Switching Costs, Brand Premia and Behavioral Pricing in the Pharmaceutical Market*

Aljoscha Janssen†

February 28, 2022

Abstract

This article examines the market power of branded prescription drugs faced with generic competition. Using prescription-level and matched socioeconomic panel data for the entire Swedish population between 2010 and 2016, I provide evidence for the key role of switching costs. A discontinuity surrounding patent expiration dates establishes that the effect is causal. Further, by comparing patients with and without medical education, I show that those without medical education experience higher brand premia. A unique feature of the Swedish market allows me to rule out patients' inattention due to information costs as a source of market power. Therefore, switching costs and perceived quality differences are the key determinants of market power. I then estimate a dynamic oligopoly model with forward-looking firms and use this model in counterfactual studies demonstrating the effects of switching costs and perceived quality differences on prices. First, lengthening the procurement period mimics a reduction of switching costs and increases prices. While switching costs increase competitive pressure and thus reduce prices on average, they also increase inequality as patients with high switching costs pay more. Second, if all patients acted like those with medical education and thus experienced lower brand premia, prices would decrease.

Keywords: Switching Costs, Brand Premia, Behavioral Pricing, Pharmaceuticals

JEL Codes: D12, I11, L13

*I am grateful to Richard Friberg and Michelle Sovinsky for detailed feedback. I thank Andreas Born, Francesco Decarolis, Anna Dreber Almenberg, Liran Einav, Michele Fioretti, Matthew Grennan, Magnus Johannesson, Johannes Kasinger, Matthew Leisten, Erik Lindqvist, Elle Parslow, Matthew Shapiro, and Giancarlo Spagnolo as well as seminar participants at the Stockholm School of Economics, the Universidad Carlos III de Madrid, the XXXIII Jornadas de Economía Industrial, the Research Institute of Industrial Economics, the University of Gothenburg, the Empirics and Methods in Economics Conference at Northwestern University, the MicroWave at the IIES, the University College London, the University of Bern, the IO workshop in St. Gallen, the University of Umea, the Norwegian School of Economics, the European Winter Meeting of the Econometric Society, the 43rd Simposio de la Asociación Española de Economía-Spanish Economic Association, the Copenhagen Business School, the Hanken School of Economics, the National University of Singapore, the Singapore Management University, the University of Sydney, the University of Melbourne, the Ecole Polytechnique, the Instituto Tecnológico Autónomo de México, the IIOC 2019 in Boston, and the APIOS 2019 for helpful comments. Financial support from the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged.

†Singapore Management University and IFN, ajanssen@smu.edu.sg

1 Introduction

Consumers in various markets often stick to products they have purchased before and prefer to buy known brands even when cheaper alternatives of similar quality are available. Previous research has suggested possible reasons for such habit persistence and brand premia: switching costs, inattention due to information costs, risk aversion, quality uncertainty, quality misconceptions or brand prestige effects.[1] Such behavior is especially surprising in the context of prescription drugs in markets where counterfeit pharmaceuticals are not a substantial issue.[2] Consumers often prefer branded and familiar products even though regulations require generic products to be medically equivalent to the branded product, and pharmaceuticals do not generally carry any value apart from their treatment abilities.

In comparison to the United States, where the share of generics exceeded 80% of the total pharmaceutical market volume in 2018, generics have a market penetration of less than 50% in most European countries (OECD, 2017). Even among generics, branded generics play a crucial role. Branded generics are off-patent products marketed under a new trade name by manufacturers that are different from the originator. Branded generics are more expensive than generics without a brand name. To increase uptake of generic pharmaceuticals, Sweden not only offers patients monetary incentives to consume the cheapest alternative among medically equivalent products but also educates patients at the pharmacy level about the equivalence of products. Yet, up to 20% of patients refuse substitution to the cheapest available generic.

In this article I answer the question of whether consumers' behavioral frictions can explain the market power of branded prescription drugs. This study uses individual choice data for prescription drugs covering the entire Swedish population from 2010 to 2016 to examine the underlying cognitive reasons why consumers are willing to pay higher prices for branded pharmaceuticals that have the same ingredients as the generic versions. The unique institutional setting and data sources allow me to identify important determinants of consumer choice: switching costs and perceived quality differences. First, I document behavioral phenomena in the data and then offer a structural estimation of a dynamic oligopoly model to allow for a systematic evaluation of the effects of counterfactual policies. In the counterfactual analysis I show that prices in equilibrium are dependent on consumers' behavioral frictions. Behavioral pricing incorporates the idea that the pricing of a product is based on the behavior of potential customers; here, this concept explains the persistence of high prices as well as the price dynamics of pharmaceuticals under generic

[1] The importance of switching costs in the retail market has been documented by Erdem and Sun (2001), Dubé et al. (2009), Shcherbakov (2016), Shy (2002), Viard (2007) and Rickert (2016). There is a substantial economic and marketing literature on brand premia and the general concept of brand equity. Following Handel and Schwartzstein (2018), one may distinguish mental gaps from behavioral frictions that are in line with the neoclassical assumption of correct beliefs. For the latter, see Hortaçsu and Syverson (2004) for a discussion of risk aversion and possible information costs. Ackerberg (2001) discusses the importance of prestige effects. The other possibility is that consumers have mental gaps or psychological distortions (Handel and Schwartzstein, 2018). Also, the media has covered habit persistence and brand premia; see, for example, Mullainathan (2017) for the *New York Times*.

[2] Counterfeit pharmaceuticals are a problem in large parts of the world. According to an OECD report (OECD, 2009), the incidence of counterfeit medicine is 1% across all medications but is an increasing problem in some Asian and African countries. The problem of counterfeit medicine in the legal pharmaceutical industry in the European Union is less severe. In Sweden, counterfeit medicine has not received attention.

competition.

This paper contributes to the health economics and behavioral economics literatures by demonstrating a behavioral phenomenon among pharmaceutical consumers that could be factored into public policy decisions aiming to decrease the costs of prescription drugs for patients and government payers.

First, the panel structure of the individual choice data allows the identification of consumers' switching costs after starting a medical treatment with a specific product. Crucially, switching costs are not apparent in all prescription markets.[3] I show that patients are 16.04% (3.09 percentage points) more likely to pay premia for painkillers and 11.52% (1.08 percentage points) more likely to pay premia for antibiotics if they have consumed the identical product in the preceding month. There is no evidence of this habit persistence in markets for antiepileptics. Further, I compare repeated purchases by patients shortly before and after the expiration of a substance's patent, showing that the estimates of habit persistence due to switching costs are robust. The evidence of switching costs over a wide range of pharmaceuticals is novel.

Second, the regulatory structure of the Swedish prescription drug market, as well as the registry data, allows the identification of mechanisms behind the brand premia. In general, the explanation for brand premia is that either some brands carry a higher quality (in real terms or caused by psychological effects, i.e., prestige effects) or consumers perceive quality differences, for example, due to uncertainty about the medical equivalence between products. For two reasons, brand premia are not likely to be due to real quality differences. First, the market for prescription drugs in Sweden is ordered into groups of exchangeable products that are medically equivalent.[4] Second, I further reduce the possibility of real quality differences by comparing purchasing behavior between experts (patients with medical education as a physician) and non-experts (patients without such medical education) in an approach similar to that of Bronnenberg et al. (2015).[5] While the comparison of experts to non-experts allows me to exclude real quality differences as a reason for brand premia, it is often not possible to differentiate between situations in which consumers are inattentive and situations in which the consumer perceives quality differences (Handel and Schwartzstein, 2018).[6] In the former, information costs prevent consumers from accessing all the relevant information. In the latter case, even when consumers have information, uncertainty about the products' equivalence drives the choice. There is a growing literature that tries to isolate inattention (and switching costs) in markets of insurance and pension choices (e.g., Ho et al., 2017). Handel and Schwartzstein (2018) argue that in the case of insurance choices the underlying mechanisms

[3] Klemperer (1995) differentiates between five types of switching costs: (1) need for compatibility with existing equipment, (2) transaction costs of switching suppliers, (3) costs of learning to use new brands, (4) uncertainty about the quality of untested brands and (5) discount coupons and similar devices.

[4] The Swedish Medical Product Agency (2010) states that the 'basic principles for substitution are that the products have the same active substance in the same amount and are otherwise medically equivalent'. The inactive (non-medical) ingredients of drugs could differ.

[5] In this article, I use the term 'medical education' to refer specifically to the academic training of physicians.

[6] Handel and Schwartzstein (2018) report examples where both phenomena may play a role. For example, the difference in purchasing behavior of over-the-counter (OTC) drugs between pharmacists and the general population in Bronnenberg et al. (2015) could be caused by the lower information costs of pharmacists (who have lower search costs when comparing products) or by differences in perceived quality.

are still based on assumptions of the authors. Given the institutional setting, this study is unique in that I can credibly rule out inattention as a possible reason for brand premia. In Sweden, not only are patients with a prescription from a primary health care provider financially incentivized to consume the cheapest available product from a predefined group of equivalent drugs (referred to as a 'substitution group'), but also the pharmacist is obliged by law to tell the patient about the cheapest product as well as its equivalence (Sveriges Riksdag, 2002). Given that the information about equivalence is provided at the time of purchase and the cheapest equivalent pharmaceutical is always presented as a default option, inattention is not an explanation for brand premia. Therefore my approach differs from the existing literature on insurance choices (Abaluck and Gruber, 2016; Bhargava et al., 2017; Handel, 2013; Handel and Kolstad, 2015; Ho et al., 2017; Ketcham et al., 2015; Marzilli Ericson, 2014; Miller, 2019), not only because I investigate consumer choices of physical consumption products but also because inattention is not an option. Instead, compared with the general population, experts are 28.46% (5.47 percentage points) less likely to purchase painkillers that carry additional costs compared with the cheapest available product. For antibiotics, experts are 26.42% (2.49 percentage points) less likely than experts to pay an additional price difference, and for antiepileptics, experts are 39.49% (1.12 percentage points) less likely to oppose substitution to the cheapest available option.

Third, I quantify the effect of switching costs and brand premia on firms' pricing strategies. I estimate the demand for pharmaceutical products from experts as well as non-experts, incorporating switching costs. In a representative substitution group of high-dosage paracetamol, switching costs and brand premia are important. Using a two-step estimator similar to that of Bajari et al. (2007), I recover firms' cost parameters under the assumption that firms compete in an infinite horizon setting. I use the structural model to investigate two counterfactual scenarios that highlight the importance of consumers' behavioral frictions for price equilibria.

The first counterfactual evaluates effects on prices and consumer welfare by changing the procurement procedure for prescription drugs in the Swedish health care system. Under the hypothetical policy, firms are allowed to change the price of a product every 12 months instead of each month. Under this scenario, firms have a lower incentive to condition their prices on switching costs of consumers.[7] Intuitively one might think that switching costs make a market less competitive and that a policy lowering their importance would result in reduced prices. In contrast, the results show that prices increase and patients' welfare decreases under the alternative regime with only annual procurement. The results are in line with theoretical (Cabral, 2016; Rhodes, 2014) and empirical (Dubé et al., 2009) research that shows that moderate switching costs may lower prices. The basic reasoning is as follows. Switching costs always have two effects. On the one hand, they increase the market power of firms with locked-in consumers. On the other hand, they increase competition for new customers. If switching costs are moderate in size and there are

[7] I increase the length of procurement contracts instead of decreasing the switching costs directly because this approach represents a realistic policy intervention. Granlund and Rudholm (2018) propose a policy of a longer contract time and suggest that it decreases the possibility of coordination in the form of price cycles, empirically described in Janssen (2018).

enough new customers, then an increase in switching costs is associated with higher competitive pressure and lower prices in equilibrium. The results of this study add to the discussion about the competitive pressure of switching costs. Firms have a lower incentive to reduce prices in order to lock in consumers, and moderate switching costs prevent prices from becoming too high in the harvest periods. Further, not every patient is a frequent consumer. New incoming patients and those who do not experience switching costs profit from behavioral pricing as firms decrease their prices sporadically to lock in patients.

The results are important for the interpretation of switching costs from an industrial organization perspective as well as for policymakers: First, switching costs are empirically important when it comes to evaluating firms' price incentives. Second, switching costs do not necessarily increase prices in equilibrium on average. Third, I show that switching costs have a heterogeneous impact across consumers. While a slight majority of patients benefit from paying lower prices when firms follow a 'lock in and harvest' pricing strategy, patients with lower income, less education, and frequent prescriptions that likely relate to worse health conditions pay more on average. Firms take advantage of these consumers while others pay less. Policymakers must carefully evaluate the size and effect of switching costs in a pricing equilibrium. Further, in some markets, such as the health care market, policymakers should consider potential heterogeneous effects that may be connected to negative welfare effects on vulnerable consumers.

The second counterfactual investigates the situation in which all consumers are medical experts, limiting the role of brand premia in patients' choices. The hypothetical situation is closely related to directly targeting consumers' perceptions of quality differences or a reform that makes substitution mandatory. In such a situation, substitution is mandatory if not opposed by a physician. Limiting brand premia is associated with price decreases and an increase in patients' welfare (costs for consumers decrease by 4.05%). With less scope for brand premia, firms have less opportunity to take advantage of patients. A decrease in brand premia is profitable for the consumers through a direct effect of lower absolute payment and an indirect effect caused by an increase in competition as firms cannot rely on brand premia.

Finally, I connect the product-specific reduced-form estimates with the counterfactual estimation using the structural model of the paracetamol market. The connection allows a back-of-the-envelope calculation that sheds light on aggregate effects of brand premia and switching costs on patients' and public health care expenditures. I first show that brand premia and switching costs are commonly observed in most product groups, especially for products such as opioids, paracetamol products and common antibiotics. From an aggregate expenditure perspective, brand premia lead to larger aggregate changes of expenditures for patients and the public.

