7-2023

The importance of the first generic substitution: Evidence from Sweden

Aljoscha JANSSEN
Singapore Management University, ajanssen@smu.edu.sg

David GRANLUND

The Importance of the First Generic Substitution: Evidence from Sweden*

Aljoscha Janssen† David Granlund‡

April 1, 2022

Abstract

We analyze changes in the willingness to substitute from prescribed pharmaceuticals to more affordable generic equivalents in response to the first experience with a substitution. Using Swedish individual-level data of prescribed and dispensed pharmaceuticals, we employ a dynamic event study and an instrumental variable approach to show that an initial substitution reduces the probability of opposing subsequent substitutions by 39 percentage points. The impact of a first substitution is especially large among elderly patients. We recommend that policy-makers target patients with a history of opposed substitution and offer additional discounts to promote substitution as long-term savings outweigh one-time costs.

JEL: D12, I11, I12

Keywords: Generic Substitution, Pharmaceuticals, Health Care Costs

*The authors are grateful to the County Council in Västerbotten for supplying the data used in this article. Granlund acknowledges support from the Jan Wallander and Tom Hedelius Foundation and the Tore Browaldhs Foundation (Grant number P2016-0113:1).

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

‡Umea University, david.granlund@umu.se

1 Introduction

Substitution of expensive prescription pharmaceuticals with cheaper but medically equivalent options is an important tool in reducing overall health care costs.[1] The majority of advanced economies have implemented reforms involving generic substitution (Panteli et al., 2016). Yet, generic take-up across countries is often still low. In most European countries, the share of generics among total pharmaceutical volume is less than 50%, but increasing (OECD, 2017). This is a stark difference to the United States, where the generic share exceeds 80%. Small-scale, qualitative surveys across countries indicate that positive experiences with prescribed products and uncertainty surrounding treatment with generics lead to a lower take-up rate of generics (Håkonsen and Toverud, 2012).

This article answers the question of whether the first experience of switching to a generic pharmaceutical changes patients' perception and behavior in future decisions involving pharmaceuticals. We take advantage of the Swedish health care system, where generic substitution is financially incentivized but not mandatory. Using rich Swedish patient-level data of prescribed and dispensed pharmaceuticals, we document a large effect of the first substitution from a prescribed pharmaceutical to a more affordable option on following substitution decisions.

We start our empirical analysis with a dynamic event study on the individual level. Using a sample of individuals with a history of opposed substitutions and consumption of non-generic pharmaceuticals, we document that the first substitution from a prescribed product to a more affordable option reduces the probability of opposing substitution and incurring additional monetary costs to receive the prescribed product in following dispenses by approximately 20 percentage points. The difference is stable for subsequent purchase occasions, which leads to the conclusion that the initial substitution, rather than multiple substitutions, has an effect on the future probability of substitution.

The estimate of the dynamic event study is based on the assumption that the initial decision to substitute is random and unrelated to unobservables. If the first substitution is driven by events that change the substitution decisions of some patients, we measure the effect of such events on those individuals rather than the effect of the first substitution. To show robustness and to demonstrate that our estimate is not based on such unobserved heterogeneity, we use an instrumental variable approach based on the fact that some products exit the market. The intuition is that patients with prescriptions for a product that has exited the market are forced to substitute due to the unavailability of the prescribed pharmaceutical. The instrumental variable

[1] Johansen and Richardson (2016) estimate, in a sample of more than 100,000 patients between 2010 and 2012 in the U.S., more than 70 billion USD in excess expenditure due to overuse of expensive, branded drug options. Shrank et al. (2010) show high savings from mandatory substitution for patients enrolled in Medicaid.

estimation shows robustness of the event study as the initial substitution has a large discontinuous effect on the probability of opposed substitutions in future substitution decisions. In detail, we show that a first substitution reduces the probability of future opposed substitutions by 39 percentage points.

We investigate the heterogeneity of our estimates. Using the linear least square estimation and our instrument variable approach, we show that gender is not related to the impact of the first substitution decision. We observe the strongest and most substantial effect of the first substitution on future substitution decisions among the oldest individuals in the sample.

The results document the importance of the first substitution. For patients who tend to oppose substitution repeatedly, co-payments due to opposed substitution account for 34.4% of all co-payments. A back-of-the-envelope calculation reveals that a first substitution could reduce this share to 26.7% of co-payments. We argue that the large savings on the patient level should encourage policy-makers to implement policies that encourage first-time substitution. As an example, insurance providers could waive co-payments for the cheapest available product for patients with a history of opposed substitutions. This would increase the price difference between the cheapest available option and the prescribed product, thereby increasing the probability of a first substitution. In practice, such a subvention of co-payments would increase the price difference between the cheapest available option and the prescribed product, thereby increasing the probability of a first substitution. Another option is to flag those patients needing a first substitution for physicians or pharmacists. The intention would be that physicians or pharmacists could put additional effort into explaining the medical equivalence and safety of cheaper generics.

Our article adds to the health economics literature by shedding light on the impact of the first generic substitution. As we investigate the impact of a first generic substitution on following substitution decisions, our paper is related to the literature that concerns learning in the pharmaceutical market. Crawford and Shum (2005) use different molecules on the market for anti-ulcer drugs and show that patients experience learning. Ching (2010), Ching and Ishihara (2010), Ching et al. (2016) and Coscelli and Shum (2004) model how consumers in the U.S. market learn about, choose, and update their information about drugs.[2] Relatedly, Ketcham et al. (2012) show learning in prescription insurance choice. Our approach differs as we do not model the learning and information updating. Instead, we are the first to show that an initial generic substitution has a large impact on following substitution decisions. The discontinuity could be explained by a positive experience with substitution to a cheaper pharmaceutical in

[2]Note also that advertisement of drugs plays a role in learning about the existence of new pharmaceuticals; see for example Anderson et al. (2013).

general.

Health economic studies in multiple countries have concluded that patients prefer originals and experience history dependence when choosing pharmaceuticals, meaning that they tend to consume products with which they have had a positive experience (Bronnenberg et al., 2015; Feng, 2020; Granlund, 2021; Janssen, 2020).[3] We add to the literature but also highlight that it is possible to break such history dependence. A first substitution could break history dependence as patients experience that the treatment effects of pharmaceuticals are not brand specific.

2 Pharmaceuticals in the Swedish Health Care System

In the following we give a brief overview of prescription drugs in the Swedish health care system. Generally, prescription drugs are reimbursable.[4] The market for off-patent drugs is arranged into groups of medically identical drugs called substitution groups (i.e., groups of pharmaceuticals with the same substance, size, strength, and delivery format), and each drug in a substitution group is interchangeable (Swedish Medical Product Agency, 2010).[5]

Prices within substitution groups can vary on the monthly level and are uniform across Sweden. Patients with a prescription are incentivized to choose the cheapest available product in the substitution group. In detail, the process works as follows (TLV, 2016; Sveriges Riksdag, 2002): When a patient fills a prescription in a pharmacy, a pharmacist or pharmacy employee has to explain the concept of substitution groups and recommend substitution if a prescription is not for the cheapest available product within a substitution group. If a patient opposes substitution, the patient has to bear the cost difference between the cheapest available product and the prescribed product entirely out of pocket. Two exceptions prevent the process of substitution. A prescriber can oppose substitution by marking the prescription accordingly. Further, a pharmacist can prevent substitution if the pharmacist has reason to believe that the consumer could be adversely affected by the substitution, for example, if the low-cost alternative has a package that could be difficult for the consumer to open. Both exceptions are rare; in our sample, a physician opposes substitution in 2.39% of dispenses, while a pharmacist prevents substitution in 2.15%

[3] In a wider sense the paper therefore adds to the general literature of behavioral frictions in health care markets (Abaluck and Gruber, 2016; Handel, 2013; Handel and Kolstad, 2015; Ho et al., 2017; Ketcham et al., 2015; Marzilli Ericson, 2014; Polyakova, 2016). Further, there is evidence that physicians also play a role in generic adoption (Camacho et al., 2011; Chan et al., 2013; Chintagunta et al., 2012; Granlund, 2009; Granlund and Sundström, 2018; Iizuka, 2012).

[4] Decisions about reimbursements are made by the Dental and Pharmaceutical Benefits Agency (TLV). TLV (2017) provides detailed information on the decision process.

[5] Package size is allowed to vary slightly; for example, substitution can be made from a 30-pill package to a package in the range of 28 to 32 pills.

Table 1: Co-payments

Accumulated Pharm. Expenses	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-payments for cumulative health care expenditures, which include prescription drugs. Reimbursement is calculated for expenses during an entire year, beginning with the first expenditure. Prices are in SEK; 10 SEK are approximately 1 USD.

of dispenses. Within this article we only consider opposed substitutions by patients.

While prescription drugs are reimbursable, patients pay co-payments. Table 1 describes how these co-payments depend on yearly costs for pharmaceucicals.[6] Note that pharmaceutical costs above a ceiling are entirely covered.[7] Note also that patients opposing substitution must pay the price difference between the prescribed product and the cheapest available product.

On the supply side, drug prices are result of a tendering system. Manufacturers that would like their product to be part of a substitution group in the pharmaceutical benefit scheme have to submit their bid two months in advance. Prices should not exceed the highest price of a substitution group in the previous month. The Swedish Dental and Pharmaceutical Benefits Agency (*Tandvårds- och läkemedelsförmånsverket*, TLV) collects the bids and announces final substitution groups as well as prices one month before the prices go into effect. Retail prices and purchasing prices follow a simple algorithm whereby the difference between the retail and purchase prices increases with the purchase price. The cheapest product in a substitution group is announced as the product of the month, meaning that prescriptions should be substituted to that product when a patient fills a prescription for any product within the substitution group.

3 Data

We use data on all non-narcotic pharmaceutical consumption by adult inhabitants of Västerbotten County between January 2014 and April 2016.[8] We connect the data to monthly prices for

[6] There are some exceptions: pharmaceuticals for children (younger than 18 years), pharmaceuticals for communicable diseases, insulin, and pharmaceuticals for persons lacking perception of their own illness are fully subsidized. In comparison, over-the-counter drugs are not subsidized.

[7] According to Bergman et al. (2012), almost half of the revenue from prescription pharmaceuticals is from coverage for patients that have reached the cost ceiling.

