Online Appendix

Generic and Branded Pharmaceutical Pricing: Competition under Switching Costs

Aljoscha Janssen*

October 17, 2022

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

Contents

A Additional Institutional Details **3**

A.1 Co-payment Function 3

A.2 Timeline of Auction 3

A.3 Trade-margins of pharmacies 4

B Reduced Form Analysis **5**

B.1 Other Experts 5

B.2 Patent Expiries, Additional Substances 5

C Demand Model **10**

C.1 Technical Details 10

D Supply Model **11**

D.1 Aggregate Transition Probabilities 11

D.2 Marginal Cost Estimates 11

E Counterfactual Analysis **13**

E.1 Reduction of State Space 13

E.2 Technical Details of Counterfactuals 15

E.3 Reduction and Increase of Switching Costs 18

F Back-of-the-envelope Calculation **19**

Table 1: Co-Payment Structure

Price	Reimbursement	Max. Total Out-of-Pocket Payment
$p \geq 4300$	100%	
$3500 \leq p < 4300$	90%	1800 SEK
$1700 \leq p < 3500$	75%	1700 SEK
$900 \leq p < 1700$	50%	1300 SEK
$p < 900$	0	900 SEK

Price	Reimbursement	Max. Total Out-of-Pocket Payment
$p \geq 5400$	100%	
$3900 \leq p < 5400$	90%	2200 SEK
$2100 \leq p < 3900$	75%	2050 SEK
$1100 \leq p < 2100$	50%	1600 SEK
$p < 1100$	0	1100 SEK

Notes: Co-payment structure for cumulative health care expenditure (including prescription drugs) before (upper panel) and after (lower panel) 2012. Reimbursement is calculated for expenses during an entire year, beginning with the first expenditure. Prices are in SEK. 10 SEK are approximately 1 USD.

A Additional Institutional Details

A.1 Co-payment Function

The following Table show the co-payment function of patients. Note that the co-payment covers all prescription drugs. Costs for pharmaceuticals that are not in the benefit scheme are not covered. OTC drugs are not subsidized. Specific pharmaceuticals for children (younger than 18 years), pharmaceuticals against communicable diseases, insulin and pharmaceuticals for persons lacking perception of their own illness are fully subsidized.

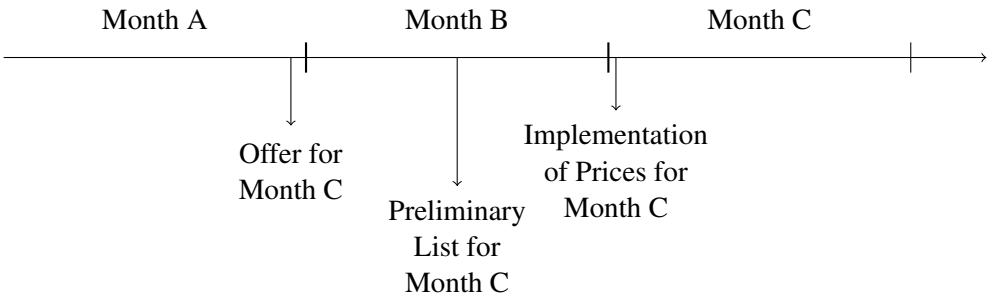
A.2 Timeline of Auction

In Figure 1 I show the details of the sealed-bid first-price auction. Timing is as as described in the main paper: "A pharmaceutical company that wishes to sell a product submits a price at the end of a month (Month A) for the month after the next (Month C). Companies have to bid a lower price than the specified ceiling.[1] The auctioneer publishes a preliminary list of prices in the middle of the next month (Month B).

[1]The usual ceiling is 35% of the original brand product price before expiration of the patent. In detail, 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

After the supplier of the cheapest product confirms the ability to service the entire Swedish market, prices are implemented.[2] Note that the pharmaceutical companies see the preliminary list for Month C before bidding for the month after that (TLV, 2016b)."

Figure 1: Timeline of Auction



A.3 Trade-margins of pharmacies

Table 2: Trade-margins of Pharmacies

Purchasing Price (PP)	Retail Price
$PP \leq 75$	$PP \times 1.20 + 30.50 + 11.50$
$75 < PP \leq 300$	$PP \times 1.03 + 43.25 + 11.50$
$300 < PP \leq 50,000$	$PP \times 1.02 + 46.25 + 11.50$
$PP > 50,000$	$PP + 1,046.25 + 11.50$

Retail prices of pharmaceuticals under generic competition in dependency to purchasing prices since 04/2016 (TLV, 2016c). Trade margins are implicitly defined. Note that the 11.50 KR apply due to the generic competition. Prices in SEK; 10 SEK are approximately 1.1 USD.

Purchasing Price (PP)	Retail Price
$PP \leq 75$	$PP \times 1.20 + 31.25 + 10.00$
$75 < PP \leq 300$	$PP \times 1.03 + 44.00 + 10.00$
$300 < PP \leq 6,000$	$PP \times 1.02 + 47.00 + 10.00$
$PP > 6,000$	$PP + 167.00 + 10.00$

Retail prices of pharmaceuticals under generic competition in dependency to purchasing prices before 04/2016 (TLV, 2016c). Trade margins are implicitly defined. Prices in SEK; 10 SEK are approximately 1.1 USD.

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, 2016a).

[2]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.

B Reduced Form Analysis

B.1 Other Experts

To test if pharmacists indeed oppose substitution more or less often compared to other health care experts, I provide additional regression evidence. The regression models are exactly like the one in the main manuscript except that I consider patients with a pharmacy degree and patients with an education in nursing, in addition to patients with medical (physician) education. Consider the following regression:

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

where only the variable *Expert* differs from the main regression model in the manuscript. *Expert* indicates the dummy for the professions I use, and I run separate regressions. In each I consider a dummy that takes the value 1 if patients have a specific education. The specific types of education are (1) medical doctor, (2) pharmacist, (3) medical doctor or pharmacist, and (4) nursing. In Table 3 I show results for painkillers. Each column refers to a different dummy, and the first column is identical to the results in the main manuscript. The results reveal that compared to the general sample, individuals with an education as a pharmacist also are less likely to oppose substitution. Nevertheless, the probability to oppose substitution is higher for patients educated as pharmacists compared with those educated as medical doctors. When considering both together (see column 3) the effect is still strong. Finally, I do not observe any effect for patients with nursing education. For those, I do not observe a higher likelihood of opposing substitution.

