F(\nu_i).$$

and in the counterfactual situation without parallel trade, it is

$$E\left[U^*_{it}\right] = \int \ln\left(\sum\nolimits_c e^{\alpha_{i1ct}}\right) dF(\nu_i)$$

so that the change in consumer welfare is

$$E[\Delta EU_{it}] = E\left[U^*_{it} - U_{it}\right] = \int \ln\left(\frac{\sum_c e^{\alpha_{i1ct}}}{\sum_c e^{\alpha_{i0ct}+\theta_{ct}\delta_{ict}}}\right) dF(\nu_i)$$

We can note that there are several effects. First, there is a *positive* effect of more diversity, which is partially due to the inclusion of an idiosyncratic random utility component for each additional product, and partially due to preference heterogeneity, where some consumers will gain due to higher valuation of the new products. Remark that retail price regulation prevents price competition to improve consumers welfare since all prices are at price ceiling. Second, there is an *negative* effect stemming from the uncertainty in choice sets because $\alpha_{i0ct} + \theta_{ct}\delta_{ict} = \alpha_{i0ct} + \theta_{ct}\ln(1 + \alpha_{i1ct} - \alpha_{i0ct}) \leq (1 - \theta_{ct})\alpha_{i0ct} + \theta_{ct}\alpha_{i1ct} \leq \max(\alpha_{i0ct}, \alpha_{i1ct})$. Then if only direct imports are available, there is a *negative* effect for consumers preferring parallel trade versions but a *positive* effect for those preferring direct imports. Those heterogenous effects have a magnitude that also depends on θ_{ct}.

Remark that in order to transform consumers surplus changes into a monetary compensating variations, we need divide by the marginal utility of income, which we don't identify in our model because consumers choose between same price alternatives. Thus, we consider the proportion of consumers who would be better off without parallel imports,

$$E[\Delta EU_{it} > 0] = \int \mathbf{1}_{\{EU^*_{it} - EU_{it} > 0\}} \, dF(\nu_i),$$

to assess the importance and role of preference heterogeneity.

Table 5.1: *Change in consumer welfare from removal of parallel imports*

	All markets	Any $\theta_{ct} < 1$
$\%\Delta EU_t$	-1.7%	-0.9%
$E[\Delta EU_{it} > 0]$	34.8%	41.9%
N	95	79

The estimated changes in welfare are presented in Table 5.1. It is clear that the loss of diversity would be the dominant effect, reducing consumer welfare on average by almost two percent. We interpret this as a modest loss to consumer welfare if parallel imports were to be removed from this market. For 16 out of these 95 markets, there is no estimated supply restriction, and considering only the markets featuring supply restriction (i.e., some $\theta_{ct} < 1$), the reduction in consumer welfare is only about 1%. We also see that a large proportion of consumers have strong enough preferences for the direct imports that they would benefit from such a ban. In 34 of the 95 markets the change in consumer surplus would actually be positive on average, which corresponds to markets where the mean valuation for direct imports is relatively high compared to parallel imports and where the pharmacy chains use supply restrictions to a large degree.

Table 5.2 shows the counterfactual results if we were to ban parallel trade. It shows that the change in profits would favor the upstream producers and penalize pharmacy chains would not be able to use intra-brand competition between parallel trade and direct imports to extract part of manufacturers profits.

Table 5.2: *Change in profits from removal of parallel imports*

Pharmacy Chain	Δq	$\Delta(p \cdot q)$	Δw	$\Delta(w \cdot q)$	$\Delta\pi$	$\Delta\Pi$
Chain 1						
Direct	55.2	162.0	1.02	186.3	-24.2	78.7
	244.1%	191.6%	40.2%	308.8%	-100.0%	46.9%
Parallel	-51.8	-150.6	-	-102.1	-48.5	-5.8
	-100.0%	-100.0%	-	-100.0%	-100.0%	-100.0%
Chain 2						
Direct	20.0	54.2	1.06	78.4	-24.2	37.3
	85.7%	66.3%	42.3%	136.3%	-100.0%	37.9%
Parallel	-20.3	-57.3	-	-39.8	-17.4	-1.5
	-100.0%	-100.0%	-	-100.0%	-100.0%	-100.0%
Chain 3						
Direct	29.5	81.8	1.03	98.7	-16.8	45.5
	206.1%	139.3%	40.4%	235.4%	-100.0%	47.9%
Parallel	-28.6	-77.9	-	-50.4	-27.4	-4.7
	-100.0%	-100.0%	-	-100.0%	-100.0%	-100.0%
Total Direct	104.7	298.1	1.04	363.3	-65.3	161.5
	173.7%	132.5%	40.9%	227.4%	-100.0%	44.7%
Total Parallel	-100.7	-285.7	-	-192.3	-93.4	-11.9
	-100.0%	-100.0%	-	-100.0%	-100.0%	-100.0%

Quantities (q) in million DDD, wholesale prices (w) in NOK/DDD and monetary sums in million NOK.