[8] Västerbotten is one of the 21 Swedish counties, and it had approximately 260,000 residents in 2015.

outpatient pharmaceuticals under generic competition, provided by the TLV. Each observation in our final data set describes one single dispense of a pharmaceutical. Panel A of Table 2 shows basic summary statistics on the product level. We observe 270 substances, measured by a unique Anatomical Therapeutic Chemical (ATC) code.[9] As a substance may be available in several strengths or different forms, there are more substitution groups (1347) than substances. On average there are 3.57 products competing in a substitution group, and 38% of the products are on average the cheapest available product (or one of the cheapest) in a month (termed the product of the month). The average price of a product is 254 SEK (25 USD), and 20% of the observable products are originals.

Besides information on prices, the final data set documents patient-specific prescriptions and dispenses. In practice, we observe a patient over time, including the prescriptions the patient received and which products were purchased. Thereby we can distinguish if a patient opposed substitution, meaning that the patient received a prescription that was not the product of the month and paid the price difference between the prescribed product and the product of the month out of pocket. Separate indicator variables show if instead the prescriber or pharmacist prevented substitution. In Table 2, Panel B describes the data on the level of the purchase occasion. We divide the sample into three subsamples. Besides the full sample, we consider only those prescriptions where substitution is an option, meaning that a physician has neither prescribed the product of the month nor the physician or a pharmacists has prevented substitution. Further, we evaluate the sample where we observe an opposed substitution, meaning that a patient opposed substitution and incurred additional out-of-pocket costs. During the two years and four months, we observe 4.5 million purchases, and in approximately half of them a substitution was possible. In approximately 270,000 occasions (5% of all purchases) a patient opposed substitution.

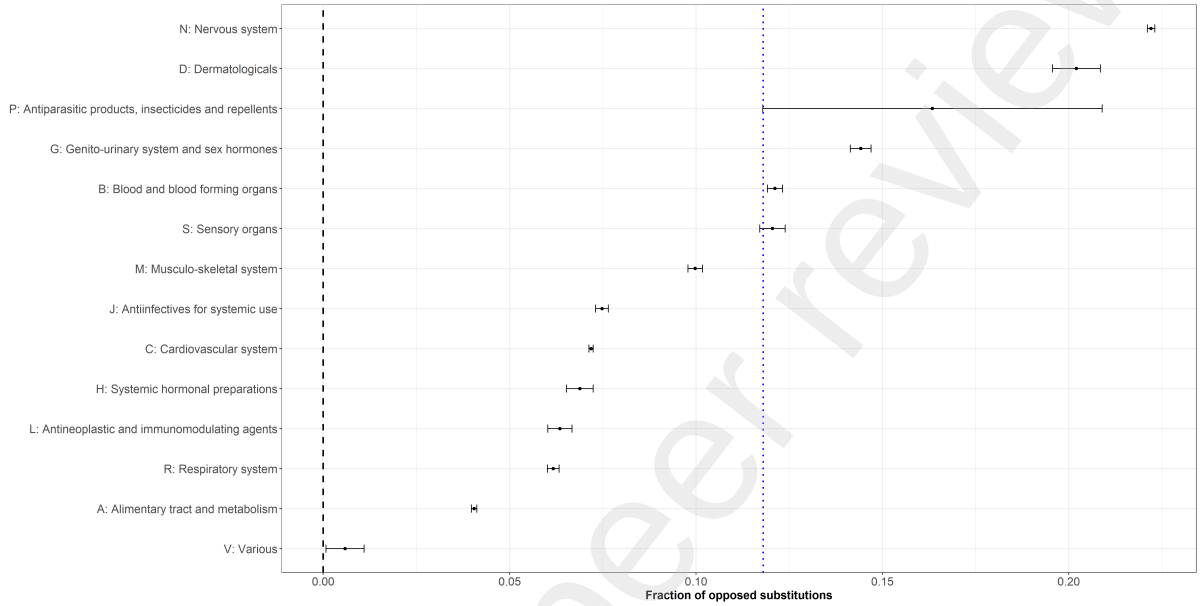
The great majority of patients (91%) in the sample have faced a substitution decision, and around a third of patients opposed substitution at least once. Panel B of Table 2 also shows that patients who opposed substitution at least once have on average opposed substitution on four purchase occasions. The cost for purchases in which a patient opposed substitution is higher than in the whole sample due to the additional out-of-pocket costs, which are on average 19.7 SEK (approximately 2 USD).

The decision to oppose substitution is heterogeneous. In Figure 1 we show the fraction of opposed substitutions conditional on the possibility of substitution across different drug types in different products targeting different anatomical main groups. The graph shows first that opposed substitution is observed in all drug segments. We observe an especially high rate of

[9]The ATC code classifies drugs by their active ingredients.

opposed substitution for drugs of the nervous system, which are mostly painkillers. In comparison, patients using drugs that target the alimentary tract and metabolism oppose substitution to a much lower extent.

Figure 1: Opposed Substitutions by Drug Type



Notes: The graph shows the fraction of opposed substitutions conditional on the possibility of substitution across different drug types. The drug types are ordered according to the first level of the Anatomical Therapeutic Chemical (ATC) classification system. The first level of the code shows the anatomical main group and consists of one letter. The dotted blue line shows the fraction of opposed substitutions for all observed pharmaceuticals conditional on the possibility of substitution. The error bars represent 95% confidence intervals.

Table 2: Summary Statistics

Panel A: Product Level	
	All
Number of Products	4811
Number of Substances	270
Number of Substitution Groups	1347
Avg. Numb. of Competitors in Substitution Group	3.57
	(2.86)
Fraction of Product of the Month	0.38
Avg. Price	254.4
	(604.6)
Avg. Share Originals	0.2
	(0.4)

Panel B: Consumer Level			
	All	Possible Subst.	Opposed Subst.
Purchase Occasions (in thousands)	4513	2299	271
Number of Unique Patients (in thousands)	174	159	69
Avg. Purchase Occasions per Patient within Sample	26	14	4
	(63.9)	(27.1)	(6.7)
Avg. Cost of Purchase	100	119	139
	(251.6)	(274.8)	(200.5)
Avg. Cost from Opposing Substitution	1.3	2.5	19.7
	(12.4)	(17.3)	(44.6)

Note: Summary statistics of the full data set. Panel A shows summary statistics on the product level. An observation is a product in a specific month. Products are ordered in substitution groups, which are groups of medically equivalent pharmaceuticals. The variable price refers to the full price in SEK (10 SEK are approximately 1 USD). Panel B shows summary statistics on the consumer level. An observation is a purchase occasion of a patient. The column for possible substitution refers to the purchase occasions when substitution was possible as the patient received a prescription for a product that was not the cheapest available product. The patient may have substituted or opposed substitution. The column for opposed substitution involves only those purchase occasions where patients opposed substitution and purchased a product that was not the cheapest available option. The variable average cost of purchase refers to the cost to the patient, i.e., the co-payment in SEK. Standard deviations are reported in parentheses.

4 Empirical Strategy

We turn to describing our empirical strategy of showing evidence that an initial substitution affects subsequent decisions. We start by presenting our results in a naive regression and an event study using panel data methods. The initial empirical analysis is intended to show a potential non-causal relationship between a first substitution and subsequent decisions to substitute. Then, we use an instrumental variable approach to show that the relationship is indeed causal.

4.1 Naive Regression

In this initial step, we show the correlation between an initial decision to substitute and subsequent decisions. Initially, we face the challenge that we do not observe dispensing of pharmaceuticals before January 2014. Thus, an individual's decision to substitute in our data could be the individual's first substitution, or a substitution could have already happened in the past and is not observed in our data. Assuming that the initial observed decision is independent of previous decisions could lead to an biased estimate of the correlation between an initial decision to substitute and subsequent decisions. To tackle the challenge, we solely consider individuals with leading purchase occasions in which the patient opposed substitution in our data.

Consider individual i consuming a product in substitution group s at a purchase occasion in month t. Again, we consider the sample of patients facing the decision of substitution, meaning that the prescriber has prescribed a product that is not the cheapest product of the month and neither the prescriber nor the pharmacist has opposed substitution. Consider a naive least square regression model:

$$Opp_{ist} = \beta \mathbb{1}\{AfterFirstSub_{it}\} + \psi_{st} + \alpha_i + \varepsilon_{ist}, \tag{1}$$

where Opp_{ist} is a dummy variable that takes the value 1 if a patient opposes substitution. $\mathbb{1}\{AfterFirstSub_{it}\}$ is an indicator that takes the value 1 if an individual has substituted in the previous observed purchase occasion.[10] ψ_{st} are substitution group $\times$ year-month fixed effects. Additionally, α_i are individual fixed effects. As a result, we use variation on the individual level controlling for factors that are specific for substitution groups within a month. Note that the fixed effects control for price differences between the cheapest available product and other products as prices are fixed and change on a monthly basis.

We estimate the models with different samples. Each sample conditions on the number of

[10] Note that we exclude the first substitution as the outcome is always an oppossed substiution.

purchase occasions with opposed substitutions. Thus we can compare the general correlation of a first substitution and subsequent decisions and decide about a suitable requirement for the number of leading opposed substitutions. A larger number of leading observed purchase occasions with opposed substitution increases the probability that the patients are indeed in an "opposing equilibrium" and never have substituted before. However, restricting the sample size also comes with the cost of losing statistical power. We evaluate the trade-off in this section.

4.2 Event Study

In the following we turn to an event study analysis, which not only allows us to observe an average treatment effect but further shows if a first substitution is related to long-lasting behavioral change. We therefore consider a dynamic version of the model in equation 1, considering a sample with leading purchase occasions with opposed substitution:

$$Opp_{ist} = \sum_{d=-1}^{d=12} \beta_d \mathbb{1}\{PurchOccToFirstSub_{it} = d\} + \psi_{st} + \alpha_i + \varepsilon_{ist}, \tag{2}$$

where Opp_{ist} is a dummy variable that takes the value 1 if a patient opposed substitution. $\mathbb{1}\{PurchOccToFirstSub_{it} = d\}$ is an indicator that takes the value 1 in case of d occasions before or after the first observed substitution. We bin all observations before the first substitution into the indicator $d = -1$ as the outcome is always an opposed substitution. The first substitution is excluded as it is by definition an opposed substitution. We do not require substitution of a specific substance. For example, it is possible that a patient's first substitution is for a product to treat high blood pressure and we subsequently evaluate the decision for a painkiller. Therefore, we are able to look across substances on the individual level. ψ_{st} are substitution group $\times$ year-month fixed effects. Further, α_i are individual fixed effects. Thus, we again control for the effect of monthly price changes.