Overall, I observe that the education as a pharmacist does not relate to different behavior as strongly as it does for patients educated as medical doctors. One possible reason for the difference could be that physician have longer education and in their work make decisions on medical issues with much greater autonomy than a pharmacist.

B.2 Patent Expiries, Additional Substances

Within the following, I first present in Table 4 the results referring to the instrumental variable results of Oxyodone. Then I present additional substances that ran off patent. As in the main analysis I use the quasi-experimental setting to show causal evidence of switching costs. For Rizatiptan (a painkiller treating

Table 3: Regression Evidence, Probability of Opposed Substitution, Different Education

	Painkillers			
	Opposed Substitution			
	(1)	(2)	(3)	(4)
D_{t-1}	0.0325***	0.0325***	0.0324***	0.0328***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Med	−0.0537***			
	(0.0017)			
Pharma		−0.0163***		
		(0.0031)		
Med or *Pharma*			−0.0455***	
			(0.0015)	
Nursing				−0.0020
				(0.0014)
$log(Inc)$	−0.0004***	−0.0005***	−0.0004***	−0.0005***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
$D_{t-1} \times Med$	0.0017			
	(0.0039)			
$D_{t-1} \times Pharma$		0.0057		
		(0.0050)		
$D_{t-1} \times$ *Med* or *Pharma*			0.0054*	
			(0.0031)	
$D_{t-1} \times Nursing$				−0.0129***
				(0.0016)
Education	Yes	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	SubGroup×Year-Month
Mean Opp. Subst.	0.192 (0.394)	0.192 (0.394)	0.192 (0.394)	0.192 (0.394)
Mean Price SEK	99.965 (138.778)	99.965 (138.778)	99.965 (138.778)	99.965 (138.778)
Mean Overpayment SEK	9.555 (26.446)	9.555 (26.446)	9.555 (26.446)	9.555 (26.446)
Med Increase	−27.96%	−8.49%	−23.71%	−1.04%
N	36,955,992	36,955,992	36,955,992	36,955,992
R^2	0.2264	0.2263	0.2264	0.2263

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

Notes: Linear least squares regression results for the segment of painkillers. One observation corresponds to one specific purchase occasion by a patient. The outcome variable is a dummy variable that takes the value 1 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. Pharma is a dummy that takes the value 1 if an individual has an education as a pharmacist. Med or Pharma is a dummy that takes the value 1 if an individual has an education as a medical doctor or a pharmacist. Nursing is a dummy that takes the value 1 if an individual has an education as a nurse. 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. In the lower part of the table I show 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 that are 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.

Table 4: IV Regression, Oxycodone

	First Stage 3 M	β_3	First Stage 6 M	β_6
204	0.437***	0.771***	0.439***	-0.197***
	(0.013)	(0.042)	(0.024)	(0.076)
206	0.108***	0.476***	0.113***	0.141***
	(0.005)	(0.045)	(0.006)	(0.051)
207	0.417***	0.464***	0.412***	0.116**
	(0.011)	(0.035)	(0.015)	(0.049)
211	0.304***	0.573***	0.211***	1.332***
	(0.023)	(0.108)	(0.031)	(0.302)
213	0.068***	0.3***	0.095***	0.015
	(0.006)	(0.104)	(0.01)	(0.083)
214	0.339***	0.501***	0.321***	0.465***
	(0.016)	(0.066)	(0.02)	(0.097)
218	0.269***	1.264***	0.277***	0.652*
	(0.05)	(0.326)	(0.067)	(0.354)
219	0.021***	-0.191	0.012*	-0.736
	(0.008)	(0.599)	(0.007)	(0.937)
220	0.29***	0.475***	0.249***	0.707***
	(0.028)	(0.132)	(0.035)	(0.224)
223	0.32***	0.774***	0.416***	0.371***
	(0.005)	(0.019)	(0.01)	(0.037)
229	0.185***	1.438***	0.123***	-0.115
	(0.045)	(0.45)	(0.046)	(0.719)

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

Results of the Instrumental Variable regression for different substitution groups of Oxycodone (painkiller, original product OxyContin). The first column differentiates between the substitution groups. The first stage of the IV regression is shown in columns two and four, first for the initial three months and second for months four to six. The coefficients for the second stage are in columns three and six. Standard errors in parentheses.

migraine) as well as Clindamycin (an antibiotic used against a number of bacterial infections) I present the first and second stage of the IV identification strategy. Table 5 and Figure 2 correspond to Rizatriptan, and Table 6 and Figure 3 describe results for Clindamycin.