2 Pharmaceuticals in the Swedish Health Care System

Health care coverage in Sweden is universal and mainly publicly funded. Prescription drugs are reimbursable.[8] Patients' co-payments depend on the yearly costs for pharmaceuticals. The higher the yearly costs, the lower the fraction a patient has to pay out of pocket. Costs above a ceiling are entirely covered.[9] Almost half of the revenue from prescription pharmaceuticals is from coverage for patients who have reached the cost ceiling (Bergman et al., 2012). Due to the low ceiling, a function of decreasing co-payments and generally lower drug prices, we do not observe nonlinear price schedules as occur in the United States under Medicare Part D.[10]

The market for off-patent drugs is arranged into substitution groups (groups of pharmaceuticals with the same substance, size and strength) that are determined by the Medical Product Agency. Each product within a substitution group is medically equivalent. Only nonactive ingredients such as coloring or sugar content may differ; however, medical equivalence is tested (Swedish Medical Product Agency, 2010). For specific drugs, such as for some antiepileptics, a incremental variation of ingredients could have adverse health effects. In those cases, each product forms its own group and substitution is not possible. In the case of pharmaceuticals with generic competition where substitution is medically appropriate, patients with a prescription for a product in a substitution group are incentivized to acquire the cheapest available generic in the substitution group. Substitution of the cheapest product happens at the pharmacy level (Sveriges Riksdag, 2002). A pharmacist is obliged to recommend the cheapest available generic to patients. Further, pharmacists are required to explain the medical equivalence of the products. However, not every patient receives the cheapest available product. Three possible reasons may prevent a substitution. First, patients may oppose substitution. In this case, patients have to pay the price difference between the cheapest available generic and the prescribed product out of pocket. Only the price of the cheapest available product is subject to the co-payment structure. Second, a physician or health care provider has the option to oppose substitution.[11] Third, the cheapest available product may not be in stock at the pharmacy. In such a case, the second cheapest product is dispensed.

The pharmaceutical market had a market capitalization of $4.15 billion in 2015, with $3.08 billion due to prescription drugs; patients' co-payments accounted for $0.64 billion (TLV, 2016*a*). Since 2002, a tendering system based on a sealed-bid first-price auction determines prices of off-patent drugs in each substitution group. A pharmaceutical company that wishes to sell a product submits a price at the end of a month (Month $t = 0$) for the month after the

[8] Note that the exact product reimbursement is subject to the decision of the Dental and Pharmaceutical Benefits Agency (TLV). Some products are only partly reimbursed (TLV, 2016*c*).

[9] If patients' pharmaceutical costs exceed 2200 SEK (approx. 220 USD), additional costs are covered without out-of-pocket expenses.

[10] Drug purchases in Medicare Part D are characterized by the 'donut hole,' a coverage gap for patients. Purchases of drugs are lower around this coverage gap as consumers decrease purchases as well as substitute drug expenses to the following year (Dalton et al., 2020; Einav et al., 2015; Zhang et al., 2009).

[11] One may ask why this possibility exists if products are equivalent. Historically, pharmacists were required to follow a physician's prescription exactly, including the choice of brand. Since the reform, substitutions are possible, but doctors can oppose them. However, physicians seldom oppose substitution.

next (Month $t = 2$). Companies have to bid a lower price than the specified ceiling.[12] The auctioneer publishes a preliminary list of prices in the middle of the next month (Month $t = 1$). After the supplier of the cheapest product confirms the ability to service the entire Swedish market, prices are implemented.[13] Note that the pharmaceutical companies see the preliminary list for Month $t = 2$ before bidding for the month after that (TLV, 2016*b*).

Pharmacy purchasing prices for reimbursable, patent-protected pharmaceuticals are also regulated. Manufacturers apply with a specific price for a patent-protected product. The Dental and Pharmaceutical Benefits Agency (TLV) determines whether the price requested by a manufacturer is reasonable. Assessment is based on a 'value-based pricing approach' that considers principles of equality, solidarity and cost-effectiveness (TLV, 2017).

Retail prices are regulated and dependent on the prices in the auction (or the value-based pricing for patent-protected products) as an almost linear function of pharmacy purchasing prices. The difference between the retail and pharmacy purchasing prices provides the pharmacy's trade margin.[14] Although some pharmacies were privatized in 2009, pharmacists are required to dispense the cheapest available product (Sveriges Riksdag, 2002). As noted, the pharmacy can dispense the second cheapest product if the cheapest product is not in stock.[15]

3 Data

I use data for painkillers (ATC code N02), antibiotics (ATC code J01) and antiepileptics (ATC code N03) from January 2010 through June 2016 in Sweden.[16] The data are provided by Socialstyrlsen, the Swedish governmental agency for health and welfare. To restrict the data to reimbursable pharmaceuticals, I connect the choice data to monthly prices/bids for outpatient pharmaceuticals under generic competition, provided by the TLV. Each individual who is covered by the universal health care system and purchases a product at a pharmacy with a prescription from a health care provider is observable.

The patient-specific data are collected on the pharmacy level and reported to health authorities. After a transaction, a pharmacist records a personal identifier, the dispensed product and the time of the transaction. If an individual receives the cheapest available generic within a substitution group, the pharmacist solely records the dispensed product. In this case, the prescribed product is not recorded. However, by choosing not to oppose substitution, the physician

[12] The usual ceiling is 35% of the original brand product price before expiration of the patent. A price ceiling exists if a branded drug has had generic competition for at least four months and the prices of the drug have fallen by 70% of the original price 12 months prior to patent expiration. If no price ceiling exists, the most expensive product of the month will form the price ceiling. If an original product does not have sufficient generic competition, prices may also be reduced (7.5% reduction) if its market authorization was at least 15 years ago (TLV, 2016*d*).

[13] If a firm confirms delivery but fails to do so, it is subject to a penalty fee. Before 2014 the confirmation was not part of the process.

[14] I describe the exact relationship between purchasing and retail prices in online Appendix A.

[15] Additionally, a pharmacy can sell the remainder of the previous month's cheapest product in the first two weeks of a new month. Afterward, pharmacies can sell the products at the pharmacy's purchasing price without profit. The pharmacy has no incentive to overstock the product.

[16] The selection of the three sub-markets is not intentional and based on data provision by Swedish authorities. The ATC code describes each pharmaceutical substance and is ordered according to five levels. The first level describes the anatomical main group (e.g., nervous system); the second level, the main therapeutic group (e.g., analgesics/painkillers); the third level, the pharmacological subgroup; the fourth level, the chemical subgroup; and the fifth level, the exact chemical substance.

has acknowledged the suitability of substitution to the cheapest available product. If a physician opposes substitution, the pharmacist records the dispensed product and the physician's decision to oppose substitution. If a patient opposes substitution, the pharmacist records the dispensed and prescribed products as well as the fact that the patient opposed substitution. Finally, also, the pharmacy can dispense a product that is not the product of the month if the cheapest product is out of stock. Also in this case, the pharmacist records why substitution to the cheapest product did not happen. Therefore, the data allow me to directly track if and why substitution is opposed.

By connecting the choices to Swedish registry data provided by Statistics Sweden (SCB), I observe the patient's place of residence (county and municipality) and some socioeconomic characteristics. The socioeconomic data consist of yearly income and length, degree and subject area of education. Finally, for non-prescription drugs, I use advertisement expenditure data provided by Kantar Sifo. These data include advertisement expenditures for non-prescription drugs by pharmaceutical brands between 2010 and 2016.

Panel A of Table I shows a basic market description. In the study period, the average number of products per substitution group was 3.22 for painkillers, 2.44 for antibiotics and 1.97 for antiepileptics. For all three therapeutic groups, the majority of products are generics. Originals are on average more expensive than generics.

Panel B of Table I shows basic demand characteristics for all three markets. The number of purchase occasions is much higher for painkillers (approximately 39.5 million) and antibiotics (approximately 13.6 million) than for antiepileptics (0.6 million). Also, the number of unique patients who received an antiepileptic is much lower than for painkillers or antibiotics (approximately 61,000 for antiepileptics, compared with approximately 3.2 million for painkillers and 4.7 million for antibiotics).[17] Whereas painkillers and antiepileptics are often repeatedly purchased by the same individual, antibiotics have a much lower average number of purchase occasions. Note also that a large variability exists among painkillers. The co-payment share is much lower for antiepileptics than for antibiotics and painkillers. One reason for the difference in co-payments is the upper ceiling for medical expenses. Patients that have a longer treatment may reach the upper ceiling earlier and therefore have lower co-payments for pharmaceuticals. Panel C of Table I shows the substitution behavior of patients for the three therapeutic groups. In line with the regulatory intent, the majority of patients receive the cheapest product, the product of the month (PoM). However, heterogeneity between and within the therapeutic groups is observable: 75.5% of patients purchased the PoM when getting a prescription for painkillers, whereas 86.7% and 92.8% of patients purchased the PoM for antibiotics and antiepileptics respectively. Patients who do not receive the PoM belong to one of the following three groups: they opposed substitution and pay the difference out of pocket, their physician opposed substitution, or the pharmacy opposed substitution because the PoM was out of stock. Patient opposition to substitution is the most common in all three therapeutic groups.

[17] The population of Sweden was approximately 9.85 million in 2016.

Table I: Summary Statistics

	Painkillers	Antibiotics	Antiepileptics
A: Product Level			
Number of Substances	10	24	4
Number of Substitution Groups	162	167	39
Number of Products	629	513	96
Average Number of Products in Substitution Groups	3.22	2.44	1.97
	(2.35)	(1.8)	(0.8)
Percent Originals	0.21	0.15	0
	(0.41)	(0.35)	(0)
Percent Generics	0.76	0.72	0.82
	(0.43)	(0.45)	(0.39)
Average Price (in SEK)	288.86	231.62	390.37
	(342.62)	(291.11)	(383.73)
Average Price of Originals (in SEK)	357.91	331.36	-
	(425.91)	(312.03)	
Average Price of Generics (in SEK)	266.45	195.46	377.77
	(313.84)	(274.02)	(395.74)
B: Consumer Level			
No. of Purchase Occasions (in millions)	39.46	13.62	0.56
No. of Patients (in thousands)	3,233	4,722	61
Avg. Purchase Occasions per Patient	12.2	2.9	9.2
	(26.14)	(3.48)	(14.08)
Avg. Monthly Total Costs (in million SEK)	47.67	23.04	2.78
Avg. Monthly Co-payment (in million SEK)	18.3	14.44	0.7
C: Substitution Decisions			
Fraction Consuming the Product of the Month	0.755	0.867	0.928
	(0.43)	(0.339)	(0.258)
Fraction with Substitution Opposed by Patient	0.192	0.094	0.028
	(0.394)	(0.292)	(0.166)
Fraction with Substitution Opposed by Physician	0.021	0.005	0.018
	(0.145)	(0.07)	(0.133)
Fraction with Substitution Opposed by Pharmacy	0.031	0.034	0.025
	(0.173)	(0.18)	(0.157)

Notes: The table shows summary statistics for the three market segments of painkillers, antibiotics and antiepileptics. Prices are in SEK (10 SEK are approx. 1 USD). Panel A describes markets on the product level. Branded generics are included in the generics category. Panel B shows summary statistics considering patients' purchasing decisions. The number of purchase occasions is the sum of purchase occasions across consumers between January 2010 and June 2016. Average purchase occasions are the number of purchases by the average patient. The total costs include costs for insurance as well as for the patient (co-payment). Panel C describes substitution decisions. The fraction of consumption of the product of the month (PoM) describes the fraction of purchase occasions in which a patient consumed the PoM. If a patient does not consume the PoM, it is because of one of the three displayed reasons (substitution opposed by the patient, substitution opposed by the physician or substitution opposed by the pharmacy). Standard deviations are in parentheses.

4 Empirical Analysis

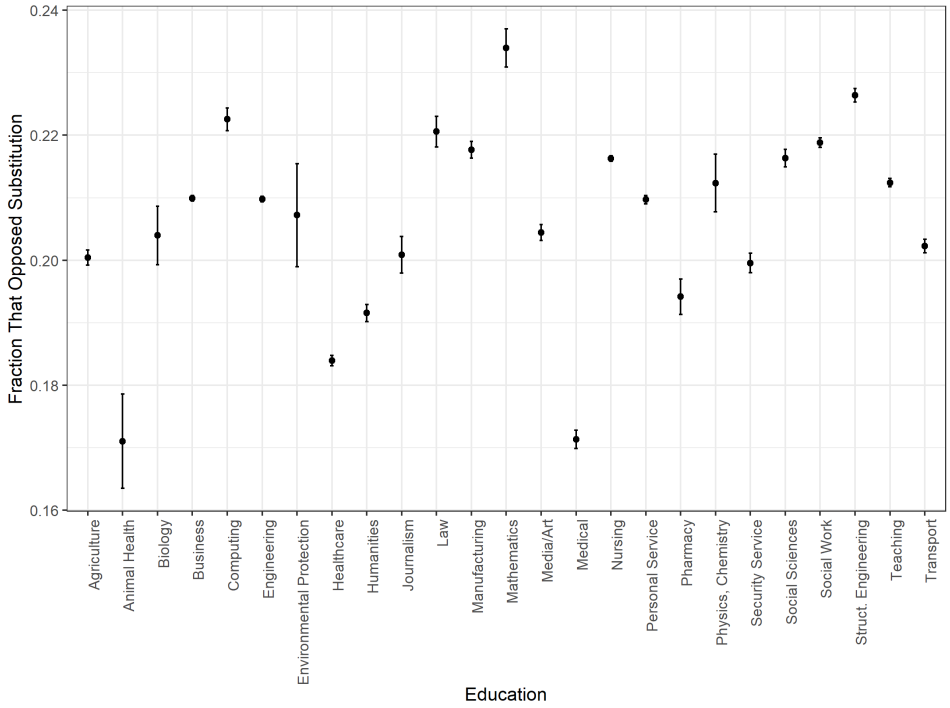
I now turn to estimate switching costs and brand premia. Considering switching, my analysis of the Swedish pharmaceutical market differs from existing analyses in several dimensions. Patients in the Swedish market for prescription drugs under generic competition get a prescription from their primary health care provider and can choose between products within a pharmacy. The substitutability of products is determined by a medical governmental agency and further implicitly acknowledged by the prescribing physician, who can oppose substitution. When a patient chooses a product at a pharmacy, the cheapest available product is first presented and recommended to the patient as the default product. As the analysis is on the substitution group level, and the products have the same ingredients, I do not consider that patients learn about the perfect match of a ingredient for their own usage. This stands in contrast to the work by Crawford and Shum (2005) who use different molecules on the market for anti-ulcer drugs and show that patients experience learning. Nevertheless, I acknowledge that patients' switching costs decrease after they switch products within a group and learn that switching is not related to worse treatment outcomes. In comparison to other retail markets, where habit persistence may also be due to inattention, the pharmaceutical market is therefore characterized by a clean and structured environment. Patients may have higher trust in a specific product after a successful treatment. Switching costs are therefore a psychological cost that a consumer bears after starting treatment with a specific product. Importantly, switching costs are induced by previous usage.