We expect that purchase occasions after the first substitution are related to a lower probability of opposed substitution. A key question is whether the coefficients of $\hat{\beta}_d$ are stable for the purchase occasions of $d = 1$ to $d = 11$, meaning that we see no further steady decline of the probability to oppose after further purchase occasions. Stable coefficients would speak in favor of a interpretation that a first substitution leads to a decline in opposing substitution, but that repeated experience with substitution does not further lead to learning about equivalence. Declining coefficients would mean instead that each additional substitution leads to additional learning about the equivalence of substances.

4.3 Instrumental Variable Design

The naive OLS regression as well as the event study uses individual variation controlling for factors that are stable within a substitution group within a month. However, also variation on the individual level may be endogenous. As an example, consider a patient that, after frequent opposed substitutions, switches to substituting pharmaceuticals. It may be that an initial substitution itself causes future decisions. However, it may also be the case that other factors affect initial as well as following substitutions. For example, an individual could experience an income shock that increases the possibility of initial substitutions as well as the possibility of following substitutions. In the following we use an instrumental variable approach to tackle the endogeneity threat. Consider the following model, which is equivalent to the approach in equation 1.

$$Opp_{ist} = \beta \mathbb{1}\{AfterFirstSub_{it}\} + \psi_{st} + \alpha_i + \varepsilon_{ist}, \tag{3}$$

where Opp_{ist} is a dummy variable that takes the value 1 if a patient opposed substitution, $\mathbb{1}\{AfterFirstSub_{it}\}$ is an indicator that takes the value 1 if an individual has substituted in a preceding purchase occasion, and ψ_{st} are substitution group $\times$ year-month fixed effects. Additionally, α_i are individual fixed effects. As before, we consider a sample where each patient has leading purchase occasions with opposed substitution. If the assumption of exogenous first substitution is invalid (i.e., $Cov(\varepsilon_{ist}, \mathbb{1}\{AfterFirstSub_{it}\}) \neq 0$), $\hat{\beta}$ may be biased. We therefore consider an instrument Z_{ist}. In detail, we consider instrument $\mathbb{1}\{PrescAfterExit_{it}\}$ which takes the value 1 for individuals that had a prescription for a product that exited the Swedish market and is not available even if a patient wishes to oppose substitution. We observe several substitution groups in which products, often originals, exit a market.[11] We expect that the instrument is correlated to the probability of first substitution. The intuition is the following: Individuals may be used to a specific product. Due to positive experience with this specific product, they do not substitute but oppose substitution repeatedly. After such a product exits the market, these individuals are exogenously forced to substitute.[12] Thus, we expect that $AfterFirstSub_{it}$ is cor-

[11] We define an exit if we do not observe a single dispense after the time of exit. The reason for such exits is usually prices that are too low in a specific market. Other European countries use external reference price systems. The system involves a pricing mechanism based on comparing and weighting prices in multiple countries (see for example Rémuzat et al., 2015, for an overview). Thus, it could be rational to exit a market such as the Swedish market as the overall revenue is too low and low prices in Sweden reduce profits in larger economies.

[12] Note that we observe a few purchase occasions in which individuals receive consecutive prescriptions for a product that has exited. In such cases we only consider the first observation of a forced substitution to avoid an upward bias of our estimate. As a result, the sample size is decreased slightly compared with the naive regression.

related with the instrument $Z_{ist} = \mathbb{1}\{PrescAfterExit_{it}\}$. The analysis recovers the local average treatment effect of those patients that were affected by forced substitution due to the exit of a product.

Besides the strength of the instrument, we build on the exclusion restriction. We assume and argue that is reasonable that the exit of a product is not correlated to a general behavior change of an individual except through the forced substitution. In Appendix A we analyze correlation between exiting products and some market characteristics such as prices or market shares of exiting products before an exit, as well as the general market size and the number of competitors. We do not find any clear trend in any variable before products exit. We argue that this is reassuring as we do not observe any clear sign of exit that could be correlated with a patient's behavior change, except the exit itself.

5 Results

5.1 Naive Regression

In Figure 2 we show results of the main estimate $\hat{\beta}$ of equation 1 for different samples. Each regression includes substitution group $\times$ year-month and individual fixed effects. In detail, we consider individuals with different numbers of leading observed opposed substitutions. Without any restriction we see that a first substitution is related to a strong decrease in the subsequent probability of opposing substitution. If we condition the estimation on leading observed opposed substitutions, the probability of opposing future substitutions increases but remains negative. The confidence interval increases because of the smaller sample size.[13]

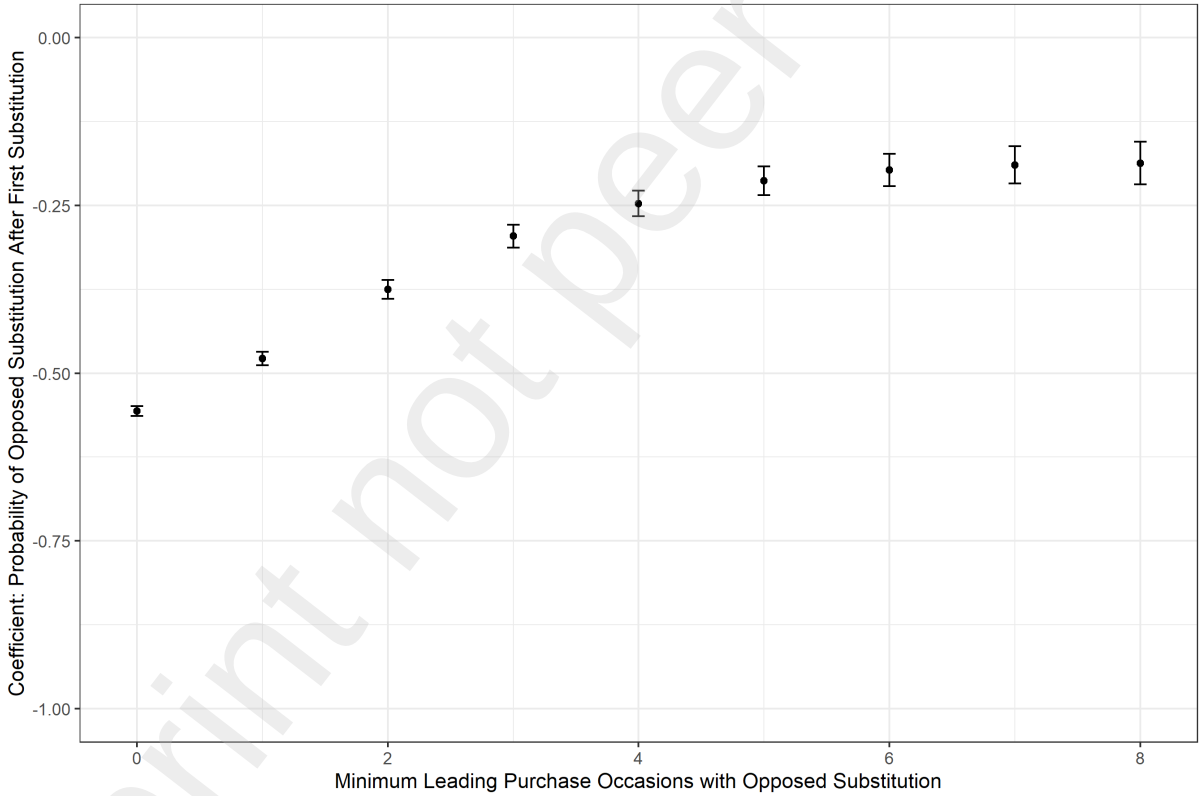
The impact of a first substitution experience in the past converges to an approximately 22-percentage-point lower chance of opposing substitutions in subsequent decisions when considering six to eight leading decisions with opposed substitutions. As we consider patients that have never have substituted before the initial change the 22-percentage-point decrease is equivalent to a change of 22%. Accordingly, we use the condition of six preceding opposed substitutions in the following sections of the article. Within the condition of six preceding periods of opposed substitution we that patients oppose substitution in 82% of the purchase occasions. Due to the smaller sample and a smaller average treatment effect, we believe that this

[13] One explanation for the decrease in effect size is that some individuals do not always oppose substitution but do so occasionally. Reasons could be individual or product specific. When not requiring multiple leading purchase occasions with opposed substitution, we would explain the following repeated substitution decisions with the effect of a first-time substitution. The more often we observe purchase occasions of opposed substitution, the more likely it is that a patient is in a steady state of opposing substitution.

choice is a conservative option. Decreasing the requirement of leading opposed substitution increases the effect size. In Appendix B we show robustness for the requirement of leading purchase occasions but considering fewer number of leading opposed substitutions. The effect size increases.

In Appendix C we show the average treatment effect, $\hat{\beta}$, with six leading purchase occasions with opposed substitution using a variety of fixed effects. Results are robust. Additionally, we show robustness to potential biases due a linear model when using two-way-fixed effect models when treatment is staggered and treatment effects are potentially heterogeneous. In Appendix D we use the estimator from De Chaisemartin and d'Haultfoeuille (2020) to show that the average treatment effect is robust.

Figure 2: Naive Regression with Different Numbers of Leading Purchase Occasions with Opposed Substitution



Notes: The graph shows the coefficients of β for OLS regressions in equation 1. An observation is a substitution decision of an individual when substitution is possible. Each coefficient refers to a sample with different minimum numbers of leading purchase occasions with opposed substitution. The regression includes substitution group × month and individual fixed effects. The error bars represent 95% confidence intervals. Standard errors are clustered at the individual level, adjusted for within-cluster correlation.