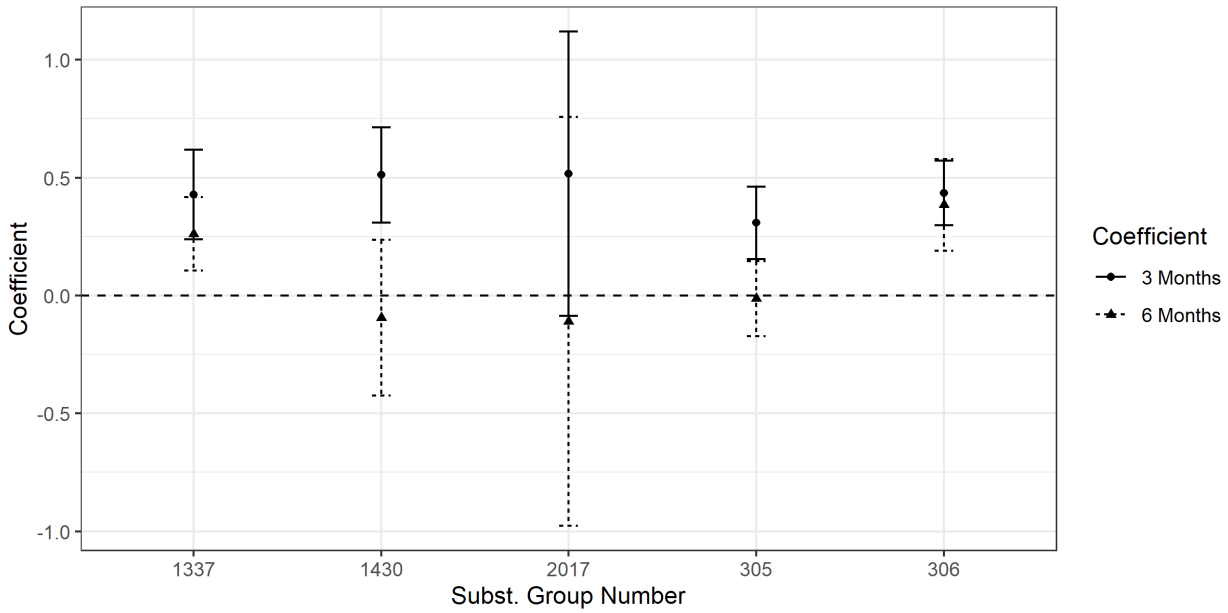
Table 5: IV Regression, Rizatriptan

	First Stage 3 M	β_3	First Stage 6 M	β_6
305	0.612***	0.308***	0.504***	-0.013
	(0.04)	(0.078)	(0.037)	(0.081)
306	0.587***	0.435***	0.562***	0.384***
	(0.035)	(0.07)	(0.048)	(0.099)
1337	0.547***	0.429***	0.664***	0.262***
	(0.043)	(0.097)	(0.041)	(0.079)
1430	0.506***	0.511***	0.386***	-0.094
	(0.039)	(0.103)	(0.042)	(0.167)
2017	0.426***	0.517*	0.278***	-0.11
	(0.087)	(0.303)	(0.082)	(0.435)

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

Results of the Instrumental Variable regression for different substitution groups of Rizatriptan (painkiller, original product Maxalt). The first column differentiates between the substitution groups. The first stage of the IV regression is shown in columns two and four, first for the initial three months and second for months four to six. The coefficients for the second stage are in columns three and six. Standard errors in parentheses.

Figure 2: IV Regression, Rizatriptan



Coefficients of second stage for different substitution groups of Maxalt. Coefficients for each substitution group are divided into coefficients for the initial three months and months four to six. A coefficient for the first three months is equal to $\beta_3 = .5$ and says that the initial consumption of a generic increases the possibility of purchasing a generic again during the following three months by 50%. Note that I include 95% confidence intervals.

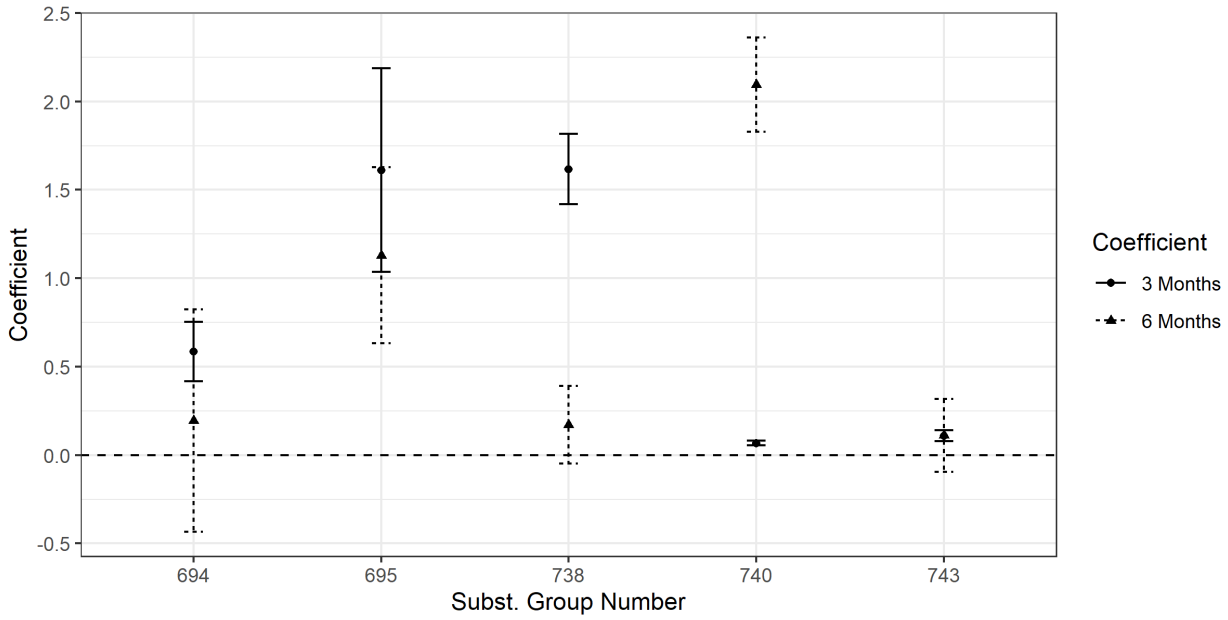
Table 6: IV Regression, Clindamycin

mod3	First Stage 3 M	β_3	First Stage 6 M	β_6
738	0.274***	1.618***	0.251***	0.171
	(0.014)	(0.101)	(0.019)	(0.112)
740	0.145***	0.068***	0.143***	2.095***
	(0.002)	(0.007)	(0.007)	(0.136)
743	0.269***	0.109***	0.223***	0.111
	(0.008)	(0.016)	(0.019)	(0.105)
694	0.152***	0.585***	0.121***	0.195
	(0.011)	(0.085)	(0.021)	(0.32)
695	0.037***	1.612***	0.179***	1.129***
	(0.006)	(0.293)	(0.025)	(0.253)

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

Results of the Instrumental Variable regression for different substitution groups of Clindamycin (an antibiotic). The first column differentiates between the substitution groups. The first stage of the IV regression is shown in columns two and four, first for the initial three months and second for months four to six. The coefficients for the second stage are in columns three and six. Standard errors in parentheses.

Figure 3: IV Regression, Clindamycin



Coefficients of second stage for different substitution groups of Clindamycin. Coefficients for each substitution group are divided into coefficients for the initial three months and months four to six. A coefficient for the first three months is equal to $\beta_3 = .5$ and says that the initial consumption of a generic increases the possibility of purchasing a generic again during the following three months by 50%. Note that I include 95% confidence intervals.

C Demand Model

C.1 Technical Details

Using the choice set I estimate the following structural equation.