In addition to the estimation of switching costs, I add to the literature by identifying the effect of perceived quality differences on consumers' pharmaceutical choices. In the Swedish health care market, exchangeable pharmaceuticals are defined by a medical product agency, and a prescriber is able to prevent substitution. Labeling and advertising are regulated, and manufacturers have limited possibilities to differentiate themselves. Advertisement of prescription drugs is not allowed in Sweden. Therefore this study differs from most of the evidence from retail markets where products could potentially have a higher degree of differentiation. Nevertheless, specific pharmaceutical brands may carry a higher quality in physical or psychological terms.[18] It is possible that products have quality differences when using the approach of Bronnenberg et al. (2015). My approach uses the same method; however, instead of evaluating drug choices dependent on the patient's occupation, I use medical education as the proxy for an informational advantage. I control for true quality differences by comparing choices made by experts (patients with a medical education) to choices made by others. Figure I provides some preliminary graphical evidence for the effect of education on substitution decisions for painkillers. Patients with medical education are less likely than other patients to oppose substitution.

After controlling for the first explanation of brand premia, I note that it is still possible that consumers cannot

[18] A original branded pharmaceutical may have a higher quality, or patients may view an original product as having a greater treatment effect than generic alternatives. Some placebo effects may be due to the branded product (Branthwaite and Cooper, 1981; Kamenica et al., 2013).

Figure I: Subject of Education and Substitution Decision



Notes: The figure shows the fraction of patients that oppose substitution when consuming painkillers, dependent on the subject of their education. 'Medical' refers specifically to the education needed to become a physician. The error bars correspond to a 95% confidence interval.

evaluate all information or that consumers experience perceived quality differences. Bronnenberg et al. (2015) are not able to differentiate between these two explanations. It is possible that the consumers of the non-prescription headache remedies in the study by Bronnenberg et al. (2015) do not evaluate all available information, like reading the ingredients or comparing all prices on the retail shelf. On the other hand, people may have perceived quality differences. Comparing experts with non-experts allows for both explanations, as experts may have lower information costs as well as fewer perceived quality differences (Handel and Schwartzstein, 2018). Because of the regulation in the Swedish prescription drug market, I am able to identify perceived quality differences and exclude information costs and inattention as explanations. Indeed, the patient has no information costs, so I am able to exclude the reasoning that patients have high information costs and are therefore inattentive. Patients who oppose substitution and pay a brand premium are doing so because they either believe in the superiority of a product or are uncertain about the real equivalence and decide to stick to a safe option.

In the following, I show reduced-form evidence. The estimation strategy allows me to identify switching costs and perceived quality differences. Within the reduced-form analysis, I evaluate choices on the dispensing level, meaning that each dispensed prescription is one observation. In the demand estimation in Section 6.1, I use a discrete-choice model where I model choices in detail. The major advantage of the reduced-form framework is that it evaluates the importance of behavioral frictions without estimating a demand system. Consider the following model:

$$P(OpposeSubst_{ist} = 1) = \alpha + \beta_1 D_{ist-1} + \beta_2 Med_{it} + \beta_3 D_{ist-1} \times Med_{it} + \rho X_{it} + D_{is0} + \gamma_{st} + \varepsilon_{ist}, \quad (1)$$

where the outcome variable is a dummy that takes the value 1 if individual i in period t opposes substitution to the cheapest available product of substitution group s. In substitution group s all products of the unique size $\times$ strength $\times$ ingredient combinations are grouped. Opposing substitution corresponds to the case in which the patient pays the price difference to purchase another product instead of the PoM. D_{ist-1} is a dummy that takes the value 1 if individual i chose the same product in the last purchase that occurred within a month of the current purchase and 0 if not.[19] Med_{it} is the variable of interest, which examines the impact of whether a patient has had medical education as a physician. X_{it} are controls such as the logarithm of income, the geographical location and the general education level.[20]

I use fixed effects of substitution groups in each time period (γ_{st}). Therefore I examine variation among individuals who purchase a product within a given month. Note that the fixed effects absorb important factors that may affect the individuals' switching behavior or perceived quality differences, for example, the price differences between products.

State dependence due to switching costs is measured by β_1 and perceived quality differences by β_2. It is important to separately identify switching costs and perceived quality differences. To identify switching costs, it is important to consider unobserved individual heterogeneity that may be correlated with the lagged product choice dummy D_{ist-1}.[21] I tackle this problem in two ways. First, I recover true state dependence due to switching costs by conditioning decisions to oppose substitution on the initial choice D_{is0}, a dummy that takes the value 1 if patient i consumed the same product as in t in the first observed period. Intuitively, unobserved heterogeneity beyond state dependence would be captured by the initial choice of a product. However, the approach assumes that the initial choice is random and not correlated with unobserved heterogeneity (initial condition problem).[22]

To estimate perceived quality differences, I evaluate whether patients with medical education have a higher or lower likelihood of opposing substitution towards a cheaper equivalent. Having excluded that inattention plays a role, as patients get their information at the pharmacy level, I measure the perceived quality differences by β_2. I include

[19] This model requires that the patient purchased the product in the previous month. If I also considered the same calendar month, the choice environment would be the same. Fixed effects would capture the corresponding differences in the outcome variable. Within this specification, I look at short-term induced state dependence with potential different choice environments, as the prices and PoM may have changed.

[20] I control for the geographical location on the county level to control for differences across counties. While prices are identical across regions of Sweden, some pharmacy retail brands differ across regions. Thus, the geographical fixed effects control to some degree for differences across pharmacy brands.

[21] For example, patients may have brand preferences and choose the same product repeatedly. See further discussion in Dubé et al. (2010).

[22] As I do not observe the entire medical history, this is unlikely the case for all patients. Following Wooldridge (2005), it is possible to use a reduced-form approach of conditioning the unobserved household effects on the initial values and exogenous variables. Rabe-Hesketh and Skrondal (2013) recommend a slight adjustment of using the within means of time-varying variables as well as including the initial periods. In comparison to approaches in the literature, such as those of Erdem and Sun (2001), Skrondal and Rabe-Hesketh (2014) or Rickert (2016), I do not observe sufficient time-variant covariates of products, as they do not differ in any aspects besides their brand. The homogeneity of products and the clean choice environment reduce the problem of unobserved heterogeneity and the endogeneity of the initial condition.

the interaction of D_{ijt-1} and Med_{it}. As state dependence is induced through past consumption and not affected by perceived quality differences, I expect that β_3 is not significantly different from zero.

I show results of the reduced-form evidence of model 1 in Table II. For each therapeutic group (painkillers, antibiotics, and antiepileptics) I show results of the full model including all fixed effects and covariates. A purchase of a product within the last month increases the probability of opposing substitution for painkillers significantly by 3.1 percentage points (16%). The increase is an average across all painkiller substances.[23] For antibiotics, previous consumption is associated with a significant increase of 1.1 percentage points (11.5%). For antiepileptics, the results are different: a previous purchase of the same product in the last month decreases the probability of opposing substitution. I interpret the positive coefficient of painkillers and antibiotics as real state dependence due to switching costs. For antiepileptics, previous consumption is negatively associated with opposed substitution, even though the size of the effect is small. One explanation for the lack of switching costs is that patients' learning plays a role. In detail, patients use antiepileptics in the long term. Those patients that use antiepileptics in the long term may have frequently switched in the past and therefore know about the equivalence within a substitution group. It is also possible that the patients are used to acting in response to the information of a pharmacist who recommends switching. Therefore, long-term usage lets patients learn that switching within a product group is possible.

Secondly, Table II shows that medical education is associated with a significantly lower probability of opposing substitution (painkillers: 5.5 percentage points, 28.5%; antibiotics: 2.5 percentage points, 26.4%, antiepileptics: 1.1 percentage points, 39.5%). I interpret the difference as perceived quality differences. Note that the interaction between previous consumption and medical education is insignificant for all three therapeutic groups. In other words, medical education is not associated with higher or lower switching costs. The lack of effect supports that perceived quality differences and switching costs induced by previous consumption are not correlated.[24]

4.1 Quasi-experimental Evidence of Switching Costs

To show robustness, I use a quasi-experimental setting to show causal evidence for state dependence due to switching costs for a small part of the substitution groups, those where a patent expired between 2010 and 2016. Similar to Feng (2022), I take advantage of the introduction of generic products after patent expiration. Correspondingly, I compare the repeated purchasing behavior of consumers who start their treatment with a drug for the first time shortly before expiration of a patent to those patients who start treatment shortly after the patent expired.

I begin by describing a specific but representative example of a substance for which a patent has expired. Figure

[23] The share of patients that purchase repeatedly, independent of the probability of opposing substitution, could be higher for some products due to a higher addiction potential. Considering evidence from actual repeated purchases, I do not observe a direct increase of repeated purchases with addiction potential. For example, compared with opioids, paracetamol products are on average more often purchased repeatedly.

[24] In online Appendix B.1, I show that patients with an education as a pharmacist are also less likely to oppose substitution. However, the effect is smaller than for medical experts. Further, patients with a nursing degree show no differences from the general population.

Table II: Regression Evidence, Probability of Opposed Substitution

	Opposed Substitution		
	Painkillers	Antibiotics	Antiepileptics
D_{t-1}	0.0309^{***}	0.0108^{***}	-0.0060^{***}
	(0.0002)	(0.0005)	(0.0005)
Med	-0.0547^{***}	-0.0249^{***}	-0.0112^{***}
	(0.0017)	(0.0007)	(0.0041)
$log(Inc)$	-0.0004^{***}	0.0006^{***}	-0.0001
	(0.0001)	(0.00003)	(0.0001)
$D_{t-1} \times Med$	0.0028	0.0018	0.0056
	(0.0039)	(0.0036)	(0.0066)
Education	Yes	Yes	Yes
Control Heterogeneity	Yes	Yes	Yes
Geographic Fixed Effects	Yes	Yes	Yes
Fixed Effects	SubGroup×Year-Month	SubGroup×Year-Month	SubGroup×Year-Month
Mean Opp. Subst.	0.192 (0.394)	0.094 (0.292)	0.028 (0.166)
Mean Price SEK	99.686 (138.94)	122.691 (83.91)	373.681 (211.07)
Mean Overpayment SEK	9.475 (25.102)	10.766 (21.791)	7.679 (26.639)
$D_{t-1}Increase$	16.04%	11.52%	-21.04%
Med Increase	-28.46%	-26.42%	-39.49%
N	36,471,580	12,695,230	530,263
R^2	0.2288	0.1215	0.0616

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Linear least squares regression results for the segments of painkillers, antibiotics and antiepileptics. One observation corresponds to one specific purchase occasion by a patient. The outcome variable is a dummy variable that takes the value one if a patient opposes substitution in order to receive a more expensive product. The patient bears the additional costs. D_{t-1} is a dummy that takes the value 1 if a patient has consumed the product in the previous purchase occasion in the last month. Med is a dummy that takes the value 1 if an individual has medical education. Log(Income) is the logarithm of income. Education indicates if the model controls for education categorized into six levels. Geographical indicates if the model controls for county-level fixed effects. The lower part of the table shows the average fraction of opposed substitution as well as the price and average payment of those that oppose substitution (in SEK; 10 SEK equal approximately 1 USD). Finally, I also state the percentage increase of opposed substitutions associated with past consumption (switching costs) and medical education (quality misconceptions). Standard errors are clustered on the individual level, adjusted for heterogeneity, and shown in parentheses.

IIa describes the monthly market shares of different brands within a substitution group of the substance oxycodone. Oxycodone is an opioid, used as a strong painkiller. OxyContin, the branded version of oxycodone, had patent protection until 2012. We see a steady reduction of the original's market share after the patent expired. The main reason for this development is that consumers began to be reimbursed for the cheapest product, which was the newly authorized generic product.

In the following, I start with identifying those substitution groups where a patent has expired.[25] I use the substitution groups of oxycodone, a strong opioid. The original product, OxyContin ran off patent in early 2012. I reduce the sample to the purchases by those patients who started purchasing a product three months before or after patent expiry. The treatment group is the group of patients who started with one of the entering generics three months after the patent ended. Correspondingly, the group of patients starting treatment three months before the expiration of the patent is the control group. The basic intuition is that the groups are randomized, as patients cannot influence their medication starting date. I consider the repeated purchases of the two different groups after the initial purchase:

$$Y_{it} = \alpha + \beta_t Y_{i0} + \varepsilon_{it}$$

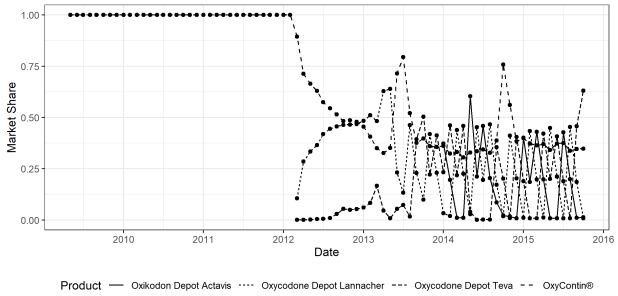
The variable Y takes the value 1 if a patient i consumes one of the newly entered products (generics). Given that I use the sample of patients whose initial consumption lies between three months before and after a patent expiry, t is the number of months after the initial consumption, which is designated as $t = 0$. I instrument Y_{i0} by Z_i, which shows whether a consumer is in the treatment group. Note that the sample only includes repeated purchases by consumers. The exclusion restriction of the approach relies on the assumption that patients are randomly selected into the treatment and control groups. Given that patients need prescriptions from a primary health care provider for all three substances and given that patients in need of these substances need immediate treatment when prescribed, it is unlikely that patients self-select into one of the two groups.