5.2 Pharmacy chain strategy and distribution of profits

With our estimates of the bargaining model, parallel importer costs and direct importer opportunity cost, it is also possible to assess the impact of the pharmacy chains' strategy of optimally choosing the probability with which a drug will be proposed to the consumer. We will consider the case where each pharmacy chain c sets $\theta_c = 1$. This could be due to the regulator enforcing an obligation to supply all varieties, or if the producer could make a take-it-or-leave-it requirement to always propose its product in the contract with each pharmacy chain. *This will have an effect on the composition of goods sold, where it is clear that the amount of direct imports sold will increase at the expense of parallel imports, though it is difficult to say by how much.* Among the chain-market combinations featuring parallel imports, the average θ is roughly three quarters, meaning that the consumer will face a restricted choice set one in four times. The quantitative effect of setting θ below one depends on the preferences of the consumers in the cases where $\theta_c < 1$, since some will likely buy the parallel imported variety also in the case where both are available. When the pharmacy chains are required to always keep both varieties, it will also have an effect on the bargained wholesale prices between the upstream firms—the direct and parallel importers—and the pharmacy chains. This implies that the wholesale prices in general will increase, since there is no longer an incentive for the upstream firms to reduce wholesale prices to increase sales. The distribution and size of this increase cannot in general be determined theoretically.

Table 5.3: *Impact of a requirement to carry all varieties* ($\theta_c = 1$, $\forall c \in \mathcal{C}$)

	Δq	$\Delta p \cdot q$	Δw	$\Delta w \cdot q$	$\Delta \pi$	$\Delta \Pi$
Chain 1						
Direct	22.8	72.6	0.19	62.7	9.9	15.6
	26.7%	14.0%	5.0%	17.0%	46.7%	9.3%
Parallel	-21.4	-67.4	0.35	-39.2	-28.2	2.3
	-41.0%	-44.8%	15.8%	-38.4%	-55.5%	40.3%
Chain 2						
Direct	2.2	4.2	0.16	7.9	-3.7	1.3
	4.3%	1.3%	4.2%	3.4%	-9.7%	1.3%
Parallel	-3.9	-10.2	0.05	-4.8	-5.5	1.2
	-15.9%	-17.8%	2.6%	-12.1%	-23.1%	87.4%
Chain 3						
Direct	8.1	24.5	0.16	21.6	2.9	6.5
	17.1%	8.3%	4.3%	10.3%	25.0%	6.8%
Parallel	-7.8	-23.8	0.31	-9.8	-14.0	2.4
	-26.5%	-30.6%	14.4%	-19.4%	-46.1%	72.9%
Total Direct	33.1	101.6	0.17	92.2	9.1	23.4
	14.4%	7.5%	4.5%	11.4%	20.2%	6.5%
Total Parallel	-33.1	-101.6	0.25	-53.8	-47.7	5.9
	-27.9%	-29.8%	11.8%	-28.0%	-46.7%	56.9%

Quantities (q) in million DDD, wholesale prices (w) in NOK/DDD and monetary sums in million NOK.

To get a sense of the quantitative impact of the chains' strategies, we calculate the market equilibrium that would arise if the chains always had available the varieties which they are observed to sell in the data.[22] For these calculations, we take consumer preferences, marginal costs of the parallel importer, the direct importer's gain on each parallel traded unit and bargaining weights as given, and solve for demand and the bargaining outcomes. In Table 5.3 we show the changes from the current situation in terms of quantities (Δq), sales revenues ($\Delta p \cdot q$), wholesale prices (Δw), wholesale expenditure ($\Delta w \cdot q$), profits of the pharmacy chains ($\Delta \pi$) and profits of the upstream firms ($\Delta \Pi$), broken down by pharmacy chain and type of upstream firm both in units and percentages. We see that such a regulatory change would have large impacts on both sales and distribution of profits in the market, where sales of parallel imports would drop by about 33 million DDD and 100 million NOK (roughly 12.5 million EUR). This change is very unevenly distributed between the chains, where the largest chain, *Apotek 1*, would experience the largest changes both in absolute numbers and relative to the status quo, which to a large extent mirrors the aggressive policies of this chain in stocking