5.2 Event Study

Figure 3 shows the results of the regression model in equation 2. The sample consists of patients with at least six leading opposed substitutions. The reference group is the time before substitution, which is an opposed substitution. After the initial first substitution of any product, the probability of opposed substitution decreases substantially by around 15 to 25 percentage points compared with the time before the initial substitution. Further, coefficients of $\hat{\beta}_d$ for $d > 1$ vary within this interval but do not show any trend. Following a first substitution, the probability of opposed substitution is around 20 percentage points lower than before the initial substitution. Thus, we observe that individuals reduce their resistance to substitution after a first substitution. This effect is a very large permanent effect, in the sense that we do not observe further reduction after subsequent purchase occasions.[14]

5.3 Instrumental Variable Design

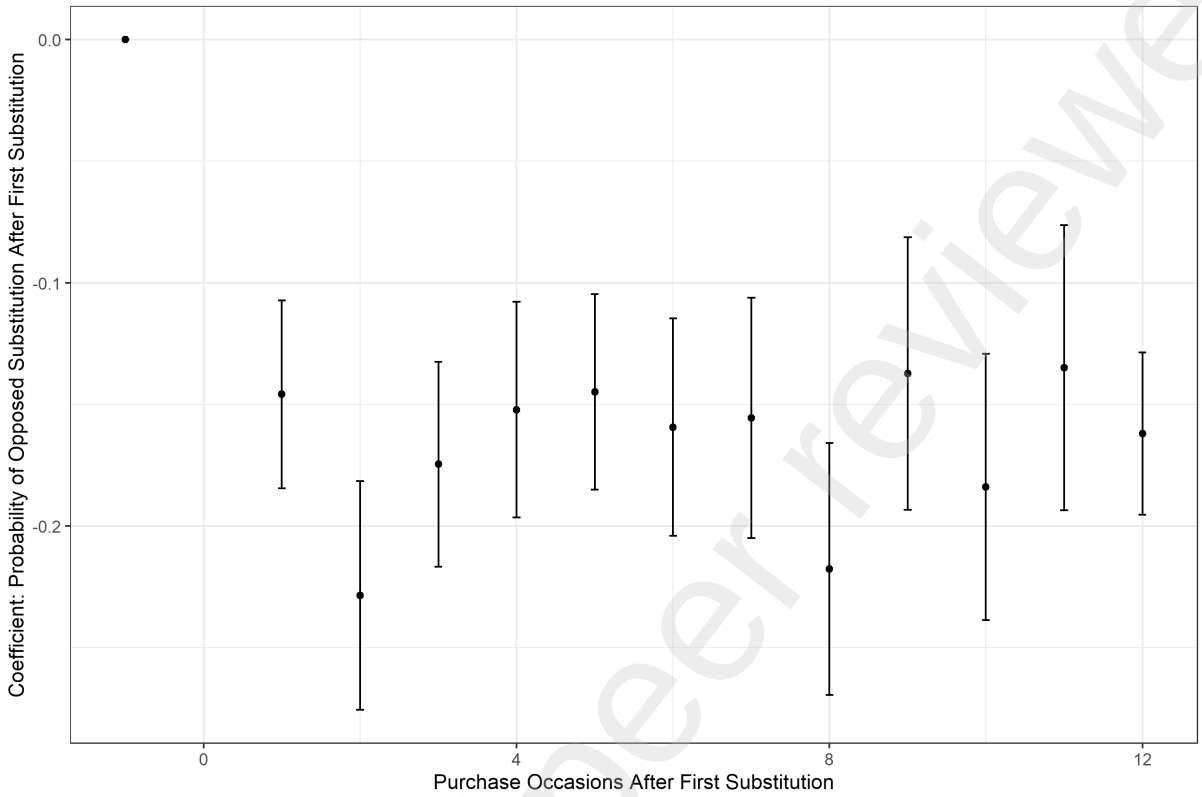
Next we turn to the instrumental variable design. Results of the regression are presented in Table 3. Again, we consider a sample of patients with at least six leading opposed substitutions. Panel A refers to the naive OLS regression.[15] Panel B describes the first stage, while Panel C shows results of the second stage. Considering the instrumental variable approach, we use a single instrument, a dummy that takes the value 1 for individuals with a current or previous prescription for a product that exited the Swedish market and is not available even if the patient wishes to oppose substitution. The F-test allows us to evaluate if the instrument is weak. In all specifications of the first stage, the instrument shows a strong correlation with the endogenous regressor. The second stage confirms a strong effect of the first substitution on the probability of opposing substitution. In detail, a first substitution experience decreases the probability of opposing substitution in the following choices in all specifications. Using individual and substitution group $\times$ year-month fixed effects, a previous first substitution decreases the probability of opposing substitution in the following choices by 39 percentage points.[16] The effect size is larger than the one in the naive regression. Overall, the instrumental variable suggests that the

[14] As for section 5.1, we show robustness for the two-way-fixed effect regression in Appendix D. As econometric literature shows that linear least square regressions that include two-way-fixed effects such as year-month and individual fixed effects could suffer from bias in case of a staggered treatment design and heterogeneous treatment effects across the cohorts, we use an estimator based on Sun and Abraham (2020) to show that results are stable.

[15] Note that the sample of this naive regression differs slightly to the sample of the average treatment effect in Appendix C as we only consider the first observation of a forced substitution to avoid an upward bias of our estimate. We show the instrumental variable regression for the sample without excluding following observations of forced substitution in Appendix E. The effect of the first substitution increases.

[16] As we consider patients that have never have substituted before the initial change the 39-percentage-point decrease is equivalent to a change of 39% on average.

Figure 3: Event Study: Opposed Substitutions Before and After First Substitution



Notes: The graph shows the results of the regression presented in equation 2. One observation corresponds to an individual pharmaceutical decision where an individual either substitutes or opposes substitution. Each patient in the sample has at least six leading decisions with opposed substitution. The outcome variable takes the value 1 if an individual opposed a substitution conditional on facing a substitution decision. Thus, we exclude cases where substitution is not an option, either because the prescription was for the cheapest available product or because the prescriber or pharmacist opposed substitution. In the graph, we show coefficient estimates of $\hat{\beta}_d$*, The estimates show the probability of opposing substitution for the dth purchase occasions after the first substitution, which is excluded. The regression includes substitution group* $\times$ *year-month as well as individual fixed effects. The default is an opposed substitution during the months before the initial substitution. The error bars represent 95% confidence intervals. Standard errors are clustered at the individual level, adjusted for within-cluster correlation.*

naive regression was downwards biased in absolute effect size. In particular, patients with a first substitution are not selected randomly but in a way that makes long-run behavioral changes less likely. Thus, we underestimate the causal effect in the naive regression.

Table 3: Instrumental Variable Regression

Panel A: OLS Regression				
	Opp_{ist}			
	(1)	(2)	(3)	(4)
$\mathbb{1}\{AfterFirstSub_{it}\}$	-0.367^{***}	-0.290^{***}	-0.198^{***}	-0.168^{***}
	(0.021)	(0.018)	(0.015)	(0.013)
Constant	1.000			
Substitution Group FE	No	Yes	Yes	No
Year-Month FE	No	Yes	Yes	No
Individual FE	No	No	Yes	Yes
Substitution × Year-Month FE	No	No	No	Yes
N	11,129	11,129	11,129	11,129
R^2	0.274	0.500	0.672	0.805
Panel B: First Stage				
	$\mathbb{1}\{AfterFirstSub_{it}\}$			
	(1)	(2)	(3)	(4)
$\mathbb{1}\{PrescAfterExit_{ist}\}$	0.676^{***}	0.436^{***}	0.274^{***}	0.274^{***}
	(0.018)	(0.026)	(0.040)	(0.041)
Constant	0.302^{***}			
	(0.017)			
Substitution Group FE	No	Yes	Yes	No
Year-Month FE	No	Yes	Yes	No
Individual FE	No	No	Yes	Yes
Substitution × Year-Month FE	No	No	No	Yes
F-statistics	1420.6	278.2	47.3	44
N	11,303	11,303	11,303	11,303
R^2	0.135	0.388	0.724	0.816
Panel C: Second Stage				
	Opp_{ist}			
	(1)	(2)	(3)	(4)
$\mathbb{1}\{AfterFirstSub_{it}\}$	-0.691^{***}	-0.665^{***}	-0.391^{***}	-0.388^{***}
	(0.070)	(0.089)	(0.102)	(0.097)
Constant	1.113^{***}			
	(0.023)			
Substitution Group FE	No	Yes	Yes	No
Year-Month FE	No	Yes	Yes	No
Individual FE	No	No	Yes	Yes
Substitution × Year-Month FE	No	No	No	Yes
N	11,129	11,129	11,129	11,129
R^2	0.061	0.313	0.650	0.787

$^{*}\ p < 0.1$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Notes: Results of the OLS and instrumental variable regression. An observation is a substitution decision of an individual when substitution is possible. We exclude cases where substitution is not an option, either because the prescription was for the cheapest available product or because the prescriber or pharmacists opposed substitution. Each patient in the sample has at least six leading purchase occasions with opposed substitution. Further, we only consider the first observation of a forced substitution due to an exiting product. Panel A shows an OLS regression. Panel B shows the first stage of the two-stage least square estimation. Here, the instrument $\mathbb{1}\{PrescAfterExit_{ist}\}$ takes the value 1 for individuals with a current or previous prescription for a product that exited the Swedish market and is not available even if the patient wishes to oppose substitution. The F-statistics refer to the test statistics of testing significance of the instrument. Panel B shows results of the second stage. The outcome variable takes the value 1 if an individual opposed a substitution conditional on facing a substitution decision. Standard errors are clustered at the individual level, adjusted for within-cluster correlation, and reported in parentheses.

6 Heterogeneity Analysis

In the following heterogeneity analysis we shed light on differences across gender and age of patients. We therefore subsample our population according to gender and four equally sized age brackets. For each sample we show regression evidence of the naive regression described in equation 1, considering patients with at least six leading decisions with opposed substitution.[17]