For each substitution group I run the regression. The first stage is strong for all specifications.[26] Figure IIb show coefficients for β for the first three months (β_3) and the fourth to sixth months (β_6) after the initial purchase for each substitution group containing the substance oxycodone. The results of the second stage suggest that there is a strong state dependence during the first three months in the oxycodone substitution groups. For the longer time horizon, the state dependence diminishes. In general, patients who start with a generic are more likely to consume a generic three months later than those who started with an original product. The analysis is in line with the general regression evidence. Patients experience habit persistence.

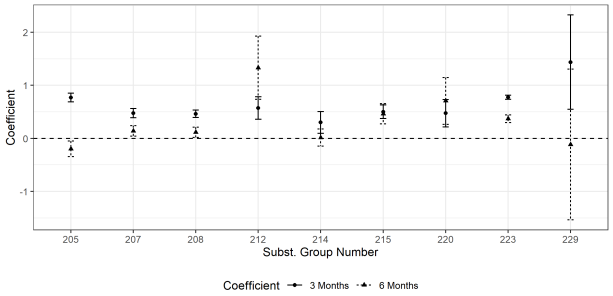
[25] In the Online Appendix I also show evidence for additional products with patent expiries: Rizatriptan (a painkiller used against migraines) and clindamycin (an antibiotic). Results are similar.

[26] Note that I include detailed tables of the first and second stage for each of the substitution groups in online Appendix B.2. Results are similar across substitution groups.

Figure II: Quasi-experimental Evidence of Switching Costs



(a) Market Shares, Oxycodone



(b) Instrumental Variable Regression Results, Oxycodone

Notes: The first subfigure shows monthly market shares in the substitution group of oxycodone before and after patent expiry of the original product, OxyContin. Each observation stands for the monthly market share of a brand. The lower graph shows second-stage coefficients for different substitution groups of oxycodone. Coefficients for each substitution group are divided into coefficients for the initial three months and for months four to six. A coefficient for the first three months equal to $\beta_3 = .5$ indicates that the initial consumption of a generic increases the possibility of purchasing a generic again during the following three months by 50%. First-stage results are reported in online Appendix B.2. Error bars show 95% confidence intervals.

5 A Structural Model of Demand and Supply

This section introduces a model that allows me to structurally estimate the demand as well as the supply side of the Swedish pharmaceutical market. The reasons for the use of a model are manifold. Most importantly, it is possible to relate the phenomenon of switching costs to general price levels, that is, one may address the open research question of whether switching costs make markets more or less competitive. Further, modeling the supply side allows me to evaluate how firms' pricing strategies and consumers' costs would change under counterfactual scenarios. This section is divided into two parts. I first present the model setup of the demand side and then present the supply-side model.

5.1 Demand

The utility of individual i purchasing product j in substitution group s at time t is defined as

$$u_{ijst} = \gamma_{ijs} + \rho_{is} y_{ijs,t-1} + \alpha_{is} p^{C}_{jst} + h_{ijs} + \varepsilon_{ijst}.$$

Consumers are myopic and not forward looking.[27] Each product j is part of a set of products that form a substitution group. As almost all prescriptions in Sweden are filled (Ekedahl and Månsson, 2004; Ax and Ekedahl, 2010), I assume

[27] A prescription is connected to a fixed quantity. Patients do not have the possibility to stockpile.

that there is no outside good. Each coefficient of utility varies at least over the substitution group. First, there is a random brand-specific intercept γ_{ijs}. The variable $y_{ijs,t-1}$ is a dummy that takes the value 1 if a consumer i has already purchased product j during the last calendar month. ρ_{is} captures the impact of switching costs and varies across each patient within a substitution group. p^C_{jst} is the price of product j at time t corresponding to the patient-specific price based on the co-payment and eventual price differences between the cheapest and more expensive products. The price coefficient α_{is} also varies across consumers within a substitution group. h_{ijs} denotes the unobservable heterogeneity of patients and product characteristics, and ε_{ijst} is an error term.

Estimation requires two adjustments to ensure identification of the price elasticity and switching cost estimates. It is possible that prices p^C_{jst} are correlated with unobserved product characteristics, that is, $E[p^C_{jst} h_{ijs}] \neq 0$. The second bias is due to the correlation of the lagged product choice and the unobserved heterogeneity of individuals. Some patients may have characteristics that lead to a repeated choice of a specific product. The repeated choice would not be due to switching costs but due to personal characteristics.

I tackle the identification threats by two methods. First, I use a control function approach to deal with the endogeneity of prices. In the first stage of the control function, instruments Z_{jst} are prices of other products from the same brand for the same therapeutic segment (painkillers). The instruments are comparable to the Hausman instruments (Hausman, 1996). If a brand produces several painkillers in different strengths, forms or sizes, the prices of the products in other substitution groups are used as instruments. The intuition of the assumption is that all products of a brand have correlated prices due to shared cost factors (supply chain, procurement of substances); however, their demand is uncorrelated. The control function takes the following form: $p^C_{jst} = Z_{jst}\gamma + \rho_j + \mu_t + \kappa_{jst}$, where ρ_j and μ_t are product and time fixed effects respectively. The exclusion restriction requires that the idiosyncratic error term κ_{jst} be independent from Z_{jst}, $E[Z_{jst}\kappa_{jst}] = 0$. As an individual prescription is for a product in a specific substitution group, it is likely that there is no demand effect between substitution groups. Also, effects of advertisement should not violate the exclusion restriction.[28] As usual for the control function, the residuals κ_{jst} enter the main estimation equation, and the error term of the main equation is adjusted accordingly.[29]

Second, I try to control for unobserved heterogeneity among consumers. As in Section 4, I control for the initial product choice of individuals that I observe in my sample. The final structural equation that incorporates the control function approach, as well as controls for unobserved heterogeneity, takes the following form:

$$u_{ijst} = \gamma_{ijs} + \rho_{is} y_{ijs,t-1} + \alpha_{is} p^C_{jst} + \mu_{is} y_{ijs,FIRST} + \lambda \kappa_{jst} + \varepsilon_{ijst}, \tag{2}$$

[28] Advertisement may affect demands for brands in all substitution groups such that prices as well as unobserved demand characteristics are correlated. The strong regulation of the Swedish pharmaceutical market does not allow for advertisement of prescription drugs, i.e., the pharmaceutical products considered in this study.

[29] In detail, let the old error be $\varepsilon_{ijst} = \lambda \kappa_{jst} + \varepsilon_{ijst}$. As p^C_{jst} is a function of Z_{jst} and u_{jst}, it is uncorrelated with the new error ε_{ijst}.

where $y_{ijs,FIRST}$ is a dummy that takes the value 1 if the patient has taken product j in the first observable period in the sample when the consumer purchases a product in substitution group s.

5.2 Supply

In each period $t = 1, ..., \infty$, there are N_{ts} firms in substitution group s. Given that supply is separate for each substitution group, I drop the subscript s. Each firm $j = 1, ..., N_t$ sets a price p_{jt} at t. The value of p_{jt} has to be lower than a regulatory price ceiling R. I model the supply side with independent substitution groups for two reasons. First, it is reasonable that demand of substitution groups is independent, as prescriptions are for a specific substitution group. Second, the assumption allows for a tractable model solely investigating a specific substitution group. Because retailers (pharmacies) get a fixed markup for each dispensed product, I do not model pharmacies as separate agents. Within the estimation I calculate for each wholesale price p_{jt} a retail price that is a structural parameter of the demand.[30] However, for simplicity I do not denote the difference between the manufacturer price and the retail price.

The per-period profit of a firm at period t is defined as

$$\pi_{jt} = [p_{jt} - c_{jt}] m_{jt} Q_t,$$

where c_{jt} represents the marginal costs of firm j in t, and m is the market share of j at time t. Finally, the exogenous quantity (measure of market size) of the substitution group s is given by Q_t.

Before turning to the continuation profits of a firm, I make two assumptions. First, costs are defined as the sum of a constant and a random privately observed shock within each period. So, beside the marginal costs c_{jt}, which differ across time and brands, a random shock $\varepsilon_{jt} \sim N(0,1)$ enters the marginal costs. The assumptions about the marginal costs are rather weak. I allow for changes of marginal costs over time as well as differences between periods.

The second assumption considers firms' beliefs about future demand. The beliefs affect the continuation payoffs as firms form expectations about future payoffs. In each time period, firm j makes a decision about setting its own price. Such a dynamic game has a continuum of Nash equilibria. Following previous literature (Maskin and Tirole, 1988; Ericson and Pakes, 1995) I reduce the equilibrium space to symmetric Markov perfect equilibria. One restricts subgame perfect equilibria to only the payoff-relevant strategies of a subgame. State variables are sufficient to determine a payoff. In detail, firms condition their strategy σ_j on the cost shock ε_{jt} and the state variables $\mathscr{S}_{jt}$, which include the lagged market shares (m_{jt-1}), the lagged number of firms ($|N|_{t-1}$), a dummy that indicates whether firm

[30] The difference between the demand-side and supply-side prices is determined by the trade margins of the pharmacies; see online Appendix A. For the substitution groups of interest (under consideration of the price ceiling), the relation between purchasing (supply side) and retail price (demand side) is linear. Given the retail price, consumers pay co-payments that depend on their yearly expenses and whether they choose the cheapest available product.

j had the cheapest product in the previous period (PoM_{jt}) and the total market size of the segment (Q_t). Formally, the strategies are defined as the mapping of the state variables and the cost shock to the prices ($\sigma_j : (\mathscr{S}_{jt}, \varepsilon_{jt}) \rightarrow p_{jt}$). Given that firms discount future profits with $\delta \in (0,1)$, the value function of firm j is

$$V_{jt}(\mathscr{S}_{jt}, \varepsilon_{jt}) = (p_{jt} - c_{jt}) m_{jt} Q_t + \delta E[V_{jt+1}(\mathscr{S}_{jt+1}, \varepsilon_{jt+1} | \mathscr{S}_{jt})].$$

The first term is the per-period profits. The second term describes the expectation from the valuation at period $t+1$. It incorporates the expectation of how the state vector evolves. I make distinct assumptions about firms' beliefs regarding the development of the market share to reduce the computational burden and the state space. I assume that firms are not able to predict future patients and their random coefficients perfectly. However, firms have knowledge about important key factors of the dynamic demand. I assume that firms (1) know the share of consumers that stay in a market at t, (2) have knowledge about the average coefficients and co-payments of the demand side described in the previous section and (3) know about the average product choice of consumers who have started a treatment in their first period. Correspondingly, the expectation about future market shares is a discrete function, dependent on the transition probability of consumers between periods as well as the demand estimates:

$$E[m_{jt+1}, \mathscr{S}_{jt+1} | \mathscr{S}_{jt}] = \phi_t E[\tilde{m}^S_{jt+1}] + (1 - \phi_t) E[\tilde{m}^{NS}_{jt+1}]. \tag{3}$$

Here, ϕ_t of the consumers at $t+1$ have been already present in t. Those consumers who were present in the last period consume product j $\tilde{m}^S_{jt+1}$ times in period $t+1$. The market share is evaluated from firm j's point of view at time period t. Firm j does not know the customer base in the forthcoming period and approximates it by the average customer of the current period. The demand model presented in Equation 2 with average coefficients among customers is used for calculating the market shares. For parameters of the patient-specific first consumption $y_{ij,FIRST}$, I use the average of consumers in t. For $\tilde{m}^S_{jt+1}$, $\bar{y}_{jt-1}$ is given by the market shares in t, m_{jt}. $(1 - \phi_t)$ presents the patients who purchase a product in the forthcoming period but are new. Here I again use the average parameter values of all parameters and coefficients except for $\bar{y}_{jt-1}$, which is set to zero as the consumers are new.[31]

All in all, firms in t estimate future profits by assuming that the average patient is the same as in t. However, they incorporate the dynamic effects of state dependence and newly entering patients. The assumptions decrease the computations described in the forthcoming section while incorporating the most important demand features. A

[31] In detail, the average coefficients are the same in each period. However, I keep track of the average first choices $\bar{y}_{jFIRST}$ and average number of patients remaining in the market ϕ_t, as well as the average value of consumers having consumed a product in the previous period, $\bar{y}_{it-1}$. When estimating the market share of a product using the averages, I keep track of key aspects determining the demand side. First, new customers do not necessary have each product in their choice set. I follow the assumption of the demand side such that the PoM is always part of the choice set while products that are not in the choice set get weighted by their fraction of choice set considerations in the actual data (i.e., the fraction of choices in the past period).

strategy forms a Markov perfect equilibrium if and only if for all $j \in N^s$ the strategy σ_j^*, $V_j(\sigma_j^*, \sigma_{-j}^*, \mathscr{S}_{jt}, \varepsilon_{jt}) \geq V_j(\sigma_j, \sigma_{-j}^*, \mathscr{S}_{jt}, \varepsilon_{jt})$, for all $\mathscr{S}_{jt}$ and ε_{jt}.

6 Estimation

6.1 Demand

At each point in time, an individual attaches a utility to a product (Equation 2). First, I make an assumption about the choice set of consumers. I assume that the patient's choice set consists of three components: the prescribed product, the PoM (which is always presented by the pharmacist) and previously consumed products (products a patient actively knows). The reduction of the choice set improves the approximation of reality as consumers do not compare all products at the pharmacy level. Note that almost all of the actual purchases in the data are covered by the chosen choice sets. Next, I assume that patients are myopic and do not form expectations of future prices such that switching cost estimates would be dependent on beliefs about future prices.