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

where ε_{ijst} is independent and identically distributed extreme value type two error. Given the assumption about the error term, one can separate the utility in a deterministic and unobservable part, that is, $U_{ijst} = V_{ijst} + \varepsilon_{ijst}$. The choice probability is then given by[3]

$$P_{ijst} = \frac{e^{V_{ijst}}}{\sum_{k=1}^{K} V_{ijst}}.$$

Note that $K = K_{it}$ denotes the choice set and varies for each consumer i with the choice set at time period t. The deterministic part of the utility relates to the coefficients and variables presented in the structural Equation 2. Let θ be the vector of parameters that is to be estimated and β be some individual specific random effects. The unconditional probability of a particular sequence of choices from an individual is then defined as

$$P_i(\theta) = \int \prod_{t=1}^{T} \prod_{j=1}^{K} [P_{ijst}]^{\mathscr{I}_{ijst}} f(\beta|\theta) \mathrm{d}\beta.$$

The unconditional probability takes into account that an individual makes several choices. $\mathscr{I}_{ijst}$ is an indicator that takes the value 1 if the patient chooses j in t from choice set K_{it}. $f(\beta|\theta)$ is the density function of β and is assumed to be normally distributed. Estimation follows by maximizing the log-likelihood function $LL(\theta) = \sum_i ln(P_i(\theta))$, the sum of the choice probabilities across individuals. The likelihood has to be simulated as it is not a closed form. The log likelihood is simulated by 'Halton-draws', meaning that for each parameter estimate θ, one draws values of β. As suggested by use 50 draws and take the average of

[3] Train (2009) provides a summary of all necessary derivations.

the likelihood function. The final likelihood that is maximized takes the following form

$$SLL(\theta) = \sum_i ln \left\{ \frac{1}{100} \sum_{r=1}^{100} \prod_{t=1}^{T} \prod_{j=1}^{K} [P_{ijst}^{[r]}]^{\mathscr{I}_{ijst}} \right\}$$

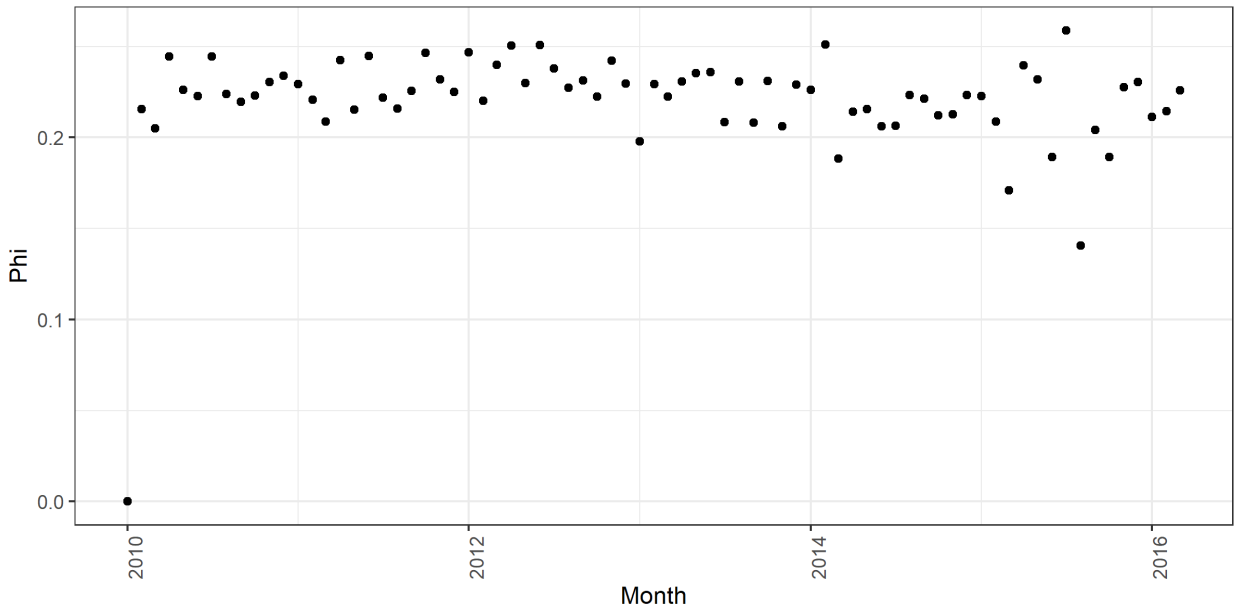
$P_{ijst}^{[r]}$ is the probability of the r-th draw for patient i. Given the estimates of the structural equation, one can compute individual-choice probabilities and market shares.

D Supply Model

D.1 Aggregate Transition Probabilities

Figure 4 shows the fraction of consumers that stay in a market at t to $t+1$.

Figure 4: Share of Staying Patients, ϕ



Monthly ϕ_t, the fraction of consumers that stay in a market at t to $t+1$, where one period corresponds to a calendar month. If a patient purchases the same product at least once in two subsequent months, he or she is considered as staying in the market.

D.2 Marginal Cost Estimates

Month and firm-specific marginal cost estimates:

Table 7: Marginal Cost Estimates, Time and Product Specific

Time Period	Original		Brand.Generic I		Brand.Generic II		Generic I		Generic II		Generic III		Generic IV	
1	9.87	(0.32)			22.89	(0.22)								
2	6.50	(0.51)			25.45	(0.19)								
3	9.07	(0.39)			24.19	(0.23)								
4	7.33	(0.5)			23.82	(0.23)								
5	8.90	(0.46)			24.40	(0.27)								
6	5.75	(0.37)			23.82	(0.18)								
7	8.47	(0.34)			24.23	(0.17)								
8	7.35	(0.48)			24.69	(0.18)								
9	8.07	(0.64)			24.75	(0.17)								
10	8.18	(0.34)			24.40	(0.18)								
11	20.87	(0.14)	19.45	(0.23)	22.72	(0.14)								
12	20.67	(0.16)	19.09	(0.27)	23.06	(0.09)								
13	20.00	(0.12)	17.68	(0.27)	22.49	(0.13)								
14	21.25	(0.13)	18.38	(0.31)	23.61	(0.12)								
15	22.08	(0.21)	11.61	(0.14)										
16	20.57	(0.19)	20.23	(0.18)							22.34	(0.16)		
17	19.48	(0.15)	19.83	(0.17)							21.33	(0.13)		
18	19.02	(0.16)	19.38	(0.22)							21.06	(0.11)		
19	20.50	(0.13)	20.64	(0.18)							22.27	(0.08)		
20	21.32	(0.19)	20.37	(0.27)							23.17	(0.11)		
21	20.82	(0.15)	19.35	(0.33)							22.58	(0.09)		
22	20.71	(0.09)	17.45	(0.36)					21.20	(0.11)	21.54	(0.08)		
23	21.03	(0.12)	17.91	(0.27)					21.68	(0.12)	21.85	(0.11)		
24	20.04	(0.08)	17.07	(0.2)					20.68	(0.09)	21.12	(0.09)		
25	21.01	(0.14)	17.56	(0.31)					21.70	(0.09)	21.78	(0.09)		
26	21.30	(0.12)	18.17	(0.4)					21.82	(0.11)	22.16	(0.08)		
27	20.51	(0.07)	16.35	(0.22)					21.36	(0.11)	21.52	(0.06)		
28	20.96	(0.1)	17.66	(0.29)					22.23	(0.12)	22.18	(0.1)		
29	20.79	(0.15)	17.08	(0.26)					21.77	(0.08)	21.50	(0.13)		
30	20.34	(0.08)	16.87	(0.29)					20.87	(0.11)	21.15	(0.09)		
31	21.02	(0.08)	17.22	(0.21)					22.02	(0.11)	22.02	(0.09)		
32	20.76	(0.13)	17.01	(0.26)					21.54	(0.09)	21.48	(0.07)		
33	21.60	(0.08)	17.21	(0.26)					22.43	(0.12)	22.38	(0.09)		
34	21.02	(0.09)	16.75	(0.28)					21.64	(0.11)	21.64	(0.06)		
35	18.90	(0.1)	15.31	(0.21)					19.60	(0.07)	19.63	(0.1)		
36	20.92	(0.1)	17.94	(0.27)					21.50	(0.14)	21.80	(0.09)		
37	20.81	(0.11)	17.50	(0.42)					21.51	(0.1)	21.65	(0.06)		
38	21.01	(0.07)	17.05	(0.41)					21.39	(0.14)	21.71	(0.09)		
39	20.86	(0.13)	17.55	(0.36)					21.72	(0.13)	21.79	(0.1)		
40	20.59	(0.06)	17.17	(0.22)					21.39	(0.13)	21.42	(0.07)		
41	19.86	(0.09)	16.32	(0.18)					20.39	(0.09)	20.71	(0.07)		
42	20.35	(0.12)	16.74	(0.27)					20.94	(0.08)	20.98	(0.09)		
43	20.29	(0.11)	15.98	(0.22)					20.84	(0.09)	20.88	(0.04)		
44	20.67	(0.12)	16.39	(0.34)					21.44	(0.13)	21.58	(0.05)		
45	20.16	(0.13)	16.45	(0.16)					20.74	(0.07)	20.95	(0.05)		
46	20.73	(0.09)	17.08	(0.32)					21.27	(0.06)	21.45	(0.08)		
47	19.43	(0.08)	16.63	(0.24)			19.92	(0.06)	19.87	(0.07)	19.76	(0.07)		
48	20.36	(0.11)	17.46	(0.26)			20.80	(0.06)	20.72	(0.05)	20.60	(0.07)		
49	18.58	(0.1)	15.36	(0.18)			18.69	(0.06)	18.92	(0.04)			18.93	(0.07)
50	18.78	(0.09)	15.71	(0.32)			19.04	(0.06)	19.35	(0.07)			19.39	(0.04)
51	18.94	(0.07)	14.86	(0.22)			19.13	(0.03)	19.30	(0.04)			19.43	(0.04)
52	18.73	(0.08)	15.42	(0.09)			18.92	(0.05)	19.32	(0.05)			19.39	(0.08)
53	18.56	(0.13)	15.29	(0.22)			18.85	(0.04)	19.14	(0.04)			19.12	(0.07)
54	19.03	(0.12)	15.51	(0.24)			19.26	(0.04)	19.50	(0.05)			19.60	(0.07)
55	19.15	(0.12)	15.44	(0.15)			19.18	(0.03)	19.57	(0.05)			19.52	(0.04)
56	18.60	(0.09)	15.21	(0.25)			18.80	(0.05)	19.01	(0.04)			19.05	(0.05)
57	18.99	(0.05)	15.78	(0.15)			19.16	(0.06)	19.56	(0.03)			19.69	(0.05)
58	18.88	(0.11)	15.96	(0.18)			19.18	(0.03)	19.43	(0.03)			19.49	(0.07)
59	18.64	(0.06)	15.27	(0.15)			18.65	(0.03)	19.02	(0.05)			18.99	(0.04)
60	18.51	(0.11)	15.02	(0.27)			18.82	(0.02)	19.03	(0.03)			19.13	(0.06)
61	17.51	(0.12)	14.69	(0.23)			17.79	(0.02)	18.02	(0.05)			18.13	(0.04)
62	19.69	(0.08)	16.41	(0.22)			20.10	(0.05)	20.23	(0.06)			20.45	(0.06)
63	19.55	(0.06)	16.11	(0.3)			19.58	(0.05)	19.90	(0.04)			19.91	(0.04)
64	18.11	(0.1)	14.90	(0.21)			18.27	(0.02)	18.58	(0.04)			18.52	(0.05)
65	19.53	(0.12)	14.26	(0.22)			20.15	(0.06)					20.09	(0.11)
66	17.05	(0.04)	12.79	(0.11)			17.57	(0.04)					17.56	(0.07)
67	19.25	(0.06)	14.22	(0.18)			19.81	(0.05)					19.75	(0.08)
68	18.80	(0.08)	14.39	(0.24)			19.62	(0.05)					19.27	(0.07)
69	20.10	(0.16)	16.73	(0.25)			21.08	(0.11)						
70	19.94	(0.11)	16.92	(0.17)			21.47	(0.09)						
71	19.09	(0.11)	15.79	(0.23)			20.65	(0.1)						
72	20.04	(0.18)	16.33	(0.22)			21.44	(0.07)						
73	19.96	(0.15)	16.81	(0.29)			21.16	(0.1)						
74	19.64	(0.09)	16.79	(0.18)			21.03	(0.11)						

Marginal costs estimates (in SEK) for each brand in the time between 2010 and 2016. Periods are monthly and sequentially numbered from 1 to 74. The monthly periods represent the time between 2010 and 2016. Within each period for each competitor marginal costs are estimated. Note that the standard errors are obtained by bootstrapping and reported in parentheses.

E Counterfactual Analysis

E.1 Reduction of State Space

For computational feasibility I reduce the state space of firms. I reduce the state space by using a machine learning method.