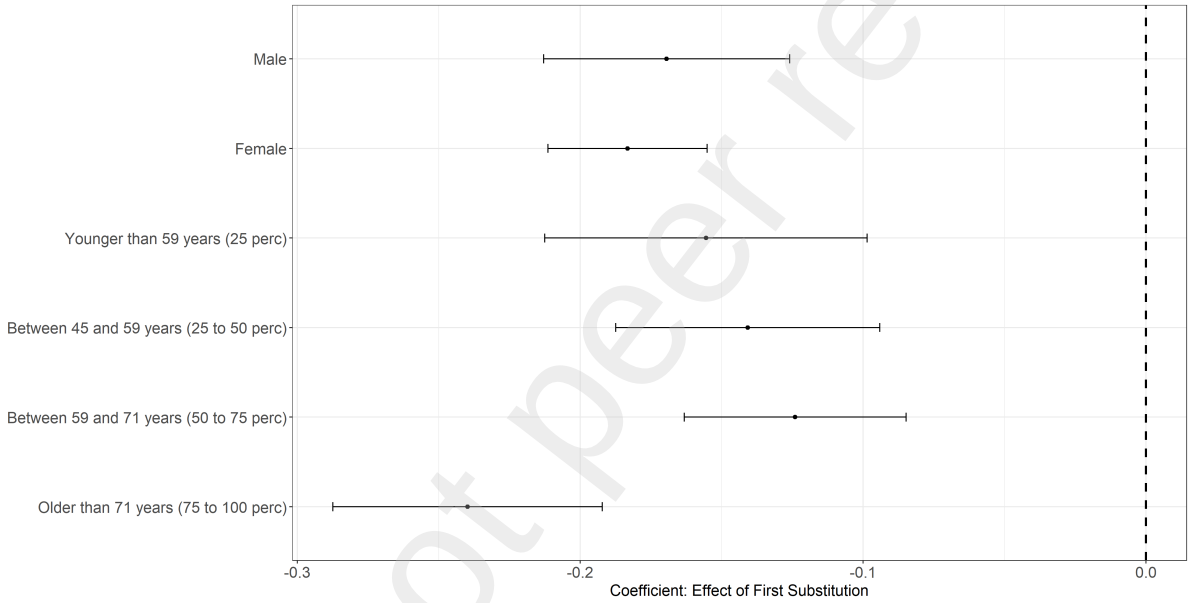
We show the results in Figure 4. As in our general analysis we observe that the results of the instrumental regression are stronger than those of the naive regression. We do not find any clear differences across gender. However, when considering age brackets, the results show that individuals older than 71 years, the top 25% of our sample, have an especially large effect size. The impact of a first substitution on the decrease in the probability of opposing substitutions in subsequent substitution decisions is more than 10 percentage points larger for individuals over 71 years of age compared with those between 59 and 71 years (the second oldest 25% of the sample). Considering only those over 71 years, an initial substitution decision decreases subsequent probabilities that an individual will oppose substitution and pay more to receive the specific product that was prescribed by 24 percentage points. In comparison, we see smaller but stable results for the other age brackets. Thus, we conclude that patients in the oldest age group benefit financially from the first experience with a substitution decision. One potential reason for the effect is that older patients are more risk averse. After having a positive experience of substitution to a cheaper product, they are more likely to stay with the cheapest options in the market.

7 Discussion

We have shown that an initial substitution has a long-lasting effect on further substitution decisions. In detail, the probability of opposing substitution reduces significantly after an initial experience with a substitution. The general decrease in the probability can be rationalized with one-time learning. Patients stick to a pharmaceutical product over longer periods. Such a behavior of state dependence is well documented for the Swedish pharmaceutical market (Granlund, 2021; Janssen, 2020). While we do not distinguish between the reasons for an initial substitution, we document a fundamental change in behavior due to the first substitution. Indeed, the stability of the effect could be due to induced state dependence due to substitution: It may be possible that after an initial substitution a patient now sticks with a newly experienced product.

[17] Note that a sample analysis following the instrumental variable regression of equation 3 is infeasible as we observe insufficient observations of the instrument, meaning that across all samples we do not observe a sufficient number of patients with prescribed products that have exited.

Figure 4: Heterogeneity Analysis: Demographics



Notes: The figure shows estimates of β for a least square regression described in equation 1. An observation is a substitution decision of an individual when substitution is possible. Each patient in the sample has at least six leading decisions with opposed substitution. The outcome variable takes the value 1 if an individual opposed a substitution conditional on facing a substitution decision. Each coefficient refers to an individual regression, divided according to gender or age group. Results for gender are shown for male and female, while we group age according to quartiles. The error bars represent 95% confidence intervals. Standard errors are clustered at the individual level, adjusted for within-cluster correlation.

If the product the patient has substituted to is always the product of the month, state dependence would explain that the patient therefore accepts substitution after an initial substitution. State dependence would then drive our estimate. In Appendix F we control for state dependence and show that all our results – the average treatment effect, the event study and the instrumental variable results – are decreasing in absolute size but are still considerable and large. Thus, we see that part but not all of the effect is due to newly induced state dependence.

7.1 Back-of-the-Envelope Calculation

Our causal estimate using the instrumental variable approach shows that after an initial substitution, patients are approximately 38.8 percentage points less likely to oppose substitution in future decisions. A back-of-the-envelope calculation reveals the strong impact on patients' spending. Considering the sample of those individuals with six leading purchase occasions with opposed substitution, on average the monthly costs of additional co-payments due to opposed substitution are 23.9 SEK (2.39 USD). Compared to the average of 69.5 SEK (6.95 USD) for all co-payments, those excess co-payments account for 34.43%. Considering the reduction of the opposed substitution probability by 38.8 percentage points, a first substitution would reduce the monthly co-payments due to opposed substitution to 14.65 SEK (1.47 USD) on average. Thus the share of co-payments due to opposed substitution would reduce to only 26.7% of all co-payments.[18]

In general, reducing opposed substitution offers the possibility of considerable out-of-pocket savings for patients. Considering the full population of Västerbotten, we observe 2.2 million SEK (200,000 USD) in yearly costs due to opposed substitution. Extrapolating the effect of a 38.8-percentage-point decrease in opposed substitution to all 10 million Swedish patients, yearly cost savings could account for 61 million SEK (6.1 million USD). While the potential effect is relatively small compared to the overall costs for all prescription pharmaceuticals (because prescription pharmaceuticals under patent and without a substitution group are especially high in costs for the public sector), the cost savings could be considerable for co-payments. This basic calculation also does not consider the reaction of firms. It is possible that a higher share of substitution would increase competitive pressure and therefore reduce prices, thereby decreasing the health care costs that patients and health care institutions bear.

[18] In detail, the new co-payments due to opposed substitution would be 14.65 SEK, while all co-payments would account for 54.86 SEK. The counterfactual share of co-payments due to opposed substitution is therefore 14.65/54.86 = 0.267.

7.2 Policy Implications

An individual patient only pays a small additional sum to get a prescribed product instead of a substitution. While one opposed substitution results only in small costs, repeated purchases over months and years increase out-of-pocket expenses and decrease patients' welfare considerably. The brief back-of-the-envelope calculation revealed a large savings potential. The change in behavior after the first substitution leads to the question of how policy makers could encourage the initial substitution.

As the first substitution has such a large effect, we argue that policy makers could further incentivize patients' first substitution. While the most obvious change would be mandatory substitution, we consider in the following only those possible policy changes that still allow consumers to make independent decisions.

A first policy change would be to increase savings for patients for a first substitution, by reducing the co-payment for the cheapest available product. Given a negative price elasticity of demand, this would increase the initial generic uptake. Targeting solely the first substitution reduces the governmental costs and has a long-lasting effect on future substitutions, as our analysis has revealed. The Swedish example allows the implementation of such a policy. Data on individual substitutions are available to governmental institutions and insurers. More than 90% of prescriptions in Sweden are handled electronically (eHälsomyndigheten, 2022), which further facilitates the implementation. In practice, co-payment waivers for the cheapest alternative could be sent to specific patients that have a history of opposed substitution. Those patients would then face larger price differences which would increase the probability of substitution. Also, reducing the co-payment to zero should give an extra strong effect on the probability of substitution according to Ching et al. (2021), who showed that for given absolute price-differences, patients are more likely to choose the cheapest alternative if the co-payment for this product is zero.

Similar to the first proposed policy change, it may be also possible to flag those patients with history of opposed substitution for the attention of pharmacists or physicians. Both could spend additional effort in providing information about the equivalence of drugs. Thereby, the pharmacist or physician would be able to target their efforts and information provision effectively.

References

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

Simon P Anderson, Federico Ciliberto, and Jura Liaukonyte. Information content of advertising: Empirical evidence from the OTC analgesic industry. *International Journal of Industrial Organization*, 31(5):355–367, 2013.

Susan Athey and Guido W. Imbens. Design-based analysis in Difference-In-Differences settings with staggered adoption. *Journal of Econometrics*, April 2021. ISSN 0304-4076. doi: 10.1016/j.jeconom.2020.10.012.

Andrew Baker, David F. Larcker, and Charles C. Y. Wang. How much should we trust staggered difference-in-differences estimates? SSRN Scholarly Paper ID 3794018, Social Science Research Network, Rochester, NY, March 2021.

Mats Bergman, David Granlund, and Niklas Rudholm. Apoteksmarknadens omreglering - effekter på följsamhet, priser och kostnader per dygnsdos. *Tillvaextanalys, Workingpaper 19*, 2012.

Kirill Borusyak, Xavier Jaravel, and Jann Spiess. Revisiting event study designs: Robust and efficient estimation. Working Paper, 2021.

Bart J Bronnenberg, Jean-Pierre Dubé, Matthew Gentzkow, and Jesse M Shapiro. Do pharmacists buy Bayer? Informed shoppers and the brand premium. *The Quarterly Journal of Economics*, 130(4):1669–1726, 2015.

Brantly Callaway and Pedro H. C. Sant'Anna. Difference-in-Differences with multiple time periods. *Journal of Econometrics*, December 2020. ISSN 0304-4076. doi: 10.1016/j.jeconom.2020.12.001.

Nuno Camacho, Bas Donkers, and Stefan Stremersch. Predictably non-bayesian: Quantifying salience effects in physician learning about drug quality. *Marketing Science*, 30(2):305–320, 2011.

Tat Chan, Chakravarthi Narasimhan, and Ying Xie. Treatment effectiveness and side effects: A model of physician learning. *Management Science*, 59(6):1309–1325, 2013.

Andrew Ching and Masakazu Ishihara. The effects of detailing on prescribing decisions under quality uncertainty. *QME*, 8(2):123–165, 2010.

Andrew Ching, David Granlund, and David Sundström. Quantifying the zero-price effect in the field: Evidence from Swedish prescription drug choices. *Journal of the Association for Consumer Research*, 2021.

Andrew T Ching. Consumer learning and heterogeneity: Dynamics of demand for prescription drugs after patent expiration. *International Journal of Industrial Organization*, 28(6):619–638, 2010.

Andrew T Ching, Robert Clark, Ignatius Horstmann, and Hyunwoo Lim. The effects of publicity on demand: The case of anti-cholesterol drugs. *Marketing Science*, 35(1):158–181, 2016.

Pradeep K Chintagunta, Ronald L Goettler, and Minki Kim. New drug diffusion when forward-looking physicians learn from patient feedback and detailing. *Journal of Marketing Research*, 49(6):807–821, 2012.

Andrea Coscelli and Matthew Shum. An empirical model of learning and patient spillovers in new drug entry. *Journal of Econometrics*, 122(2):213–246, 2004.

Gregory S Crawford and Matthew Shum. Uncertainty and learning in pharmaceutical demand. *Econometrica*, 73(4):1137–1173, 2005.

Clément De Chaisemartin and Xavier d'Haultfoeuille. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96, 2020.

eHälsomyndigheten. How prescriptions for medicinal products work in Sweden. *https://www.ehalsomyndigheten.se/other-languages/english/eu/how-prescriptions-work-in-sweden/*, 2022. Accessed: 2022-02-08.