Given the choice set, a patient compares the products and decides among those with the higher utility. Utility is estimated by the demand-side equation (Equation 2). Given each individual i and time t, the choice set may change. I assume that ε_{ijst} is independent and identically distributed extreme value type 1, such that choice probabilities follow a logit distribution. A patient chooses a product j over k at t if $U_{ijst} \geq U_{ikst}$. I follow standard discrete-choice literature in estimating coefficients of the structural equation (Equation 2). Technical details are provided in online Appendix C.1.

6.2 Supply

On the supply side I use a two-step estimator. In the initial step I estimate the policy function that characterizes the pricing of firms. In the second step I use forward simulations and the assumption that the firms play a Markov perfect equilibrium to estimate marginal costs. The approach is based on methods proposed by Hotz and Miller (1993) and Bajari et al. (2007). The main idea is to initially recover the conditional choice probabilities from observed prices within the data. The parameters of the policy function are the state variables. For firm j, the state variables $\mathscr{S}_{jt}$ are the lagged market shares m_{jt-1}, the lagged number of firms ($|N|_{t-1}$), the market size (Q_t) and a dummy that indicates whether j was the PoM in the previous month (PoM_{jt-1}). Conditional on the state variables, firms set their prices in t. In practice, I estimate the pricing policy σ^* in a reduced-form least squares regression:

$$p_{jt} = \alpha + \beta m_{jt-1} + \eta |N_{t-1}| + \rho Q_t + \gamma PoM_{jt-1} + \varepsilon_{jt}. \tag{4}$$

Note that the number of competitors in the previous period is treated as a factor variable in order to increase flexibility. Therefore, η corresponds to a vector of coefficients.

The second stage of the estimator uses the optimal policy function (Equation 4), which is assumed to be generated by the play of a Markov perfect equilibrium to estimate unobservables that rationalize the optimal policy. I estimate the marginal costs for each competitor within each period. I can recover the marginal costs because of several key assumptions. First, I assume that firms play a single Markov perfect equilibrium. Further, I assume that the profit function is correctly specified and known up to the marginal costs. I assume that firms discount future monthly profit with $\delta = .995$. Transition probabilities from the firms' points of view are estimated based on the assumptions presented in Section 5.2. The market share of the forthcoming period is a function of the demand characteristics within a period t (Equation 3). The value ϕ_t, estimated from the data, describes the share of customers that stay in a market at period t, that is, the share of patients who are the same in $t+1$ as in t. Within the forward simulation at a given period t, the share ϕ_t is constant. All in all, the transition probability estimates incorporate the dynamic factors of prices on market shares in the future but decrease the complexity of the demand system as firms take period-specific demand as an approximation of the future. Finally, I assume that the distribution of the private shocks ε_{jt} is known and given by $N(0,1)$. The assumptions are in line with those of Bajari et al. (2007).

With the assumption that the optimal pricing strategy σ^* is a Markov perfect equilibrium, it has to hold that the expected valuation given a state vector ($\mathscr{S}_j^t$) is higher than any other pricing strategy σ. The time superscripts are dropped as the equation binds in each period.

$$V_j(\sigma_j^*, \sigma_{-j}, \mathscr{S}_\mathbf{j}, c_j) \geq V_j(\sigma_j, \sigma_{-j}, \mathscr{S}_\mathbf{j}, c_j)$$

The valuation function at each period is dependent on the marginal costs. It is therefore possible to use the theoretical assumption of this inequality of a Markov perfect equilibrium when estimating marginal costs. Before making use of the objective function, I simulate the continuation function by forward simulation. I start with 50 initial parameters of the state vector $\mathscr{S}_\mathbf{j}^\mathbf{0}$. For each initial state vector I forward simulate the valuation function over 100 periods using the optimal pricing policy σ^*, using the transition probabilities. Furthermore, in each period a private ε_j shock is drawn. Correspondingly, the simulation of the valuation function given a marginal cost factor $\hat{c}_j$ is given by $\hat{V}_j(\mathscr{S}_j, \sigma_j, \hat{c}_j)$.

I use 200 alternative policy functions σ that are different to σ^*. Here also I simulate valuation functions by forward simulation. I denote one of the 50 initial draws of the state vector with $\mathscr{S}_\mathbf{j}^\mathbf{R}$ ($R = 50$) and the 200 non-optimal policy functions with σ^k (K=200). Given a marginal cost parameter c^{tj}, the difference between the optimal and non-optimal valuation function is described by $g(\mathscr{S}_\mathbf{j}^\mathbf{R}, \sigma_j^k, \hat{c}_j) = \hat{V}_j(\mathscr{S}_j^R, \sigma_j^k, \hat{c}_j) - \hat{V}_j(\mathscr{S}_j^r, \sigma_j^*, \hat{c}_j)$.

Given that the optimal strategy represents the equilibrium, I can construct the objective function in order to esti-

mate the marginal costs. I search for the marginal costs that minimize the following function:

$$\min_{\hat{c}_j} Q(\hat{c}_j) = \frac{1}{K}\frac{1}{R}\sum_{k=1}^{K}\sum_{r=1}^{R} \mathscr{I}\{g(\mathscr{S}_{\mathbf{j}}^{\mathbf{R}}, \sigma_j^k, \hat{c}_j) > 0\} g(\mathscr{S}_{\mathbf{j}}^{\mathbf{R}}, \sigma_j^k, \hat{c}_j)^2.$$

The indicator function $\mathscr{I}$ takes the value 1 if $g(\cdot) > 0$. Thus, I minimize the squared difference between the estimated valuation functions for those cases when the alternative policy function is greater than the valuation function for the optimal policy function that represents the equilibrium.

Overall, I estimate marginal costs for every period and every company. The path is calculated given the demand within the period. The approach allows for different marginal costs within time as well as across companies. Further, the companies have knowledge about the general key factors of demand. Nevertheless, I do not require knowledge of the entire demand system, that is, individual patients and random coefficients of future patients.

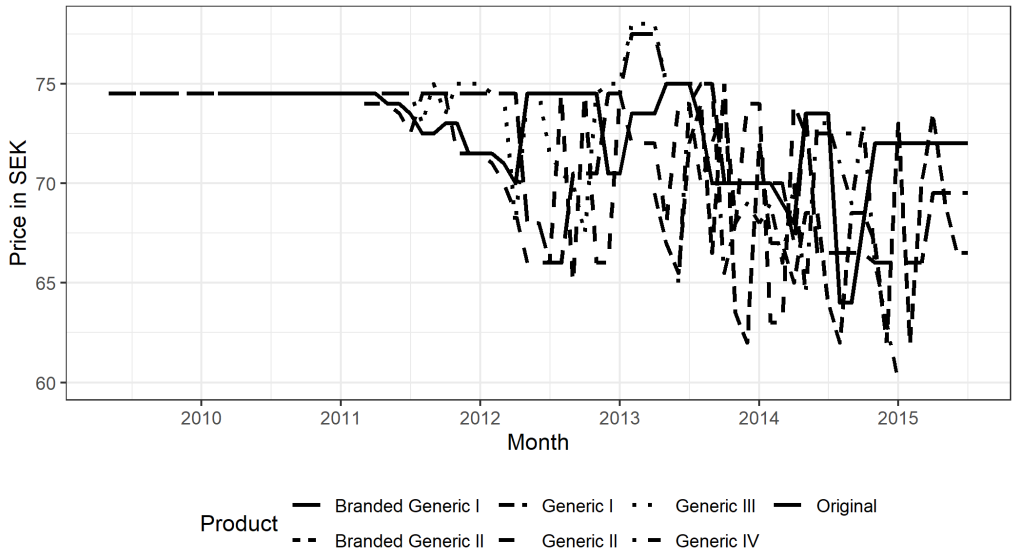
7 Results

In the following I describe the results for the demand and supply sides separately. I start by analyzing the estimation of the model for paracetamol tablets in a high dosage of 1 gram, with a package size of 30 tablets. Paracetamol is a common drug; however, it requires a prescription in high dosage. The substitution group of high-dosage paracetamol tablets is one of the largest in the substitution group system. On average more than 50,297 unique patients purchase paracetamol in a month.

Because paracetamol is no longer covered by a patent, the market for it has at least two competitors for the entire time period. Prices in the substitution group show some volatility, as shown in Figure III. On average, a paracetamol package in the time period has a price of 72.13 SEK (approx. 7.2 USD). Over the study period I observe on average 3.72 competitors. Visual examination suggests that competition has lowered prices over the study period.[32] The figure shows that the cheapest products one month often drastically increase in price in the subsequent period, compared with the overall variation in prices in the substitution group. Theoretically, the price patterns could be rationalized as the actions of forward-looking firms that lock in customers with low prices and then increase their prices and 'harvest' consumers who do not switch despite the higher costs.

[32] Note that there were only two competitors before 2012. Both firms do undercut the price ceiling, which was 74.25 SEK (approx. 7.4 USD) for most of the study period. For two months, the price ceiling was increased to 78 SEK (approx. 7.8 USD). I do not model the increase for those two months separately.

Figure III: Monthly Prices in the Paracetamol Substitution Group



Note: Monthly prices in the substitution group for paracetamol, 1 g, 30 tablets, between 2010 and 2016. Paracetamol is off patent during the entire period. There is a price ceiling of 74.25 SEK (approx. 7.4 USD) that increased for two months to 78 SEK (approx. 7.8 USD).

7.1 Demand

I now turn to the presentation of the demand side. I estimate the demand side as well as the counterfactual on a randomized sample covering one-sixth of the population as well as on the sample of all individuals with a medical education. The reduction in sample size is needed to ensure a sufficiently flexible demand estimation.

Table III shows the results of the control function as well as the random coefficient model. Note that the random coefficients are assumed to follow a normal distribution. First, the linear regression of the first stage shows that prices of other painkillers from the same firm have a strong impact on prices. Next, I show results of the demand estimation.

Models 1 and 2 evaluate the demand of the random sample, whereas Model 1-Med and Model 2-Med consider only those patients with a medical education. First, the upper part of Table III presents the brand-specific intercepts. The middle part of Table III shows the mean and the standard deviation of the random coefficients for prices and the previous consumption (corresponding to α_i and ρ_i in the estimation in Equation 2, respectively). In the lower part, I describe the specification of the different models. Model 1 and Model 1-Med use the random coefficients of previous choice and prices, as well as non-random brand intercepts. Model 2 and Model 2-Med further control for unobserved heterogeneity by controlling for the first observation within the sample. I add the control function approach to control for potential price endogeneity in Model 2 and Model 2-Med.[33]

The main results of Table III are the following: First, the mean of the price coefficient for each specification is

[33]Note that a small brand does not have other products within the group of painkillers during the entire panel period, such that I run out of instruments for part of the sample. Accordingly, the sample gets smaller.

negative and significant. Note that individuals with a medical education are more price sensitive than the general sample. Second, consumers who purchased a specific drug in the preceding period are significantly more likely to purchase a product in the current period. This effect is significant in all specifications, for the full sample as well as the sample with a medical education. Note that the willingness-to-pay estimates are very reasonable in size. On average, patients are willing to pay approximately 0.175 USD more for a product when they have consumed it before (between 2% and 3% of the average price).[34] For the difference between the entire sample and the sample with a medical education, it is important to investigate the differences in terms of the brand intercepts, which translate to perceived quality differences. The default value is the only original product in the market. The intercepts show that branded generics have a generally higher demand within all model specifications. On average, patients are willing to pay a 9.9% higher price for the most preferred branded generic compared with a non-branded generic.[35] These brand premia are similar to those of other consumer goods.[36] However, brand premia higher for the entire sample than for the sample with medical education, who are only willing to pay 2.8% more for the branded generic. The results are in line with the general result that brand premia, as a result of perceived quality differences, are observable. Patients with a medical education are less prone to perceived quality differences.

7.2 Supply

I now present supply-side results that include the estimation of the policy function and the marginal cost estimation. First, Table IVa presents the reduced-form estimate of the policy function as described by Equation 4. It is the first step of the two-step estimation. The outcome variable of the least squares regression is firm j's price in period t (p_{jt}). I explore three different models. Model 1 solely includes the previous market shares (m_{jt-1}) and a constant. Model 2 further includes the product of the month (PoM) dummy of the previous month (PoM_{t-1}), and Model 3 also considers the dummies for the number of competitors in the previous period ($|N_{t-1}|$) as well as the market size (Q_t). The policy function is an estimate of the equilibrium strategy of firms. Each firm plays a symmetric Markov perfect equilibrium. Therefore, the results in Table IVa do not carry any causal interpretation. However, the three models show that the previous market share and the previous PoM are correlated with a higher price in the next period. Furthermore, a higher number of competitors is correlated with lower prices, while the market size (higher demand) is correlated with higher prices. Given the policy function estimate in Model 3 of Table IVa and the transition function in Equation 3, I

[34] Formally, the willingness-to-pay for a consumer i is defined as the state dependence coefficient (ρ_i) divided by the negative of the price coefficient (α_i). Here, I take Model 2, $WTP = 0.313/0.179 = 1.75$ SEK, which corresponds to 0.175 USD.

[35] I calculate the value by taking the coefficient of *Branded Generic I* minus the average coefficient of *Generic I* to *Generic IV*. I calculate the willingness to pay for this difference.

[36] For example, Goldfarb et al. (2009) show brand premia in for ready-to-eat cereal with 19% or lower; Wiggins and Raboy (1996) report brand premia up to 5% for bananas; and Ailawadi et al. (2003) find brand premia lower than approximately 20% for different consumer packaged goods.