[22] The implication of the last point is that a chain is not set up to carry parallel imports in a given market in our counterfactual simulation if we didn't observe any sales in our data. This might seem like a strange artefact, though this is due to our goal of quantifying the importance of the optimal stock-out probabilities, and thus avoiding contamination with elements of assortment choice and parallel export opportunities.

out of the direct imports. Note that sales, both in quantity and value, after the regulation will be absent any strategic behavior by the firms, less the decision to be present in the market which we take as given. Thus, sales will be given only by consumer preferences. Overall, the wholesale prices of the parallel importers increases the most, which reflects that the parallel importers especially were in a position to increase their sales by reducing prices to the pharmacy chains, thus giving the pharmacy chains incentives to distort supply. We see that both the direct importer and parallel importers would gain from such a change, especially the parallel importers in relative terms. This is due to the fact that there is no longer an element of competition between the upstream firms when bargaining over wholesale prices with the chains, in the sense that there is no longer any opportunity to affect pharmacy chains' choice of θ. Since the parallel importers earlier had very small margins, the increased wholesale prices would have a large enough impact on profits to more than offset the reduction in volume.

6 Conclusion

In this paper, we have investigated the incentives of pharmacy chains in selling parallel traded drugs. Our estimates show that the availability of direct imported drugs are plausibly reduced to a large extent in the specific market we study. This is driven by the two factors of constrained pricing and parallel importers generally giving the pharmacy chains lower wholesale prices than the direct importer. Our counterfactual estimates indicate that the gains to consumers from parallel imports in our market of study is small, mostly due to reduced availability of direct imports resulting in lower utility for a large portion of consumers with preference for the direct imported variety. In countries where pharmaceutical prices are less regulated, this could very well be overturned. Further, we find that the possibility of reducing availability benefits the pharmacy chains at the expense of both the direct importer and parallel importers. In this market, where prices are constrained by regulation, being able to distort supply between the varieties introduces competition between the upstream firms through the pharmacy chains.

The specific mechanism we highlight—where pharmacies can distort availability of drugs for which they have differing margins—have not been formalized in the previous literature, though pharmacists' incentives have been mentioned as a plausible factor impacting sales of drugs for which substitution at the pharmacy level is available (see e.g., Caves et al., 1991). The incentives to distort availability seems particularly important in many European countries, where price regulation is prevalent. Furthermore, to the best of our knowledge, the addition of consumer expectations over available choices to the choice model we utilize have not been studied in the previous literature using discrete choice models.

For future studies, it would be interesting to know how prevalent the kind of behavior that we have studied here are in other pharmaceutical markets, both where prices are tightly and less tightly regulated.

References

Brekke, Kurt R., T. H. H. and O. R. Straume (2015). Price regulation and parallel imports of pharmaceuticals. *Journal of Public Economics 129*, 92–105.

Caves, R., M. Whinston, M. Hurwitz, A. Pakes, and P. Temin (1991). Patent expiration, entry, and competition in the us pharmaceutical industry. *Brookings papers on economic activity. Microeconomics 1991*, 1–66.

Crawford, G. S. and A. Yurukoglu (2012). The welfare effects of bundling in multichannel television markets. *American Economic Review 102*(2), 643–85.

Danzon, P., Y. R. Wang, and L. Wang (2005). The impact of price regulation on the launch delay of new drugs evidence from twenty-five major markets in the 1990s. *Health Economics 14*(3), 269–292.

Duso, T., A. Herr, and M. Suppliet (2014). The welfare impact of parallel imports: A structural approach applied to the german market for oral anti-diabetics. *Health Economics 23*, 1036–1057.

Ganslandt, M. and K. E. Maskus (2004). Parallel imports and the pricing of pharmaceutical products: evidence from the european union. *Journal of Health Economics 23*(5), 1035 – 1057.

Gowrisankaran, G., A. Nevo, and R. Town (2015). Mergers when prices are negotiated: Evidence from the hospital industry. *American Economic Review 105*(1), 172–203.

Ho, K. and R. Lee (2015). Insurer competition in health care markets. *Working Paper*.

Horn, H. and A. Wolinsky (1988). Bilateral monopolies and incentives for merger. *The RAND Journal of Economics 19*(3), pp. 408–419.

Kanavos, P., J. Costa-Font, and E. Seeley (2008). Competition in off-patent drug markets: Issues, regulation and evidence. *Economic Policy 23*(55), pp. 499–544.

Kanavos, P., J. Costa-i Font, S. Merkur, and M. Gemmill (2004). *The economic impact of pharmaceutical parallel trade in European Union member states: A stakeholder analysis.* London School of Economics and Political Science London.