Josh Feng. History-dependence in drug demand: Identification and implications for entry incentives. *Review of Economics and Statistics, Forthcoming*, 2020.

Andrew Goodman-Bacon. Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277, December 2021.

David Granlund. Are private physicians more likely to veto generic substitution of prescribed pharmaceuticals? *Social science & medicine*, 69(11):1643–1650, 2009.

David Granlund. A new approach to estimating state dependence in consumers' brand choices applied to 762 pharmaceutical markets. *The Journal of Industrial Economics*, 69(2):443–483, 2021.

David Granlund and David Sundström. Physicians prescribing originals causes welfare losses. *Economics Letters*, 170:143–146, 2018.

Helle Håkonsen and Else-Lydia Toverud. A review of patient perspectives on generics substitution: What are the challenges for optimal drug use. *Generics and Biosimilars Initiative Journal*, 1(1):28–32, 2012.

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

Benjamin R Handel and Jonathan T Kolstad. Health insurance for "humans": Information frictions, plan choice, and consumer welfare. *American Economic Review*, 105(8):2449–2500, 2015.

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

Toshiaki Iizuka. Physician agency and adoption of generic pharmaceuticals. *American Economic Review*, 102(6):2826–58, 2012.

Aljoscha Janssen. Switching costs, brand premia and behavioral pricing in the pharmaceutical market. 2020.

Michael E Johansen and Caroline Richardson. Estimation of potential savings through therapeutic substitution. *JAMA Internal Medicine*, 176(6):769–775, 2016.

Jonathan D Ketcham, Claudio Lucarelli, Eugenio J Miravete, and M Christopher Roebuck. Sinking, swimming, or learning to swim in Medicare Part D. *American Economic Review*, 102(6):2639–73, 2012.

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

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

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

Dimitra Panteli, Francis Arickx, Irina Cleemput, Guillaume Dedet, Helene Eckhardt, Emer Fogarty, Sophie Gerkens, Cornelia Henschke, Jennifer Hislop, Claudio Jommi, et al. Pharmaceutical regulation in 15 European countries. Review. 2016.

Maria Polyakova. Regulation of insurance with adverse selection and switching costs: Evidence from Medicare Part D. *American Economic Journal: Applied Economics*, 8(3):165–95, 2016.

Cécile Rémuzat, Duccio Urbinati, Olfa Mzoughi, Emna El Hammi, Wael Belgaied, and Mondher Toumi. Overview of external reference pricing systems in Europe. *Journal of Market Access & Health Policy*, 3(1):27675, 2015.

William H Shrank, Niteesh K Choudhry, Jessica Agnew-Blais, Alex D Federman, Joshua N Liberman, Jun Liu, Aaron S Kesselheim, M Alan Brookhart, and Michael A Fischer. State generic substitution laws can lower drug outlays under Medicaid. *Health Affairs*, 29(7): 1383–1390, 2010.

Liyang Sun and Sarah Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, December 2020. ISSN 0304-4076. doi: 10.1016/j.jeconom.2020.09.006.

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

Swedish Medical Product Agency. Comments on the list of substitutable medicinal products. *https://lakemedelsverket.se/english/product/Medicinal-products/Substitution/*, 2010. Accessed: 2019-09-17.

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

TLV. Ändring i tandvårds- och läkemedelsförmånsverkets allmänna råd (tlvar 2003:2) om ekonomiska utvärderingar. *https://www.tlv.se/download/18.467926b615d084471ac31df8/1510316395145/Tillfallig_subvention_tidigare_licenslakemedel.pdf*, 2017. Accessed: 2021-03-04.

Appendix

A Exits of Products

In our instrument variable estimation we use a dummy $\mathbb{1}\{PrescAfterExit_{ist}\}$ that takes the value 1 for individuals with a current or previous prescription for a product that exited the Swedish market and is not available even if the patient wishes to oppose substitution. Thus, we assume that exits of products are as good as random and unrelated to other factors that impact patients' substitution decisions except the first, forced substitution from the exiting product to a new product. In the following we show some results from an analysis on the product level that exiting products are not different to non-exiting products before exiting. Further, substitution groups of exiting products are similar to those without exits. Finally, we also do not observe clear changes of prices or market shares of exiting products before the actual exit. Thus we do not expect that the exit itself is correlated with other changes in patients' behavior.

To analyze if exiting products or substitution groups of exiting products differ, we evaluate within an event study if the behavior of exiting products or substitution groups of exiting products change in key parameters in the months before an exit. The event study thereby allows us further to assess if exiting products or those substitution groups with exiting products change in the months before an exit.

Considering product j in substitution group s in month t, we show the following dynamic event study designs for the product and substitution group level respectively:

$$y_{jst} = \sum_{d=-18}^{d=0} \beta_d \mathbb{1}\{PeriodsBeforeExit_{jt} = d\} + \rho_j + \xi_t + \varepsilon_{jt} \tag{4}$$

$$y_{st} = \sum_{d=-18}^{d=0} \beta_d \mathbb{1}\{PeriodsBeforeExit_{st} = d\} + \mu_s + \xi_t + \varepsilon_{st} \tag{5}$$

where y_{jt} and y_{st} are sets of outcome variables on the product or substitution group level, such as (1) the price of a product, (2) the market share of a product, (3) the market size of a substitution group, or (4) the number of competitors in a substitution group. $\mathbb{1}\{PeriodsBeforeExit_{jt} = d\}$ is a dummy that takes the value 1 if product d is d months before the exit. The same holds for $\mathbb{1}\{PeriodsBeforeExit_{st} = d\}$, considering substitution groups with one exiting product. We take month $d = 0$ as the reference level. ρ_j are product fixed effects, μ_s are substitution group fixed effects, and ξ_t are year-month fixed effects.

We show the results of the four different event studies in Figure A.1. For none of the four different outcomes do we observe a significant tendency of exiting products before their time

of exit in comparison to non-exiting products. As we do not observe any significant trends, we think it is also unlikely that patients would change their behavior suddenly.

B Leading Decisions of Opposed Substitution

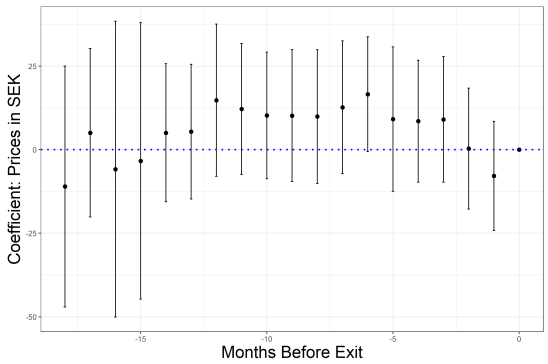
Our empirical analysis is based on a selective sample where we observe at least six purchase occasions with opposed substitution. Our intention is to ensure that we find an effect for those individuals that consistently choose to oppose substitution. Figure 2 shows that fewer purchase occasions usually lead to a larger effect. However, it is possible that some individuals are not consistently choosing to oppose substitution. Instead, it is possible either that six leading decisions with opposed substitution are not sufficient or that the requirement of six leading periods is too strict.

In this section we show that the event study of equation 2 with fewer leading purchase occasions with opposed substitution would increase not only the sample size but also the results. Thus, the results of this paper are a lower bound of the effect size. In Figure B.1 we show the event study after the first substitution for four different requirements: (a) one leading purchase occasion, (b) three leading purchase occasions, (c) six leading purchase occasions, or (d) nine leading purchase occasions with opposed substitution. We observe that results differ and are similar to the average treatment effects shown in Figure 2. The stronger the requirements, the smaller the effect from the first substitution. However, the differences vanish for higher numbers of opposed substitutions, such that we observe almost no differences when comparing the results when restricting the sample to six or when requiring nine leading purchase occasions with opposed substitution. We therefore believe that we use the sample of individuals that are in a steady state of opposed substitution and we measure how the initial observed opposed substitution affects future decisions. In the trade-off between a large sample and a high probability of observing individuals without previous active substitution decisions, we decide to use a requirement of six leading decisions with opposed substitution.

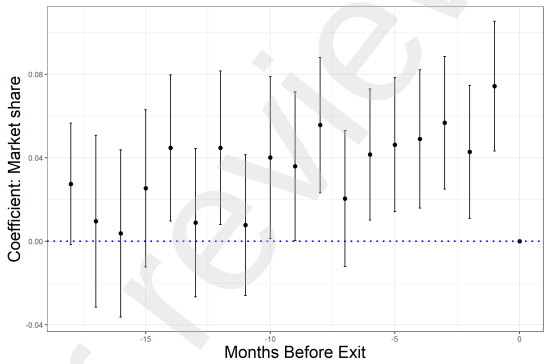
C Average Treatment Effect

In this section we present evidence of the average treatment effect of the least square regression using different fixed effects. In detail, we follow the naive OLS regression presented in equation 1 and evaluate results with different fixed effects. Our sample is the same as the one we use in our event study (equation 2) and the instrumental variable approach (equation 3). Thus, we only consider patients with six preceding purchase occasions with opposed substitution. The

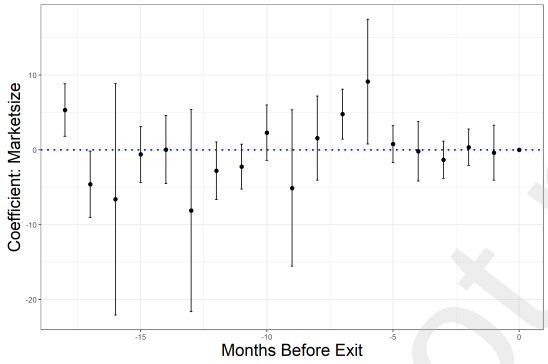
Figure A.1: Event Study: Exiting Products