Table III: Regression Evidence, Demand Model

	First Stage	Model 1	Model 2	First Stage Med.	Model 1-Med	Model 2-Med
Price of Other Painkillers	0.103***			0.102***		
	(0.002)			(0.008)		
Branded Generic I		0.422***	0.687***		0.112***	0.362***
		(0.005)	(0.008)		(0.025)	(0.033)
Branded Generic II		-0.523***	-1.501***		-1.049***	-1.725***
		(0.013)	(0.028)		(0.075)	(0.145)
Generic I		-1.159***	-0.556***		-0.98***	-0.436***
		(0.01)	(0.014)		(0.049)	(0.063)
Generic II		-1.205***	-0.812***		-1.042***	-0.67***
		(0.008)	(0.01)		(0.037)	(0.043)
Generic III		-0.894***	-0.499***		-0.866***	-0.497***
		(0.014)	(0.016)		(0.071)	(0.077)
Generic IV		-1.54***	-1.033***		-1.082***	-0.766***
		(0.016)	(0.018)		(0.07)	(0.107)
Random Brand Intercepts		No	Yes		No	Yes
Price Mean		-0.16***	-0.179***		-0.444***	-0.411***
		(0.002)	(0.002)		(0.016)	(0.016)
σ		0.384***	0.255***		0.501***	0.427***
		(0.003)	(0.003)		(0.02)	(0.02)
State Dependence Mean		0.536***	0.313***		0.325***	0.176***
		(0.009)	(0.009)		(0.059)	(0.064)
σ		0.338***	0.124***		-0.294**	-0.251
		(0.019)	(0.037)		(0.149)	(0.248)
Control Function			-0.003***			-0.012***
			(0.001)			(0.003)
Including Control Function		No	Yes		No	Yes
Unobserved Heterogeneity		No	Yes		No	Yes
Log-Likelihood		-308,422	-269,276		-13,027.23	-11,490.62
Year-Month FE	Yes			Yes		
Product FE	Yes			Yes		
F-statistics Instruments	3,601.74			150.66		
N	1,260,594	1,108,765	1,006,481	51,893	44,631	40,134
R^2	0.06			0.057		

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results from the mixed logit estimation and the control function. One observation is a patient choice in the substitution group of paracetamol, 1 gram, 30 tablets. The first three columns consider a random sample (one-sixth of the full sample, random selection), whereas the remaining columns solely consider patients with a medical education. The first stage shows results of the control functions, while the models show results of the demand model where the outcome variable is a dummy that indicates if an individual has chosen a product. The upper part of the table shows product-specific intercepts, dependent on branded generics and generics; the default value is the original. Note that the coefficients are partly random and estimates of standard deviations are excluded. Random coefficients are assumed to follow a normal distribution. The lower part of the table shows the random coefficients for price and the state dependence. Note that I also report the standard deviation of the random coefficients. Control Function indicates if the control function approach for endogenous prices has been used. Unobserved Heterogeneity indicates if the model controls for problems due to unobserved heterogeneity. The F-statistic shows the result of a weak instrument F-test. Standard errors are in parentheses.

estimate marginal costs.[37]

Table IV: Supply Side Results

(a) Policy Estimation

	Price		
	(1)	(2)	(3)
Share$(t-1)$	2.598***	2.568***	1.589*
	(0.925)	(0.871)	(0.889)
$I(NoComp.(t-1)=3$			0.062
			(0.851)
$I(NoComp.(t-1)=4$			−0.286
			(0.856)
$I(NoComp.(t-1)=5$			−1.091
			(0.942)
$PoM(t-1)$		2.482***	1.752***
		(0.416)	(0.469)
$Quantity$			0.0001***
			(0.00003)
Constant	71.066***	70.080***	61.178***
	(0.331)	(0.353)	(2.313)
N	272	272	272
R^2	0.028	0.142	0.262

* $p<0.1$, ** $p<0.05$, *** $p<0.01$

Notes: Linear least squares regression results for the estimation of the policy function. One observation corresponds to the monthly price of a product in the substitution group of paracetamol, 1 gram, 30 tablets. The outcome variable is the price of a product in period t. All regressors are state variables of the supply side: Share$(t-1)$ is the market share in the preceding period. $I(NoComp.(t-1)=3)$, $I(NoComp.(t-1)=4)$, and $I(NoComp.(t-1)=5)$ are dummies that take the value 1 if in the preceding period the number of firms was equal to 3, 4 or 5. $PoM(t-1)$ is a dummy that takes the value 1 if the firm was the cheapest product in the previous month. Standard errors are in parentheses.

(b) Marginal Cost Estimation

Firm	Mean Marginal Costs
Original	18.32
	(4.30)
Branded Generic I	16.69
	(1.87)
Branded Generic II	23.89
	(1.02)
Generic I	19.57
	(1.09)
Generic II	20.52
	(1.19)
Generic III	21.51
	(0.78)
Generic IV	19.27
	(0.68)

Notes: Summary of marginal cost estimates (in SEK; 10 SEK equal approximately 1 USD) for different firms in the market of paracetamol, 1 gram, 30 tablets. Note that the marginal costs are the means of the marginal costs estimated for each period in which a product was present. I present standard deviations of marginal cross across monthly pricing periods in parentheses.

Given the two initial steps, Table IVb shows the estimates from the marginal costs for each firm across time on average. The marginal cost estimates are heterogeneous across firms and vary over time.[38] I present the point estimates as well as standard errors for the marginal costs of each brand within each period. I obtain the standard error by bootstrapping. Estimates of marginal costs in each period for each firm are statistically different from zero. In Table IVb I show the average marginal costs for each firm over the periods in which the firm was present in the market. I also present the standard deviations for the estimates across time. The marginal costs vary only slightly over time. One observes some differences in the marginal costs across products; however, they are not large. Especially the first branded generic, which also has the highest brand premia in the demand model, has lower market share. As it is part of a well-known generic brand that has a high market share in a lot of submarkets, it is possible that

[37] Note that the share of patients who stay in the market between t and $t+1$ (ϕ_t) is inferred from the descriptive statistics in the data. On average across the time periods, 22% (standard deviation, 0.032) of the patients stay in the market.

[38] Details of the estimates are presented in online Appendix D.2.

a large product base decreases marginal costs due to distribution. The estimates of the marginal costs are lower for the branded generics than for generics with lower market shares because of economies of scale. In addition to the variability of marginal costs, one may assess the suitability of the marginal cost estimates by comparing variability of prices, marginal costs and markups. Biased estimates of marginal costs may be characterized by capturing all the variability or price changes such that markups do not change over time. Dynamic pricing would therefore be explained by marginal costs. However, I do not observe such a correlation as products with large market shares (such as the first branded generic) tend to have higher prices but lower marginal costs.

8 Counterfactuals

In the counterfactual analysis, I first present the implementation procedure and then show that an extension of the contract length that mimics a reduction of switching costs increases prices in equilibrium. In the second counterfactual scenario, all patients act as if they had a medical education, and prices decrease.

8.1 Implementation

Technically the implementation of counterfactuals requires me to estimate the policy that represents a Markov perfect equilibrium. Consistent with the Markov perfect strategies in the previous section, firms condition their strategy on all state variables. Because the environment or demand parameters change within the counterfactual scenario, I cannot use the policy function presented in Section 7.2. Instead, I need to compute Markov perfect equilibria by value (or policy) iterations and solve for a pricing equilibrium during each iteration. Because the computational burden increases exponentially with the number of state variables, I simplify the environment as follows.

I reduce the state space $\mathbf{S}_t$ to only one variable, the PoM status. Thus, firms condition their prices on the knowledge of whether their product was the cheapest in the preceding month. The reduction of the state space is strong. The reduction is motivated by the need for a simplified environment. Further, I show in online Appendix E.1 a model selection method (LASSO) based on the policy function.[39] I show technical details of the algorithm and its implementation in online Appendix E.2. In general, the algorithm works as follows: I perform a value iteration that incorporates the equilibrium conditions due to the assumption that firms play a Markov perfect equilibrium. The following steps are done in each period: As before, I take advantage of the demand parameters. I take the average consumer and assume that firms know how many patients stay in the market. Firms assume that the average consumer is constant. I further use the marginal cost estimates from the previous sections. The marginal cost estimates are different across

[39] Performing a model selection, I show that PoM status in the previous period and the quantity are the most predictive regressors. I therefore choose the previous PoM as the new state space. The method of regularization using machine learning methods is related to the solution concept offered by Thiel (2019).

Table V: Benchmark Model, Products

	Mean Data	SD	Mean Benchmark Model	SD
Prices				
Mean Price	71.76	3.63	70.11	2.19
Mean Price Original/Mean Price Avg. Generic	1.01		1.01	
Market Shares				
Branded Generic I	0.25	0.1	0.27	0.07
Branded Generic II	0.44	0.24	0.42	0.22
Generic I	0.06	0.06	0.05	0.03
Generic II	0.04	0.04	0.02	0.02
Generic III	0.01	0.03	0.01	0.01
Generic IV	0.03	0.04	0.02	0.02
Original	0.19	0.12	0.19	0.11

Notes: The table compares the data and the prediction from the benchmark model. The upper part of the table shows comparison of prices. Prices are reported in SEK (10 SEK are approx. 1 USD). The second and third columns show the mean prices and standard deviations from observable prices in the substitution group of paracetamol, 1 gram, 30 tablets. The fourth and fifth columns show the prediction from the model with one state variable. The mean prices correspond to the mean across all periods and all available products. The lower part of the table shows statistics for market shares conditional on being present in a market. Note that not all products are available in each month.

firms. For each possible state (each firm's product could have been the PoM in the previous month), I start with an initial guess of the value function for each firm (V^0). In each iteration k for each state I search for an equilibrium in prices. Note that the search for the mutual best reply (equilibrium) incorporates the static prices, the transition to the state of the next period ($\pi(\mathbf{S}^{t+1}|\mathbf{S}^t, p_j, p_{-j})$) and the value function that is dependent on the state ($V^k(\mathbf{S}^t)$). Given continuation values I update the best replies for each player in each state. I update the Bellmann equation and get a new estimate for the value function V^k. During each iteration I update the value function until convergence.

Note that the simplification and estimation come with two major concerns. The first concern is a computational one. The grid of used prices may lead to different equilibria and therefore prices. Further, equilibria may not be unique, and the employed algorithms may lead to a different equilibrium than the ones chosen by firms. The second concern is the simplification of the state space. To tackle both concerns, I explore the results of the simplified model and compare it with the observable prices before turning to the counterfactuals. Within the benchmark model I do not change the environment, and I use the demand estimates as well as the marginal cost estimates.

Table V shows some basic statistical measures of the prices and market shares in the data as well as the benchmark model. The simulated prices in the benchmark model are slightly lower than the ones in the data. In the lower part of Table V, I show a comparison of the market shares in data and in the simulated benchmark model. The relation between the different brands in the benchmark model fits the data reasonably well. Overall, the benchmark model's key characteristics are close to those of the actual data.

8.2 Extension of the Procurement Contract Length

In the first counterfactual I do not change the demand side but change the institutional background. Firms in this counterfactual are allowed to change their prices only once each year. Note that consumers' switching costs are still relevant, as some consumers may start a treatment within one year and continue into the next. However, behavioral pricing that intends to lock in consumers and harvest them in the forthcoming period is more expensive for firms. Firms would need to have lower prices over a longer term and therefore forgo profits. Overall, this counterfactual is motivated by two aspects. First, it incorporates a realistic policy change, as it solely changes the timing of the current pricing policy. Second, the counterfactual directly reduces switching costs as (1) the number of consumers whose use of the product continues over several years gets smaller and (2) consumers tend to oppose substitution less often when the time since the last purchase of a more expensive product increases.[40]

Technically, I use the same demand model as presented before. The frequency of price changes is reduced from 72 to 6. Within each new period, consumers are treated equally. The state variable still shows whether a firm had the cheapest product in the preceding period. The effect on the next period's demand, however, is reduced. The market share is equivalent to the one in the benchmark model. On the supply side, I assume that firms that are present in at least two months of the year are present for the entire year. This assumption increases competition over the studied time periods. However, the possible policy change would come with increased competition, because firms would be able to enter and exit only once a year and it is reasonable that firms would stay in a market longer. As the length of the periods changes, I adjust the share of patients that stay in a market over two subsequent months (ϕ) as well as the discount rate ($\delta^{NEW} = .95$). Finally, the estimation of the prices in equilibrium is equivalent to that in the benchmark model, as I change only the sample of a single period and reduce the number of periods.[41]

Table VI compares market outcomes of the different counterfactuals and the benchmark model. The average price is higher in the scenario with a 'product of the year' rather than a 'product of the month' (70.59 instead of 70.11 SEK). In Panel B of Table VI, I show results that incorporate the behavior of consumers and corresponding market shares of products. The results of the price for the average consumer are comparable to the general price statistics. In detail, the average purchase price for a consumer is slightly higher in the scenario of a different procurement process compared with the benchmark model. Further, the market cap for firms increases in the counterfactual scenario.

The counterfactual leads to the conclusion that lowering the possibility of reacting to switching costs by reducing the frequency of price changes is, on average and overall, not welfare-enhancing for the consumer. This result is in line with the research of Arie and Grieco (2014), Cabral (2016), Dubé et al. (2009), Fabra and García (2015) and

[40] In online Appendix E.3 I estimate a counterfactual in which I directly reduce switching costs on the consumer side. I show that the results are analogous.

[41] Note that I do not model entries and exits. I assume existence of a product within a year if the product is available in at least two months of the respective year. Therefore I implicitly increase competition, which could in principle reduce prices in equilibrium.

Rhodes (2014). These authors show theoretically as well as empirically that moderate switching costs may increase competitive pressure. Even though switching costs increase the market power of a firm with locked-in patients and therefore induce 'lock in and harvest' behavior, prices on average are lower. Because firms have an incentive to decrease prices in order to lock in patients, and moderate switching costs (see demand-side estimates) prevent prices from becoming too high in the harvest phase, prices may be lower when switching costs are present. Also, the lower standard deviation of prices in the counterfactual with a different procurement contract length confirms the lower variability of prices.