First, I get back to the first stage of the previous approach of Bajari et al. (2007). I re-estimate the policy function that has been used to recover marginal costs. However, I use the LASSO that selects state variables by shrinking coefficients within the policy function to zero. Econometrically the LASSO methods uses a standard least squares function with a penalty term on the coefficients. Due to the penaly some coefficients are exactly zero such that only the most important variables are non-zero. The method is related to the solution concept presented by Thiel (2019) who uses the two stage model by Bajari et al. (2007) with a lasso estimation in the first stage. Thiel (2019) further builds on the assumption that firms play a Sparse Markov Perfect Equilibrium, such that the second stage is build on the policy function estimated by the LASSO. In comparison, I use all state variables to recover marginal costs. However, I use a LASSO to reduce the state space in order to make the computation of a new Markov Perfect Equilibrium feasible. Consider the the same policy function as in the first stage of the two-stage estimation.

$$p_{jt} = \alpha + \beta m_{jt-1} + \eta |N_{t-1}| + \rho Q_t + \gamma PoM_{jt-1} + \varepsilon_{jt}. \tag{3}$$

I estimate equation 3 using a LASSO with a unit ($\lambda = 1$) strength of regularization. Table 8 shows three models. Model (1) is a regular OLS regression, i.e. the same as the policy estimation used for the BBL method of the main paper, model (2) shows coefficients of the LASSO, and model (3) an OLS regression using solely regressors with non-zero coefficients of the LASSO. The results show that the lagged product of the month status has a non-zero coefficient in the LASSO estimation. Thus, I assume that the lagged product of the month status incorporates the important inter temporal association from past to current periods. In my counterfactual analysis I compute Markov Perfect Equilibria considering the previous PoM status as the state space.

Table 8: Policy Estimation, LASSO Regression

	Price		
	OLS	LASSO	Restricted OLS
	(1)	(2)	(3)
Share$(t-1)$	1.589*		
	(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)		
$PoM(t-1)$	1.752***	0.51	2.548***
	(0.469)		(0.415)
$Quantity$	0.0001***		
	(0.00003)		
Constant	61.178***	71.58	70.753***
	(2.313)		(0.264)
N	272	272	278
R^2	0.262		0.120
Notes:	*** $p<0.001$, ** $p<0.001$, * $p<0.05$		

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 g., 30 tablets. The outcome variable is the price of a product in period t. All regressors are potential 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. Model (1) is a linear least square regression. Model (2) is a LASSO regression with $\lambda=1$. Model (3) is a OLS regression using only the regressors with non-zero coefficients of the LASSO specification. Standard errors are reported in parentheses.

E.2 Technical Details of Counterfactuals

Within this section I explain the technical details of the counterfactual derivation. Given the supply-side estimation I have firm- and period-specific marginal costs estimates. Knowing the market structure and not allowing for entry and exit, I reduce the state space significantly. In detail, I assume that firms condition their pricing solely on the knowledge of which firm was the PoM (has offered the cheapest product) in the last period. In the following I explain the details of deriving a Markov perfect equilibrium. Note that I the following steps of computation are performed in each period each period; I drop time subscripts for convenience.

There are $N = \{1, ..., n\}$ in the market. Each firm $j \in N$ sets a price $p_j \in P$. For the computation I use discrete values for the price, namely $P = \{15.125, 15.375, ..., 27.875, 28.125\}$. Note that the supply prices are directly linearly translated into retail prices for patients. Given the linear translation into retail prices I do not have to consider problems of double marginalization. Note also that the highest possible price is equal to the price ceiling. When estimating marginal costs I have considered a wide state space on which firms N condition their pricing strategy on past market shares, the number of competitors, the quantity of products and the past PoM status. To reduce the intense computational burden in the counterfactual analysis I reduce the state space remarkably to solely the PoM status, such that $\mathscr{S} = PoM$.[4] The state variable takes the value one if a firm has been PoM in the past period and zero otherwise, $PoM = \{1, 0\}$. Intuitively, firms based their strategy if they have been PoM in the previous period. I allow that strategies differ across different firms and within each periods.

Prices as well as the state variable determine supply-side determinants for demand $\hat{D}$. The demand is divided in two parts. A first part of the demand was already part of the market in the last period. A second part consists of new consumers. The share of the initial part is given by ϕ, a direct share from the data. Further, for the first part the state variable matters, as the demand estimates of state dependence is important. In practice I calculate $\hat{D}$ by considering both demand parts, incorporating the respective state variable. Finally, I simplify demand estimation in a similar fashion as done for the marginal cost estimation as in the previous sections. So for both demand fractions I consider the demand estimates and use the average consumer when evaluating random coefficients. In difference to the estimation of marginal costs I do not adjust choice sets as I assume that firms do not know about the identity of the previous firm which indeed

[4] See Section E.1 for a motivation.

would directly increase the probability of being in the choice set. The second computational simplification on the demand side is that I do not keep track about potential correlation between the past choices of the average consumer and their initial choices. It is likely that initial choices are correlated with past choices such that the inter-temporal effects are higher. I therefore do not integrate the initial choices and reduce the effect of state dependence.[5] Overall the procedure of estimating equilibira requires three deviations from the marginal cost estimation: (1) A reduction of state space, however, (2) allowing for firm specific strategies. (3) An adjustment of choice sets.

Static one-period profits are $\pi_j = \hat{D}_j(p_j - \hat{mc}_j)$ where $\hat{mc}_j$ are estimates of the marginal costs.

The structural model incorporates the dynamic feature of prices and demand. As in the estimation of marginal costs I assume that firms in a given period are forward looking. To decrease the computational burden, I have assumed that firms estimate that the consumers do not change. Also in the counterfactuals I consider the same demand for the next period, while still incorporating the effect of being the PoM for the next period. Formally the prices of a period $(p_j)_{j \in N}$ affect the transition of the state variables. Denoting the state variable of the forthcoming period with PoM', the transition can be formally described by

$$PoM' = \begin{cases} j & \text{if } p_j < p_{-j} \\ j & \text{if } p_j = p_{-j} \quad \text{and} \quad PoM = 1 \end{cases}$$

So the strictly cheapest product or gets the PoM status. If multiple firms are the cheapest, the firm that has been PoM in the previous period gets the status.[6] Having specified the transition function, I can write the dynamic continuation payoff for firm j as

$$V_j(p_j, p_{-j}|PoM) = \hat{D}_j(p_j - \hat{mc}_j) + \delta V_j(p'_j, p'_{-j}|T(PoM, P^n))$$

Note that δ is a discount factor. The first term is the static profit $\pi(P^n, PoM)$; the second term is the continuation payoff given the static prices.

I am searching for a Markov perfect equilibrium. In line with (Maskin and Tirole, 2001), a Markov perfect equilibrium restricts subgame perfect equilibria only to the pay-off relevant strategies of a sub-

[5] Note that both simplifications decrease the computational burden immensely when forward simulating profits and therefore the demand. I access the fit when estimating the benchmark model where equilibrium prices are comparable to the actual data.