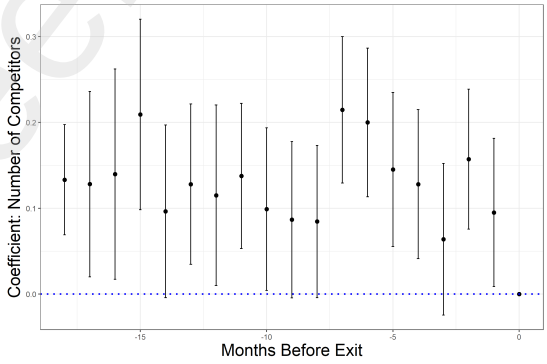
(a) Effect on Prices



(b) Effect on Market Share



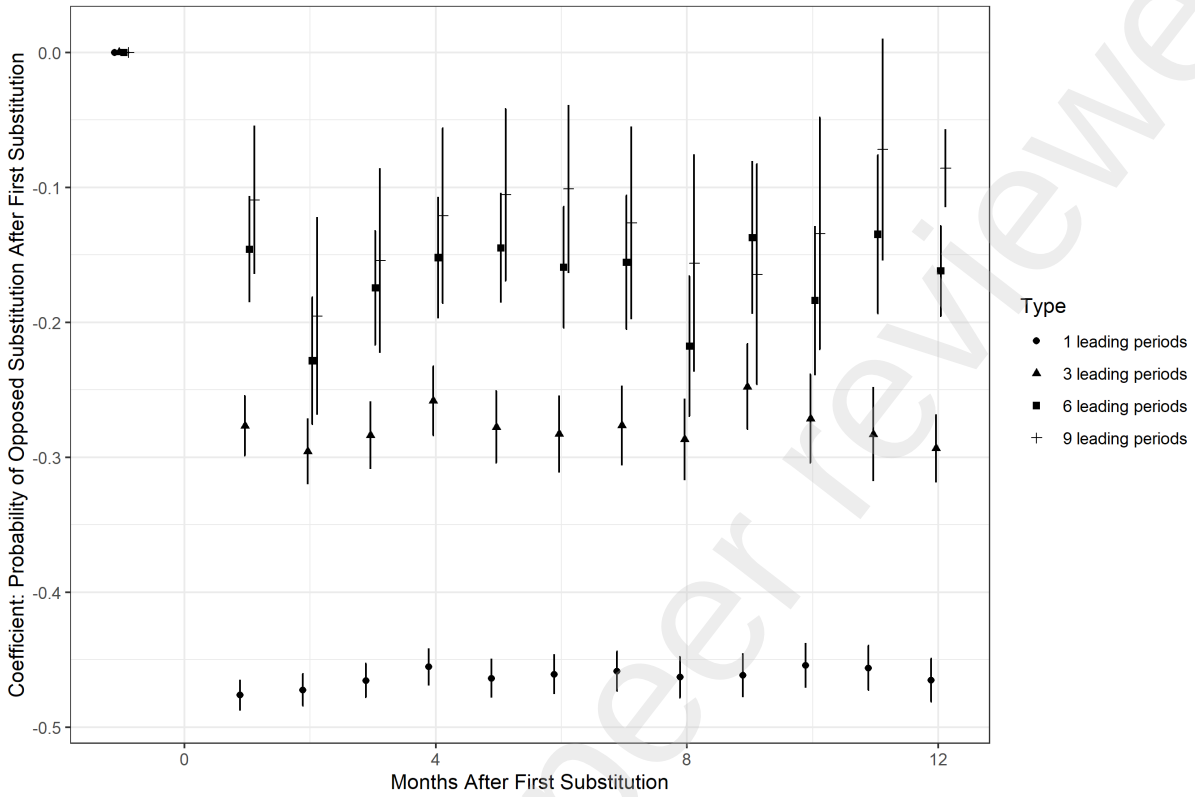
(c) Effect on Market Size



(d) Effect on Number of Competitors

Notes: The figures present coefficients from the event study on the product group level. One observation corresponds to a product within a month for the event study in A.1a and A.1b and to a substitution group within a month for the event study in A.1c and A.1d. We consider four different outcome variables: (1) the price of a product, (2) the market share of a product, (3) the market size of a substitution group, or (4) the number of competitors in a substitution group. The plotted coefficients from $d = -18$ to $d = 0$ correspond to months before the exit. The event study regressions include product and year-month or substitution group and year-month fixed effects. The period $d = 0$ is the reference period. The error bars represent 95% confidence intervals. Standard errors are clustered on the product or substitution group level and adjusted for within-cluster correlation.

Figure B.1: Event Study, Different Number of Leading Decisions of Opposed Substitution



Notes: The graph shows the results of the regression presented in equation 2 for four different samples. For each we require a different number of leading decisions with opposed substitution. One observation corresponds to an individual pharmaceutical decision where an individual either substitutes or opposes substitution. The outcome variable takes the value 1 if an individual opposed a substitution conditional on facing a substitution decision. Thus, we exclude cases where substitution is not an option, either because the prescription was for the cheapest available product or because the prescriber or pharmacist opposed substitution. In the graph, we show coefficient estimates of $\hat{\beta}_d$, The estimates show the probability of opposing substitution for the dth purchase occasions after the first substitution, which is excluded. The regression includes substitution group $\times$ year-month as well as individual fixed effects. The error bars represent 95% confidence intervals. Standard errors are clustered at the individual level, adjusted for within-cluster correlation.

evaluation with different fixed effects allows us to assess whether the treatment effect of a first substitution is stable across specifications.

Considering the decision by patient i for a product in substitution group s in month t, the

final model looks like the following:

$$Opp_{ist} = \beta \mathbb{1}\{AfterFirstSub_{it}\} + \psi_{ist} + \varepsilon_{ist}, \tag{6}$$

where Opp_{ist} is a dummy variable that takes the value 1 if a patient opposed substitution. $\mathbb{1}\{AfterFirstSub_{it}\}$ is an indicator that takes the value 1 if an individual substituted in the last purchase occasion for the first time. ψ_{ist} are a variety of fixed effects. We show models (1) without any fixed effects, (2) substitution group fixed effects, (3) substitution and year-months (i.e., price period) fixed effects, (4) substitution, year-months, and individual fixed effects, (5) individual and substitution group $\times$ year-month fixed effects, and (6) individual and substitution group $\times$ year-month fixed effects where we additionally control if a patient was prescribed an original.

We show results in Table C.1. The average treatment effect over models (1) to (5) decreases with the number of fixed effects. However, despite a high model fit, we observe a strong treatment effect of almost $\hat{\beta} = -0.2$, meaning that a first substitution reduces the subsequent possibility of substitution by 20 percentage points. Controlling for the original prescriptions, we still see an effect size of 15.5 percentage points. Overall, we conclude that the effect of a first substitution is stable.

D Robustness of Two-Way-Fixed-Effect Models

When estimating the average treatment effect of a first substitution, we use a linear least square regression and use two-way-fixed effects. Recent econometric articles show that average treatment effects from linear regressions with individual and time fixed effects could be biased in case of a staggered treatment design and heterogeneous treatment effects (Athey and Imbens 2021; Baker et al. 2021; Borusyak et al. 2021; Callaway and Sant'Anna 2020; De Chaisemartin and d'Haultfoeuille 2020; Goodman-Bacon 2021). To show that our estimates in equation 1 do not suffer from a bias, we therefore also estimate a robust estimator based on De Chaisemartin and d'Haultfoeuille (2020).

Consider for simplicity that we estimate the regression model of equation 1 with individual and year-month fixed effects. We do so by (1) using a standard linear least square estimator and (2) using the estimator of De Chaisemartin and d'Haultfoeuille (2020). Both estimations are based on the sample of patients with at least six preceding purchase occasions of opposed substitution. The average treatment effect of the initial substitution of subsequent substitutions using a two-way-fixed linear regression is equal to $\hat{\beta}^{TWFE} = -0.334$ $(s.e. = 0.016)$, and the

Table C.1: Linear Regression Results

	Opp_{ist}					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}\{AfterFirstSub_{it}\}$	-0.371^{***}	-0.296^{***}	-0.292^{***}	-0.241^{***}	-0.197^{***}	-0.155^{***}
	(0.019)	(0.015)	(0.016)	(0.014)	(0.012)	(0.012)
Substitution Group FE	No	Yes	Yes	Yes	No	No
Year-Month FE	No	No	Yes	Yes	No	No
Individual FE	No	No	No	Yes	Yes	Yes
Subst. $\times$ Year-Month FE	No	No	No	No	Yes	Yes
Original Prescribed	No	No	No	No	No	Yes
N	13,490	13,490	13,490	13,490	13,490	13,195
R^2	0.261	0.482	0.485	0.608	0.749	0.778

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

Notes: Results of a linear least square regression. An observation is a substitution decision of an individual when substitution is possible. Each patient in the sample has at least six leading decisions with opposed substitution. The outcome variable takes the value 1 if an individual opposed a substitution conditional on facing a substitution decision. Thus, we exclude cases where substitution is not an option, either because the prescription was for the cheapest available product or because the prescriber or pharmacist opposed substitution. $\mathbb{1}\{AfterFirstSub_{it}\}$ *is an indicator that takes the value 1 if an individual has substituted in the last period for the first time. Standard errors are clustered at the individual level, adjusted for within-cluster correlation, and reported in parentheses.*

estimate of De Chaisemartin and d'Haultfoeuille (2020) is equal to $\hat{\beta}^{DID-M} - 0.224$ $(s.e. = 0.0345)$. Thus the effect size is smaller but results are still strong, negative and significant.

The potential threat to identification also relates to the estimation of the event study, where we observe staggered treatment on the individual level. We therefore also estimate a robust estimator based on Sun and Abraham (2020). Consider for simplicity the regression as in equation 2 with individual and year-month fixed effects. Figure D.1 shows results from a linear least square regression and a robustness check based on Sun and Abraham (2020). Visually, results are almost identical. Indeed, the coefficients only differ slightly; for example, the coefficient of the purchase occasion after the first substitution in the case of two-way fixed effects is equal to -0.296, while the estimate based on Sun and Abraham (2020) is equal to -0.302. Thus, we conclude that the results do not indicate that the main results are affected by biased estimators.

E Instrumental Variable Regression Including Repeated Forced Substitutions

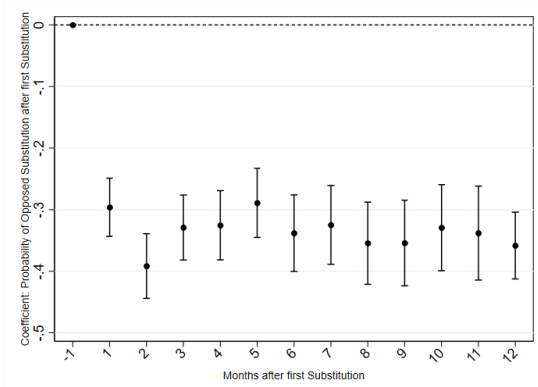
Our main analysis excludes repeated observations of forced substitutions due to exiting products. Thus, for a patient who receives multiple prescriptions for a product exiting the Swedish market, we only consider the first forced substitution. The intuition for reducing the sample is that we do want to avoid overestimating the impact of a first substitution.