However, several factors need to be considered when evaluating the overall effect of switching costs. First, note that higher prices are not uniform for all consumers. Indeed, consumers with high switching costs and frequent purchases profit from the new procurement process as there is no 'harvesting' by firms that charge high prices, whereas new consumers or those with no switching costs suffer from higher prices because there is no 'lock in' period with low prices. Figure IV describes the observation.[42] First, Figure IVa shows the distribution price differences between the counterfactual of longer procurement length and the benchmark (a positive difference means that a consumer pays more when switching costs play a less important role). While most patients pay slightly more after the increase of procurement length, some pay less. Figures IVb, IVc and IVd present some characteristics of consumers that pay more or less in the counterfactual. First, observe that patients who consume paracetamol more frequently are more likely to have lower costs in the counterfactual as repeated purchases are necessary to suffer from dynamic pricing. Second, lower income and lower education correlate with higher costs for patients in the counterfactual that mimics lower switching costs.

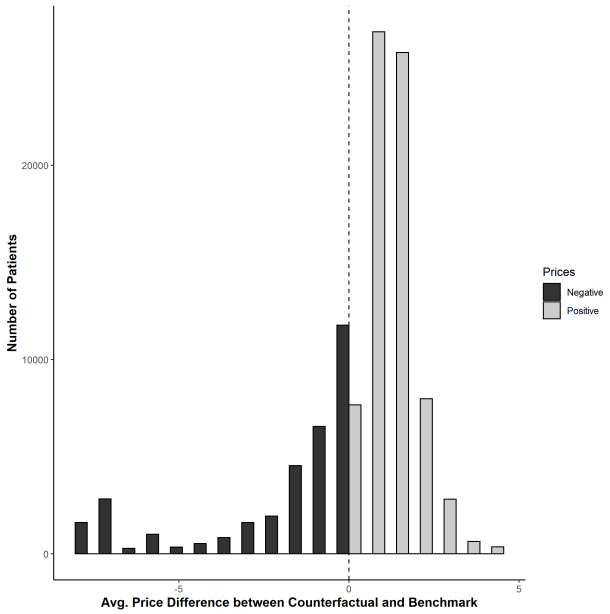
In a market for switching costs, average prices may lead to the idea that policies reducing the impact of switching costs are non-optimal as they increase prices. However, the analysis shows that switching costs are pro-competitive on average and increase prices for a substantial subset of patients. In a market where firms engage in dynamic pricing, patients with high switching costs get exploited. A policymaker may have the objective to reduce such inequality, especially as patients with lower education, greater need for pharmaceuticals, and lower income are more likely to have switching costs.

8.3 Brand Premia

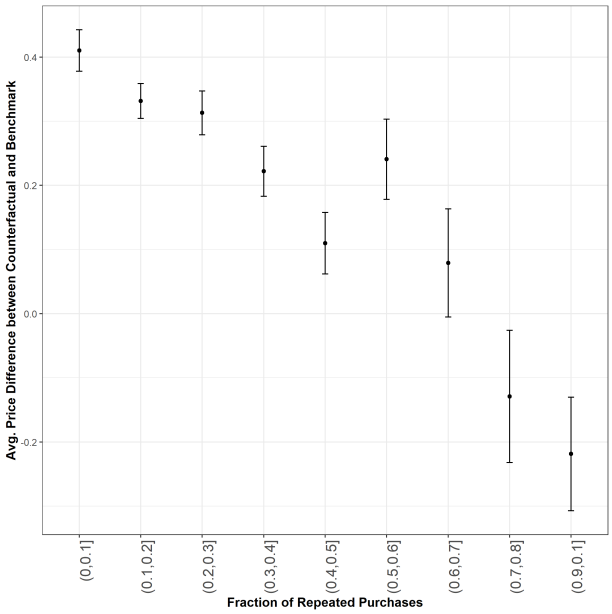
In the second counterfactual, I show the impact of brand premia due to perceived quality differences on pricing and consumers' welfare. In practice, I use the demand estimates from the sample of patients with medical education and estimate the supply side, holding the original quantity and cost factors fixed. I use the state dependence coefficients from the sample of experts but do not change the sample of patients that stay in a market over two subsequent months.

[42] In detail, I use the new prices of the counterfactual and coefficients of the demand model to calculate each conditional choice probability.

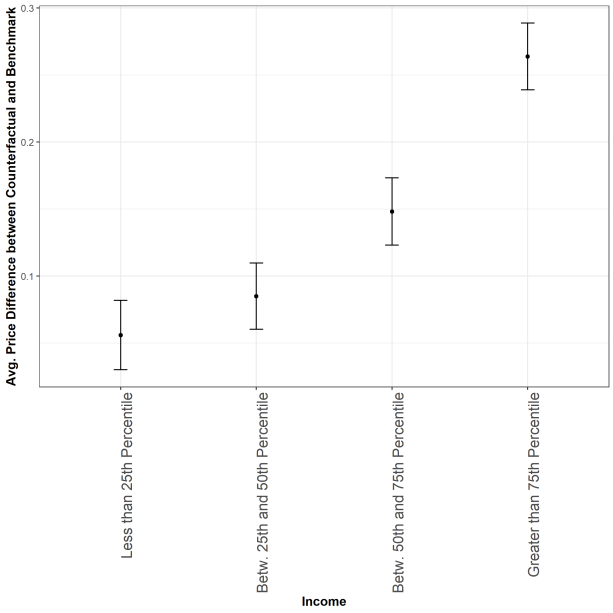
Figure IV: Profit and Loss from Procurement Counterfactual



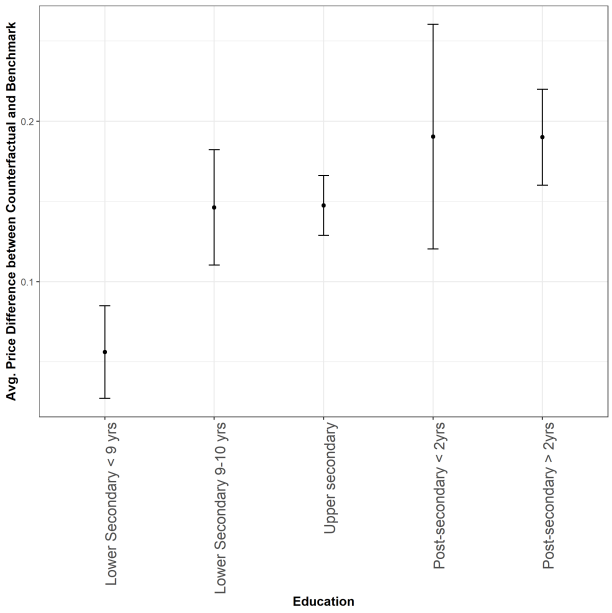
(a) Distribution of Patients



(b) Impact of Purchase Frequencies



(c) Impact of Income



(d) Impact of Education

Notes: The four figures evaluate the heterogeneity of patients' payment in the procurement counterfactual compared with the benchmark scenario. Each observation considers a patient's individual purchase occasion given the prices in the counterfactual and the benchmark. Choice probabilities are estimated using results of the full demand model presented in Equation 2. Figure IVa shows the distribution of purchase distributions according to the difference between the counterfactual and benchmark scenario. Figure IVb evaluates the mean price differences across the fraction of repeated purchases across consumers. Figure IVc presents the mean price differences across purchase occasions across income percentiles. Finally, Figure IVd shows the price differences for different education levels. The error bars correspond to a 95% confidence interval.

Table VI: Results of Counterfactuals

	Benchmark Model	Counterfactual Procurement	Counterfactual Brand Premia
Panel A: Average Prices			
Mean price	70.11	70.59	67.49
	(2.19)	(1.84)	(2.45)
Branded Generic I	69.66	70.42	66.09
Branded Generic II	74.35	74.53	72.01
Generic I	68.41	69.39	65.84
Generic II	69.5	70.59	67.88
Generic III	71.42	71.86	69.12
Generic IV	67.67	69	66.37
Original	70.78	70.25	67.81
Panel B: Average Expenditures and Revenue			
Avg. Price for Consumer	70.49	70.77	67.63
Compared to Benchmark		0.4%	-4.05%
Avg. Revenue per Month (in million SEK)	4.6	4.74	4.41
Compared to Benchmark		3.15%	-4.03%

Notes: The table shows a comparison between the benchmark model and the two counterfactual scenarios. The first counterfactual is the different procurement process. The second counterfactual mimics the case of a decrease in brand premia when all patients act as though they have a medical education. Panel A reports prices in SEK (10 SEK are approx. 1 USD). Panel B shows measures of consumer costs as well as total revenues. Standard deviations are in parentheses.

Table VI shows descriptive statistics of the pricing equilibrium in the counterfactual with a decrease in brand premia. On average, prices in the counterfactual, with a decrease in perceived quality differences, are lower than in the benchmark model (67.5 SEK vs. 70.1 SEK). Thus, the results go in the opposite direction than in the counterfactual with a different procurement process. The reasons are that experts have lower brand preferences and have a higher price elasticity. The new demand leads to stronger competition between existing firms and therefore lower prices.

Panel B of Table VI shows the effects for the average consumer. With the decrease in perceived quality differences, the consumer pays on average 67.6 SEK for a product. Two effects play a role. First, prices are lower, as firms have less incentive to engage in behavioral pricing and to take advantage of brand preferences due to perceived quality differences. Second, consumers themselves are less willing to pay brand premia, as perceived quality differences have decreased. Thus, they tend to consume cheaper products. Overall, the average consumer spends 4.05% less for a product in the counterfactual scenario.

9 Discussion

This article provides causal evidence for switching costs and perceived quality differences in the Swedish pharmaceutical market. I estimate the impact of both behavioral phenomena on pricing through counterfactuals in the substitution group of paracetamol, 1 gram. In the following, I provide a back-of-the-envelope calculation that quantifies aggregate

effects of the counterfactuals across substitution groups on patients and public health care expenditure. Afterward, I discuss policy implications.

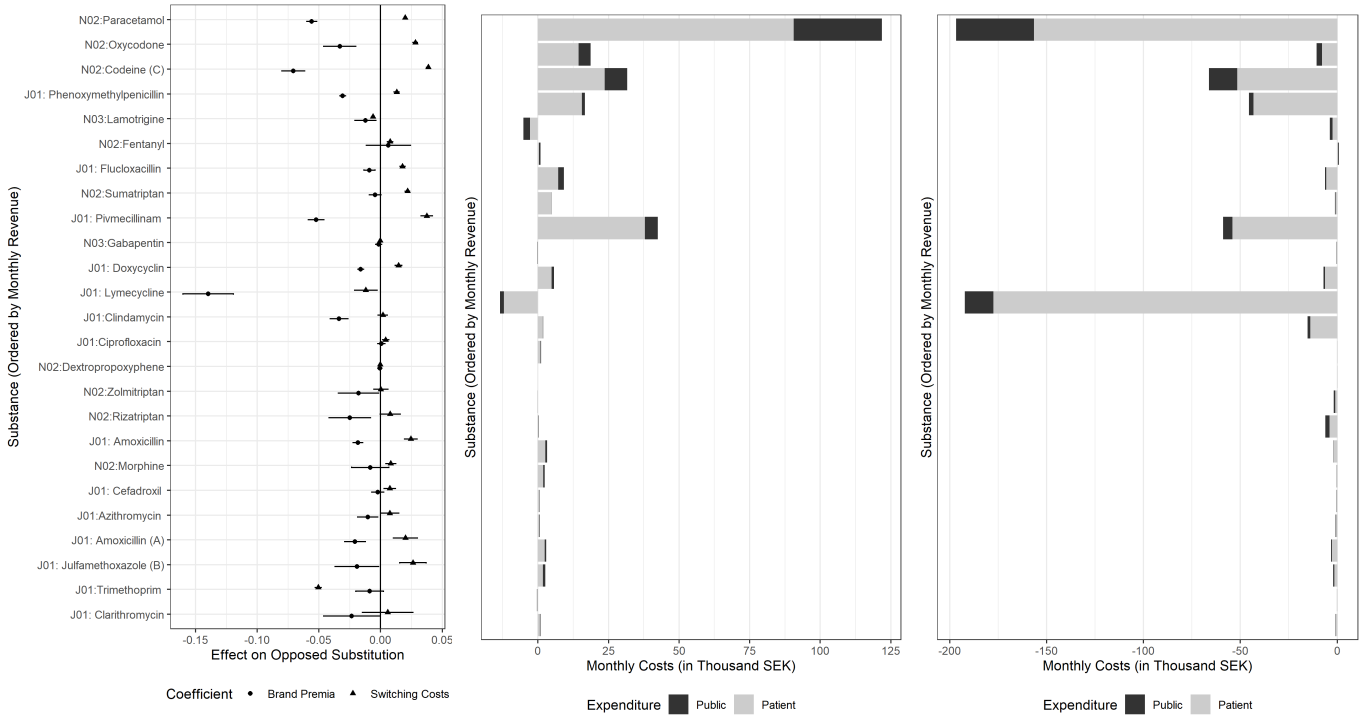
9.1 Aggregate Effects

The effects of the increased length in procurement and the reduction in brand premia are based on the careful estimation that involves the estimation of demand, costs of firms, and their maximization procedure that involves competition. Since it is not feasible to estimate the counterfactuals for each substitution group individually, I estimate the effects of the counterfactual using reduced-form estimates and connect them to the structural model of paracetamol. The procedure is the following: (1) For each substance group I observe (different substances of painkillers, antibiotics, and antiepileptics), I estimate a model equivalent to Equation 1 including year-month fixed effects. Estimates of β_1 give a raw estimate of the importance of switching costs, while β_2 shows the impact of brand premia due to perceived quality misconceptions. (2) I use the reduced-form and structural model estimates of the paracetamol substitution group to create a linear mapping between reduced-form estimates of behavioral frictions and the counterfactual impact on consumers and on public health care costs. (3) Given the mapping, I calculate the counterfactual impact for each substitution group separately. Overall, the calculation allows me to estimate the aggregate effects of both counterfactuals in a simplified environment. Nevertheless, the simplified environment comes with strong assumptions. For example, it does not consider differences in market structure.