[6] The result is robust to several adjustments. For example one could build a tie breaker rule such that in case several products have the same price and one of those products is an original, the original.

game. Each firm conditions its strategy $\sigma_j \in S_j$ to the state variable, i.e. $S_j : PoM \to \Delta(P)$. A strategy $\sigma_j^* \in S_j$ forms a Markov perfect equilibrium if and only if for all $\sigma_j \in S_j$ and $PoM \in \mathscr{S}$ it holds that $V_j(\sigma_j^*, \sigma_{-j}^* | PoM) \geq V_j(\sigma_j, \sigma_{-j}^* | PoM)$. Note that I allow strategies to differ across the firms. This is important as firms have different demand parameters that affect their strategies.

To estimate the equilibrium strategies $(\sigma_j^*)_{\forall j}$ I use a value function iteration. While the dynamic programming approach is easy to implement, the difficulty arises as the computation also involves a computation of a game theoretic equilibrium. I use an algorithm in the fashion of Pakes and McGuire (1992). The general value function algorithm is defined by the following three steps that take time in each period separately (In each period different marginal costs from the marginal cost estimates are used):

Step 0: I create an educated initial guess of the value function, i.e. V^n for each possible state variable and for each firm the value function is defined. Set n=0.

Step 1: For each possible state variable *PoM* (zero or one), I compute a the best reply in terms of prices P given the continuation payoff V^0 and the prices of opponents. In other words, I search for each player for the best prices given the transition function, the defined continuation value as well as the prices of opponent firms. For each firm the new valuation function is given by $V^{n+1} = \pi(\sigma^*, PoM) + V^n(\sigma^*, PoM)$.

Step 2: I test if $|\frac{V^{n+1}}{V^n}| < Toler$ where $Toler$ is a tolerance level. If the value function and strategies (prices) have not converged, go back to *Step 1*.

After convergence I have found equilibrium strategies as well as valuation functions for each firm. During each iteration n increases by one.

The main concerns of the equilibrium estimation are the following: First, the algorithm does not select equilibria. Indeed I only search for pure strategy equilibria. In only three periods I do not converge to a pure equilibrium. Another concern is that there exists multiple equilibria and the algorithm just uses one. The second concern is that the lack of equilibrium selection impacts the valuation function iteration. If there are multiple equilibria and the algorithm in one iteration gives me one equilibrium and in the next iteration another one, the valuation function may not converge. I recognize the potential problems of the algorithm and try to test several robustness checks: First, I start with different initial guesses. The equilibrium outcome for different initial guesses does not change. Second, I consider if a 'bang-bang' behavior between different equilibria during the iteration is a problem. Indeed, the convergence mostly seems smooth and I do not observe any jumps between equilibria between iterations.

A very similar problem arises through wrong inference by equilibrium selection. In comparison to the actual equilibria in the data it may be possible that the counterfactuals select a different equilibrium solely because of the algorithm or starting values. I tackle the concern via the benchmark model. Within the benchmark model I evaluate the market without making changes.

After having establised an equilibrium in each period, for each firm, and in all states I continue to simulate prices. I simulated prices over all monthly (or yearly) periods one thousand times to avoid mistaken inference from a specific equilibrium past. In each simulation I start with a random state. Between time periods, the states evolves in the following manner. If a firm has the strictly lowest price in t, it has state dependent consumers in the next period. However, if two firms set the same price, I randomize the state in the forthcoming periods. The average of the prices across the thousand simulations of the sequential time periods are used for inference.

E.3 Reduction and Increase of Switching Costs

Within this section I evaluate three additional counterfactuals: (1) I evaluate a counterfactual with more repeated patients. In detail, I do not change switching costs but only the share of patients that stay over periods is tripled. In theory, the effect should be similar to higher switching costs. As firms have higher base of patients staying over periods, there is a higher incentive for a "lock-in and harvest" pricing strategy. (2) I evaluate a counterfactual without any switching costs. (3) I consider a counterfactual scenario where the switching costs are doubled.

The motivation of the first and second counterfactual exercise is to analyze if results of the increased contract length match those of a decreased switching costs. In comparison to the counterfactual of an increased procurement contract length I change the preference parameter of consumers directly. I set the coefficient of previous choices (the coefficient of state dependence in the discrete choice model) equal to zero such that past choices do not effect consumers' current utility. Besides the change of the demand calculation I estimate the prices as in the benchmark model.

In Table 9 I show the results. The general results are as expected given the results of the policy scenario with longer procurement contracts. Compared to the benchmark, more repeated consumers as well as higher switching costs (higher ρ) decrease prices while the case of no switching costs increases prices. Also the standard deviation speak in favor of dynamic pricing. Increasing switching costs or repeated con-

sumers increase the standard deviation while it is lower without switching costs. Intuitively, the sporadic decrease of prices to lock-in consumers increases the standard deviation. This is also confirmed by lower minimum prices on average in the case of higher switching costs or more repeated patients. In contrast, the no switching costs case shows increased prices. Finally also the average price patients pay shows the same results.

Table 9: Additional Counterfactual Results

	Benchmark	High ϕ	High SC	No SC
Mean Price	70.11	69.38	69.47	71.26
	(2.19)	(2.49)	(2.54)	(1.83)
Mean Minimum Price	69.84	68.82	68.74	71.29
Average Price for Consumer	70.49	69.51	69.62	71.82
Compared to Benchmark		-1.39%	-1.23%	1.89%

Notes: The Table shows additional counterfactual results. The benchmark refers to the benchmark model. In the first counterfactual, I triple the share of repeated patients (ϕ). In the second counterfactual, the coefficient of state dependence (ρ) are tripled . Therefore, the counterfactual considers the case of high switching costs. In the last counterfactual I set the state dependence coefficient (ρ) to zero. Standard deviations are reported in parentheses.

F Back-of-the-envelope Calculation

Within the back-of-the-envelope calculation, I establish a connection between reduced form estimates and the impact of counterfactuals on prices. To do so, I assume there is a linear relation between estimates of the behavioral frictions and the effect on prices, assuming that if we do not observe any behavioral friction, the impact on prices is zero. The linear relation is summarised in Figure 5 where the two coordinates define the linear mapping: The origin, where I assume that of the paracetamol, 1 gr substitution group, where the switching cost coefficient in the demand model is -0.313.