Kyle, M. K. (2007). Pharmaceutical price controls and entry strategies. *The Review of Economics and Statistics 89*(1), pp. 88–99.

Kyle, M. K. (2011). Strategic responses to parallel trade. *The BE Journal of Economic Analysis & Policy 11*(2), pp. 1–32.

Sakshaug, S., K. Furu, Ø. Karlstad, M. Rønning, and S. Skurtveit (2007). Switching statins in norway after new reimbursement policy – a nationwide prescription study. *British Journal of Clinical Pharmacology 64*(4), 476–481.

7 Appendix

7.1 Parallel trade products

Figure 7.1: *Example of parallel trade and direct imported products (outside)*

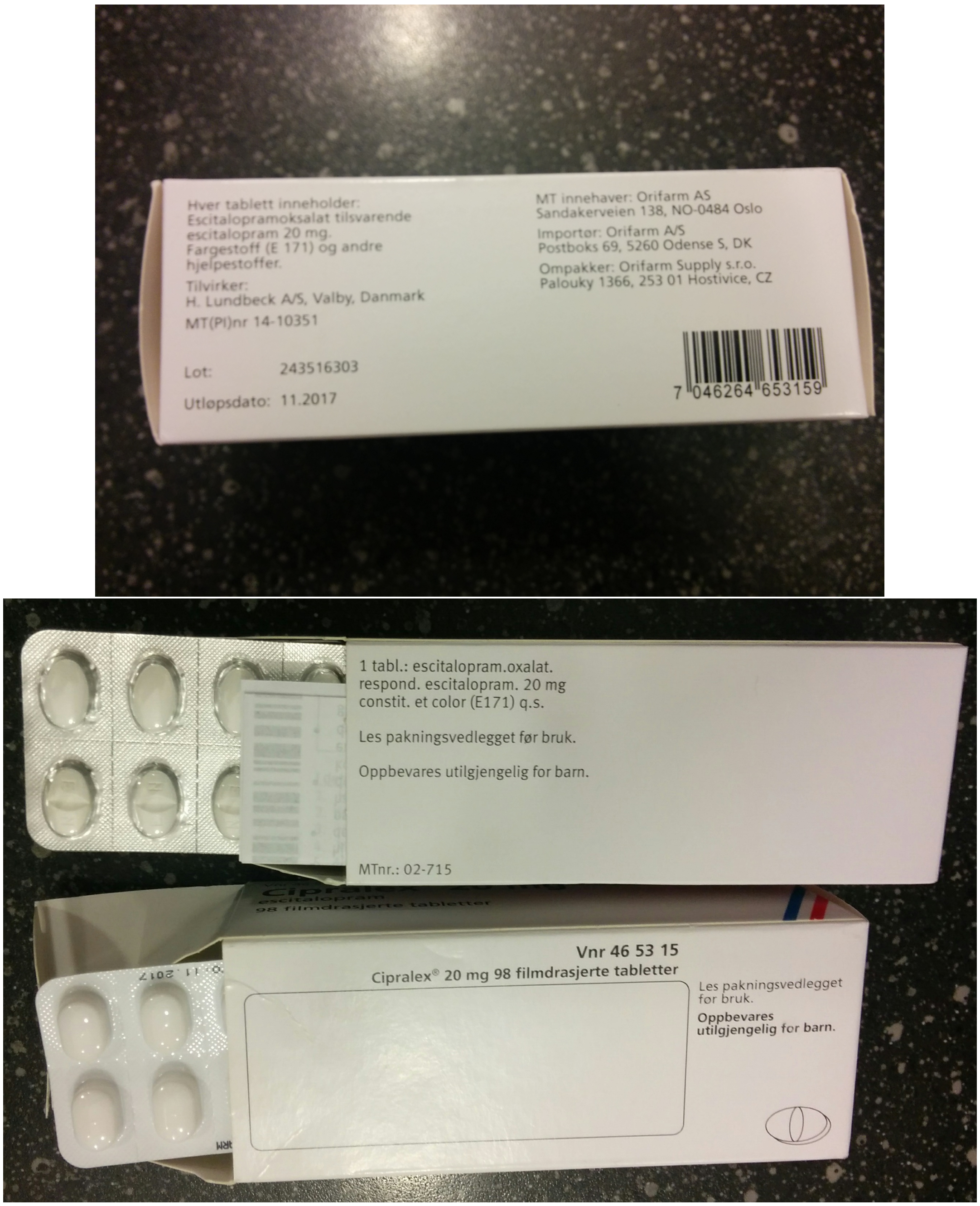


Figure 7.2: *Example of of parallel trade and direct imported products (inside)*