Figure V shows the results of the back-of-the-envelope calculation. Figure Va shows the reduced-form estimates for each substance. We see some heterogeneity in results. However, two observations are generally observable in most substitution groups: (1) Consumption in the last month increases the probability of a patient opposing substitution, which would mean positive switching costs. We see an especially high impact on painkillers such as opioids. (2) Having a medical education decreases the probability that a patient opposes substitution. Thus, we observe brand premia due to perceived quality misconceptions. Figure Vb shows results of the procurement counterfactual on monthly expenditures in SEK. Paracetamol has the highest effect on overall costs, meaning that costs for the public and patients increase. We see that the counterfactual especially has effects in the segments of painkillers and antibiotics, products with a large quantity and switching costs. Patients bear the majority of the additional costs. Figure Vc shows results of the counterfactual of reduced brand premia on monthly expenditures in SEK. The counterfactual effects are large in the substitution groups of painkillers and antibiotics, where brand premia due to misconceptions are common. In comparison to switching costs, brand premia have a larger effect. The effects on patients' co-payments are large. Still, we observe savings for the public system as price differences between products decrease and the price of the cheapest available product decreases.

Across all products, the procurement counterfactual increases patients' cost by 0.2% and public costs by 0.06%. If all patients acted as though they had medical education, costs for patients across all groups would decrease by 0.7% and governmental costs would decrease by 0.1%. We observe differences between cost reduction for patients and the public spending in the counterfactual where everyone acts as though they had medical education. The intuition is as follows. If the price of the product of the month decreases, governmental and private costs decrease as the government only reimburses (to some degree) the cost of the lowest priced product. If a branded product decreases its price but the price of the lowest-priced product remains constant, only patients' co-payments decrease. In the case of the reduction in brand premia, we see less expensive brands and more choices of the cheapest product. However, we observe fewer price changes of the cheapest available product.

Figure V: The Impact of State Dependence and Brand Premia across Drug Segments



(a) Estimates of State Dependence and Brand Premia

(b) Procurement Counterfactual

(c) Reduced Brand Premia Counterfactual

Notes: The figures show the impact of an increased procurement length and a reduction of brand premia (in the case where all patients act as though they had medical education) on patients' costs and public expenditure. Figure Va shows coefficients of the reduced-form model estimating brand premia and switching costs. The error bars correspond to a 95% confidence interval. I use a mapping of reduced-form estimates of brand premia and switching costs to estimate the impact of the counterfactuals on public and private expenditures in the segment of paracetamol, 1 gram. Given this mapping I calculate the effects of procurement (Figure Vb) and reduced brand premia due to quality misconception (Figure Vc) in the counterfactual for each individual segment. Both subfigures show monthly costs where positive costs refer to higher costs for patients and the public health care system. Effects on public health care expenditures are shown in black while those on the patients' co-payments are shown in grey.

9.2 Policy

The specific setting of the Swedish health care system shows that information provision and suggestion of substitution on the pharmacy level is not necessarily sufficient to ensure substitution to cheap generics. In numerous health care systems, the substitution decision is less strict as substitution is not suggested or presented as an option to the consumer. This analysis has shown that even in such a regulated market, where inattention is not an issue, we may see firms that build their pricing on behavioral phenomena. To quantify the impact of the brand premia and switching costs, I have presented two counterfactuals to evaluate the impact on consumers' decisions. Both counterfactuals may help policymakers that want to reduce health care costs for the public system or patients.

From an aggregate perspective, reducing perceived quality differences has considerable effects. As medical experts tend to choose cheaper drugs, it would be one option to let physicians rather than patients decide about product choice. Another possibility would be to make substitution mandatory. However, this restricts freedom of choice substantially.

The counterfactual of an increased procurement length has shown that prices could increase if switching costs play a smaller role in a firm's pricing objective. Thus, markets with switching costs could have competitive effects and seem to be desirable. However, this paper documents an important inequality that comes with switching costs. In a market with switching costs, individuals who consume products repeatedly and experience switching costs will overpay as they fall victim to firms' 'lock in and harvest' pricing. In comparison, individuals who do not experience switching costs can take advantage of low-price periods of the 'lock in and harvest' strategy. The problem for policy gets especially clear when observing that patients with lower education and lower income are likely to have worse health conditions and frequently need pharmaceuticals. Thus, switching costs may increase competitive pressure but could increase inequality across buyers. Policymakers should carefully consider whether a portion of the consumers require protection. Increasing the procurement length is a realistic policy option that could help protect consumers who experience switching costs.

References

Abaluck, J. and Gruber, J. (2016), 'Evolving choice inconsistencies in choice of prescription drug insurance', *American Economic Review* **106**(8), 2145–84.

Ackerberg, D. A. (2001), 'Empirically distinguishing informative and prestige effects of advertising', *The RAND Journal of Economics* pp. 316–333.

Ailawadi, K. L., Lehmann, D. R. and Neslin, S. A. (2003), 'Revenue premium as an outcome measure of brand equity', *Journal of Marketing* **67**(4), 1–17.

Arie, G. and Grieco, P. L. (2014), 'Who pays for switching costs?', *Quantitative Marketing and Economics* **12**(4), 379–419.

Ax, F. and Ekedahl, A. (2010), 'Electronically transmitted prescriptions not picked up at pharmacies in Sweden', *Research in Social and Administrative Pharmacy* **6**(1), 70–77.

Bajari, P., Benkard, C. L. and Levin, J. (2007), 'Estimating dynamic models of imperfect competition', *Econometrica* **75**(5), 1331–1370.

Bergman, M., Granlund, D. and Rudholm, N. (2012), 'Apoteksmarknadens omreglering - effekter på följsamhet, priser och kostnader per dygnsdos', *Tillvaextanalys, Workingpaper 19* .

Bhargava, S., Loewenstein, G. and Sydnor, J. (2017), 'Choose to lose: Health plan choices from a menu with dominated option', *The Quarterly Journal of Economics* **132**(3), 1319–1372.

Branthwaite, A. and Cooper, P. (1981), 'Analgesic effects of branding in treatment of headaches.', *British Medical Journal* **282**(6276), 1576–1578.

Bronnenberg, B. J., Dubé, J.-P., Gentzkow, M. and Shapiro, J. M. (2015), 'Do pharmacists buy Bayer? Informed shoppers and the brand premium', *The Quarterly Journal of Economics* **130**(4), 1669–1726.

Cabral, L. (2016), 'Dynamic pricing in customer markets with switching costs', *Review of Economic Dynamics* **20**, 43–62.

Crawford, G. S. and Shum, M. (2005), 'Uncertainty and learning in pharmaceutical demand', *Econometrica* **73**(4), 1137–1173.

Dalton, C. M., Gowrisankaran, G. and Town, R. J. (2020), 'Salience, myopia, and complex dynamic incentives: Evidence from Medicare Part D', *The Review of Economic Studies* **87**(2), 822–869.

Dubé, J.-P., Hitsch, G. J. and Rossi, P. E. (2009), 'Do switching costs make markets less competitive?', *Journal of Marketing Research* **46**(4), 435–445.

Dubé, J.-P., Hitsch, G. J. and Rossi, P. E. (2010), 'State dependence and alternative explanations for consumer inertia', *The RAND Journal of Economics* **41**(3), 417–445.

Einav, L., Finkelstein, A. and Schrimpf, P. (2015), 'The response of drug expenditure to nonlinear contract design: Evidence from Medicare Part D', *The Quarterly Journal of Economics* **130**(2), 841–899.

Ekedahl, A. and Månsson, N. (2004), 'Unclaimed prescriptions after automated prescription transmittals to pharmacies', *Pharmacy World and Science* **26**(1), 26–31.

Erdem, T. and Sun, B. (2001), 'Testing for choice dynamics in panel data', *Journal of Business & Economic Statistics* **19**(2), 142–152.

Ericson, R. and Pakes, A. (1995), 'Markov-perfect industry dynamics: A framework for empirical work', *The Review of Economic Studies* **62**(1), 53–82.

Fabra, N. and García, A. (2015), 'Market structure and the competitive effects of switching costs', *Economics Letters* **126**, 150–155.

Feng, J. (2022), 'History-dependence in drug demand: Identification and implications for entry incentives', *Review of Economics and Statistics, Forthcoming* .

Goldfarb, A., Lu, Q. and Moorthy, S. (2009), 'Measuring brand value in an equilibrium framework', *Marketing Science* **28**(1), 69–86.

Granlund, D. and Rudholm, N. (2018), 'Risker och kostnader för prissamordningar på den svenska generikamarknaden1', *Report* .

Handel, B. R. (2013), 'Adverse selection and inertia in health insurance markets: When nudging hurts', *American Economic Review* **103**(7), 2643–82.

Handel, B. R. and Kolstad, J. T. (2015), 'Health insurance for "humans": Information frictions, plan choice, and consumer welfare', *American Economic Review* **105**(8), 2449–2500.

Handel, B. and Schwartzstein, J. (2018), 'Frictions or mental gaps: What's behind the information we (don't) use and when do we care?', *Journal of Economic Perspectives* **32**(1), 155–78.

Hausman, J. A. (1996), Valuation of new goods under perfect and imperfect competition, *in* 'The economics of new goods', University of Chicago Press, pp. 207–248.

Ho, K., Hogan, J. and Scott Morton, F. (2017), 'The impact of consumer inattention on insurer pricing in the Medicare Part D program', *The RAND Journal of Economics* **48**(4), 877–905.

Hortaçsu, A. and Syverson, C. (2004), 'Product differentiation, search costs, and competition in the mutual fund industry: A case study of S&P 500 index funds', *The Quarterly Journal of Economics* **119**(2), 403–456.

Hotz, V. J. and Miller, R. A. (1993), 'Conditional choice probabilities and the estimation of dynamic models', *The Review of Economic Studies* **60**(3), 497–529.

Janssen, A. (2018), 'Price dynamics of Swedish pharmaceuticals', *Unpublished* .

Kamenica, E., Naclerio, R. and Malani, A. (2013), 'Advertisements impact the physiological efficacy of a branded drug', *Proceedings of the National Academy of Sciences* **110**(32), 12931–12935.

Ketcham, J. D., Lucarelli, C. and Powers, C. A. (2015), 'Paying attention or paying too much in Medicare Part D', *American Economic Review* **105**(1), 204–33.

Klemperer, P. (1995), 'Competition when consumers have switching costs: An overview with applications to industrial organization, macroeconomics, and international trade', *The Review of Economic Studies* **62**(4), 515–539.

Marzilli Ericson, K. M. (2014), 'Consumer inertia and firm pricing in the Medicare Part D prescription drug insurance exchange', *American Economic Journal: Economic Policy* **6**(1), 38–64.

Maskin, E. and Tirole, J. (1988), 'A theory of dynamic oligopoly, II: Price competition, kinked demand curves, and Edgeworth cycles', *Econometrica* pp. 571–599.

Miller, K. (2019), 'Estimating costs when consumers have inertia: Are private Medicare firms more efficient?', *Unpublished* .

Mullainathan, S. (2017), 'Why trying new things is so hard to do', https://www.nytimes.com/2017/12/01/business/why-trying-new-things-is-so-hard.html .

OECD (2009), 'The economic impact of counterfeiting and piracy', *Report* pp. 1–397.

OECD (2017), 'Health at a glance, pharmaceutical sector', *OECD Publishing, Paris* .

Rabe-Hesketh, S. and Skrondal, A. (2013), 'Avoiding biased versions of Wooldridge's simple solution to the initial conditions problem', *Economics Letters* **120**(2), 346–349.

Rhodes, A. (2014), 'Re-examining the effects of switching costs', *Economic Theory* **57**(1), 161–194.

Rickert, D. (2016), 'Switching costs, dynamic pricing, and market power. evidence from a market with a consumer life cycle', *Unpublished* .

Shcherbakov, O. (2016), 'Measuring consumer switching costs in the television industry', *The RAND Journal of Economics* **47**(2), 366–393.

Shy, O. (2002), 'A quick-and-easy method for estimating switching costs', *International Journal of Industrial Organization* **20**(1), 71–87.

Skrondal, A. and Rabe-Hesketh, S. (2014), 'Handling initial conditions and endogenous covariates in dynamic/transition models for binary data with unobserved heterogeneity', *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **63**(2), 211–237.

Sveriges Riksdag (2002), 'Lag (2002:160) om läkemedelsförmåner m.m.'.

Swedish Medical Product Agency (2010), 'Comments on the list of substitutable medicinal products', *https://lakemedelsverket.se/english/product/Medicinal-products/Substitution/* .

Thiel, J. (2019), 'Price discrimination, switching costs and welfare: Evidence from the Dutch mortgage market', *Unpublished* .

TLV (2016*a*), 'Läkemedelsmarknaden', *http://www.tlv.se/lakemedel/Lakemedelsmarknaden/* .

TLV (2016*b*), 'Periodens varor', *http://www.tlv.se/apotek/utbyte-av-lakemedel-pa-apotek/periodens-varor/* .

TLV (2016*c*), 'Pris och subvention av läkemedel', *http://www.tlv.se/lakemedel/pris-och-subvention-av-lakemedel/* .

TLV (2016*d*), 'Takpriser', *http://www.tlv.se/lakemedel/takpriser/* .

TLV (2017), 'Ändring i tandvårds- och läkemedelsförmånsverkets allmänna råd (tlvar 2003:2)', *https://www.tlv.se/download/18.467926b615d084471ac31df8/1510316395145/Tillfallig_subvention_tidigare_licenslakemedel.pdf* .

Viard, V. B. (2007), 'Do switching costs make markets more or less competitive? the case of 800-number portability', *The RAND Journal of Economics* **38**(1), 146–163.

Wiggins, S. N. and Raboy, D. G. (1996), 'Price premia to name brands: an empirical analysis', *The Journal of Industrial Economics* pp. 377–388.

Wooldridge, J. M. (2005), 'Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity', *Journal of Applied Econometrics* **20**(1), 39–54.

Zhang, Y., Donohue, J. M., Newhouse, J. P. and Lave, J. R. (2009), 'The effects of the coverage gap on drug spending: A closer look at Medicare Part D: Beneficiaries who entered the "doughnut hole" decreased their monthly prescriptions by about 14 percent per month.', *Health Affairs* **28**(Suppl1), w317–w325.