In the following analysis, I use the substitution group of paracetamol, and reestimate the counterfactual of a longer procurement length for multiple switching cost estimates. I vary the switching cost estimates between -0.3 to 1.2, where a larger coefficient refers to higher switching costs. The variation is similar to potential estimates in other substitution groups. For each hypothetical switching cost, we reestimate the counterfactual of a longer procurement length and recalculate the relative price difference between a benchmark and counterfactual scenario.[7] The estimation of prices follows the same procedure as the coun-

[7] In detail, we recalculate for each potential switching cost coefficient the benchmark and counterfactual price.

terfactual of the main article. Figure 5 shows point estimates of the results. Given the point estimates, I also add a linear regression line that is dotted.

First, we observe the general negative relation between switching costs and the relative price differences. Higher switching costs indeed lead to a stronger increase in prices in the counterfactual. The estimates are not exactly linear. However, multiple reasons could explain the phenomenon. For example, the counterfactual calculations do not use continuous prices but a grid of prices. The linearization of switching costs to relative price differences could be slightly more negative than the actual relationship. For most products, we observe switching cost coefficients between 0.3 and 0.5; therefore, the linearization used is a reasonable approximation for a back-of-the-envelope calculation.

References

Patrick Bajari, C Lanier Benkard, and Jonathan Levin. Estimating dynamic models of imperfect competition. *Econometrica*, 75(5):1331–1370, 2007.

Eric Maskin and Jean Tirole. Markov perfect equilibrium: I. observable actions. *Journal of Economic Theory*, 100(2):191–219, 2001.

Ariel Pakes and Paul McGuire. Computing markov perfect nash equilibria: Numerical implications of a dynamic differentiated product model, 1992.

Jurre Thiel. Price discrimination, switching costs and welfare: Evidence from the dutch mortgage market. *Unpublished*, 2019.

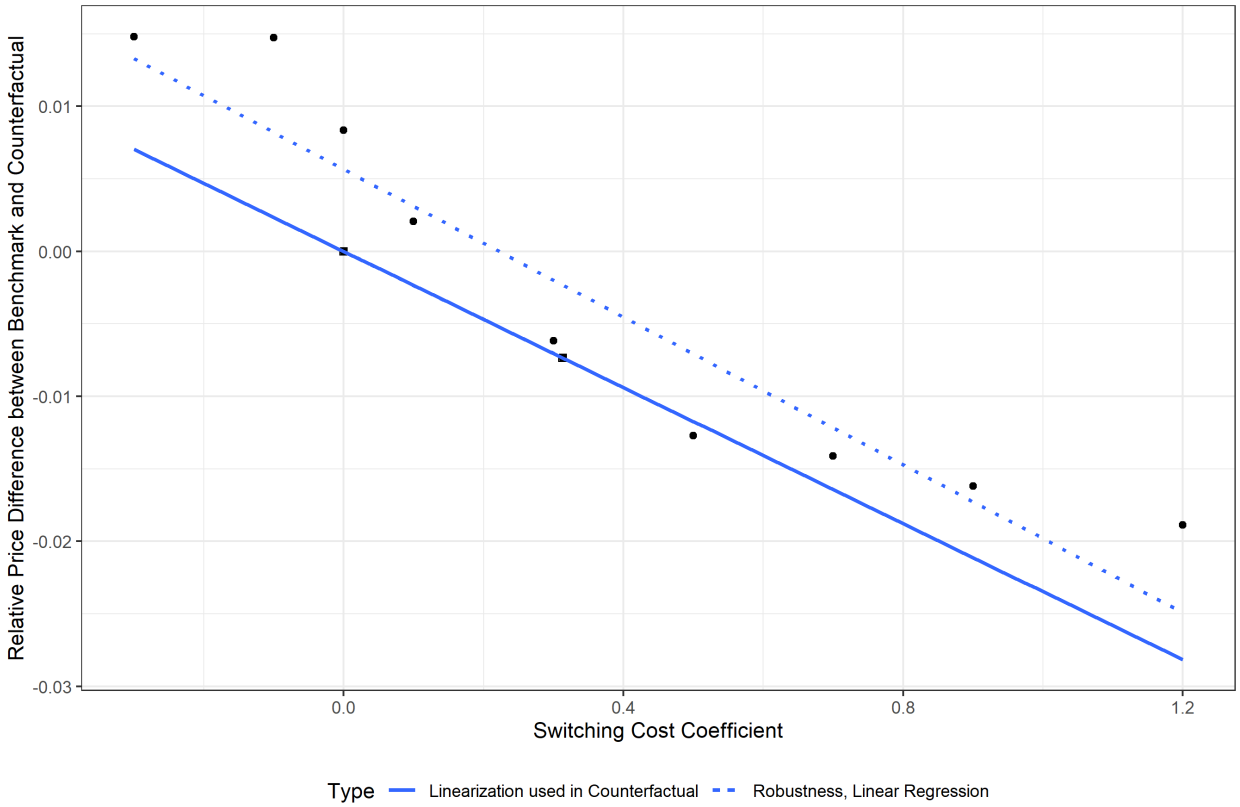
TLV. Takpriser. *http://www.tlv.se/lakemedel/takpriser/*, 2016a. Accessed: 2017-02-23.

TLV. Periodens varor. *http://www.tlv.se/apotek/utbyte-av-lakemedel-pa-apotek/periodens-varor/*, 2016b. Accessed: 2017-02-23.

TLV. Apotekens marginaler. *http://www.tlv.se/apotek/Apotekens-marginaler/*, 2016c. Accessed: 2017-02-23.

Kenneth E Train. *Discrete choice methods with simulation*. Cambridge university press, 2009.

Figure 5: The Relation of Switching Costs and Prices in the Counterfactual of Longer Procurement Length



The Figure describes the association between the switching cost coefficient within the demand model and the relative price difference between the benchmark and the counterfactual model. A negative price difference corresponds to higher prices in the counterfactual scenario. The solid line refers to the linear mapping used in the back-of-the-envelope calculation. The linear mapping is based on the origin as well as demand and supply estimation in the 1 gr paracetamol substitution group. Each dot corresponds to a robustness calculation: Between hypothetical switching cost coefficients of -0.3 to 1.2 in steps of 0.2 and for switching costs of 0 I calculate relative price differences between the benchmark model and the counterfactual in the paracetamol substitution group. A positive price difference refers to higher average prices in the benchmark model. Note that the prices are average prices across all products and periods. The dotted line shows a linear regression of the relative price difference on the hypothetical switching cost coefficients.