In the following, we show robustness for the case without excluding multiple forced substitutions. We, therefore, consider the same sample as in the naive regression presented in Appendix C. Figure E.1 shows the results of the instrumental variable regression. Panel A considers the first stage, and Panel B the second stage. We observe that second stage results are stronger than in the main paper. The difference between the effect in the main analysis and the effect in this robustness check is driven by repeated forced substitutions.

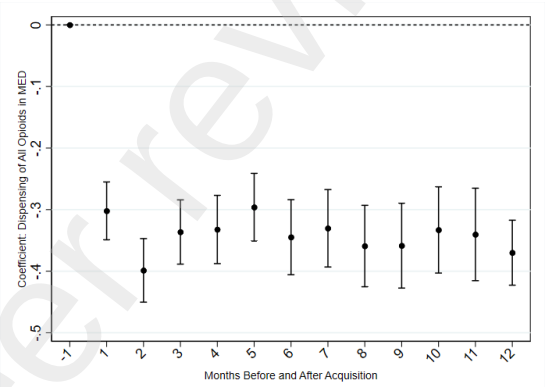
F State Dependence

Another factor that may influence the interpretation of our estimate is habit persistence or state dependence. State dependence describes the behavioral persistence of patients to stick to a pharmaceutical product they know. In our example it is possible that individuals oppose substitution in repeated purchases to receive the product of the previous purchase occasions. At some time, the patient substitutes to a different product, the product of the month. It may be possible that the patient's behavior with regard to substitution does not change and that the individual now

Figure D.1: Event Study: Opposed Substitutions Before and After First Substitution, Robustness



(a) Probability of Opposed Substitution, Two-Way Fixed Effects



(b) Probability of Opposed Substitution, Sun and Abraham (2020)

Notes: The graphs show the results of the regression presented in equation 2. The left graph is based on a linear least square regression. One observation corresponds to an individual pharmaceutical decision where an individual either substitutes or opposes substitution. Each patient in the sample has at least six leading decisions with opposed substitution. The outcome variable takes the value 1 if an individual opposed a substitution conditional on facing a substitution decision. Thus, we exclude cases where substitution is not an option, either because the prescription was for the cheapest available product or because the prescriber or pharmacist opposed substitution. In the graph, we show coefficient estimates of $\hat{\beta}_d$*. The estimates show the probability of opposing substitution for the dth purchase occasions after the first substitution, which is excluded. The regression includes year-month as well as individual fixed effects. The default is an opposed substitution during the months before the initial substitution. The error bars represent 95% confidence intervals. The second subfigure shows results from the method based on Sun and Abraham (2020). Standard errors are clustered at the individual level to control for within-cluster correlation.*

Table E.1: Instrumental Variable Regression, Considering Repeated Forced Substitutions

Panel A: First Stage				
	$\mathbb{1}\{AfterFirstSub_{it}\}$			
	(1)	(2)	(3)	(4)
$\mathbb{1}\{PrescAfterExit_{ist}\}$	0.667***	0.454***	0.261***	0.259***
	(0.017)	(0.025)	(0.037)	(0.032)
Constant	0.320***			
	(0.017)			
Substitution Group FE	No	Yes	Yes	No
Year-Month FE	No	Yes	Yes	No
Individual FE	No	No	Yes	Yes
Substitution × Year-Month FE	No	No	No	Yes
F-statistics	1522.5	330.8	48.6	63.9
N	13,490	13,490	13,490	13,490
R^2	0.202	0.423	0.731	0.814
Panel B: Second Stage				
	Opp_{ist}			
	(1)	(2)	(3)	(4)
$\mathbb{1}\{AfterFirstSub_{it}\}$	−0.608***	−0.573***	−0.681***	−0.699***
	(0.051)	(0.062)	(0.096)	(0.101)
Constant	1.095***			
	(0.018)			
Substitution Group FE	No	Yes	Yes	No
Year-Month FE	No	Yes	Yes	No
Individual FE	No	No	Yes	Yes
Substitution × Year-Month FE	No	No	No	Yes
N	13,490	13,490	13,490	13,490
R^2	0.155	0.389	0.505	0.657

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

Notes: Results of the instrumental variable regression. An observation is a substitution decision of an individual when substitution is possible. Each patient in the sample has at least six leading decisions with opposed substitution. Thus, we exclude cases where substitution is not an option, either because of a prescription was for the cheapest available product or because the prescriber or pharmacist opposed substitution. For this robustness we consider all and not only the first observation of a forced substitution. Panel A shows the first stage of the two-stage least square estimation. Here, the instrument $\mathbb{1}\{PrescAfterExit_{ist}\}$ *takes the value 1 for individuals with a current or previous prescription for a product that exited the Swedish market and is not available even if a patient wishes to oppose substitution. The F-statistics refers to the test statistics of testing significance of the instrument. Panel B shows results of the second stage. The outcome variable takes the value 1 if an individual opposed a substitution conditional on facing a substitution decision. Standard errors are clustered at the individual level, adjusted for within-cluster correlation, and reported in parentheses.*

sticks with the new product. If the new product is always the product of the month and the patient therefore accepts the substitution, state dependence drives our estimate.

State dependence is therefore one mechanism that could drive the estimates. We show this in a robustness check of the average treatment effect, the event study and the instrumental variable approach when controlling for state dependence. In practice we include an indicator that takes the value 1 if the prescribed product is the same as the product that the individual bought the last time the individual made a purchase in a specific substitution group. If an individual only substitutes more often due to state dependence and we assume a persistence in preference for the product of the month, we would expect a reduction of the effect size.

We start by showing the average treatment effect, which corresponds to the results of equation 6. Considering the decision of patient i for a product in substitution group s in month t, the final model looks like the following:

$$Opp_{ist} = \beta \mathbb{1}\{AfterFirstSub_{it}\} + SD_{ist} + \psi_{ist} + \alpha_i + \varepsilon_{ist}, \tag{7}$$

where Opp_{ist} is a dummy variable that takes the value 1 if a patient opposes substitution. $\mathbb{1}\{AfterFirstSub_{it}\}$ is an indicator that takes the value 1 if an individual has substituted in the last purchase occasion for the first time.[19] In comparison to the average treatment effect of equation 6, we now include a dummy SD_{ist} which takes the value 1 if the prescribed product is the same as the product that the individual bought the last time the individual made a purchase in a specific substitution group. ψ_{ist} are a variety of fixed effects. As before, we show models (1) without any fixed effects, (2) with substitution group fixed effects, (3) with substitution and year-months (i.e., price period) fixed effects, (4) with substitution, year-months and individual fixed effects, (5) with individual and substitution group $\times$ year-month fixed effects, (6) and with individual and substitution group $\times$ year-month fixed effects where we additionally control if a patient was prescribed an original.

We show results in Table F.1. In all specifications we see that the state dependence dummy is positive and significant, meaning that the product the consumer bought last time positively affects the probability of opposed substitution. Importantly, our estimated effect of a first substitution on subsequent substitutions, $\mathbb{1}\{AfterFirstSub_{it}\}$, remains negative and significant but is smaller in size. Thus, part but not all of the effect of the first substitution is driven by introduction of a new state dependence. In detail, using full fixed effects in model (5), we show that even after controlling for state dependence, a first substitution decreases the probability of opposed substitution by 11 percentage points. Overall we conclude that state dependence is one

[19] Note that we exclude the first substitution as it does not provide any variation. We require at least six preceding purchase occasions with opposed substitution before the initial substitution.

mechanism; however, it does not explain the entire average treatment effect.

Table F.1: Linear Regression Results, Controlling for State Dependence

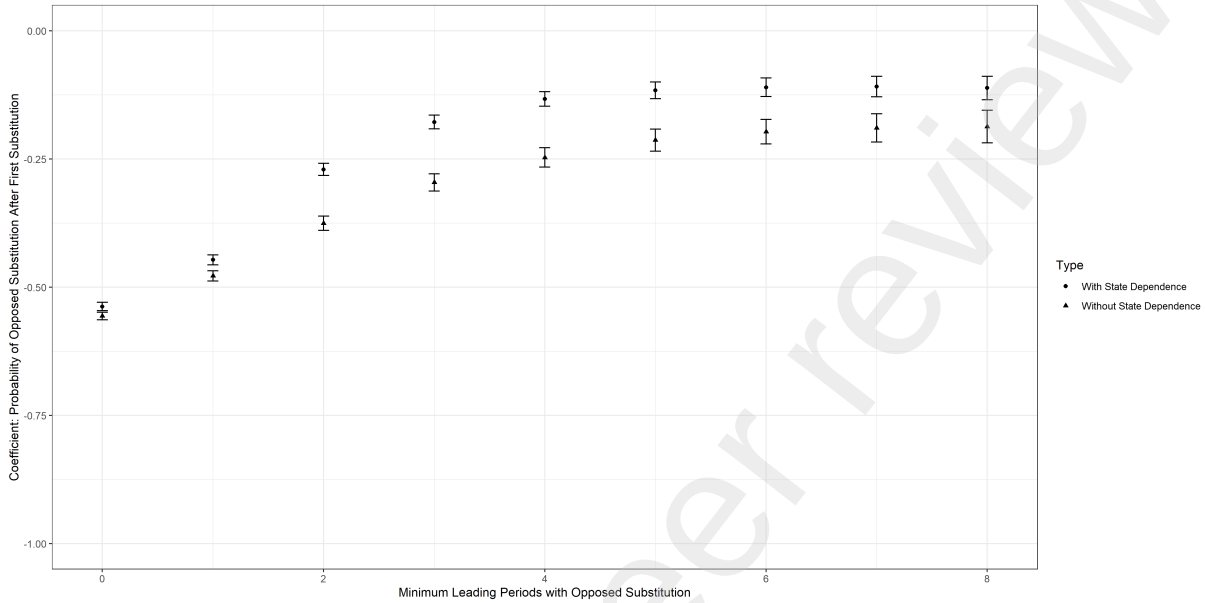
	Opp_{ist}					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}\{AfterFirstSub_{it}\}$	−0.280***	−0.252***	−0.188***	−0.138***	−0.110***	−0.095***
	(0.011)	(0.010)	(0.010)	(0.011)	(0.009)	(0.009)
SD_{ist}	0.390***	0.303***	0.364***	0.318***	0.328***	0.264***
	(0.016)	(0.013)	(0.014)	(0.012)	(0.013)	(0.013)
Substitution Group FE	No	Yes	Yes	Yes	No	No
Year-Month FE	No	No	Yes	Yes	No	No
Individual FE	No	No	No	Yes	Yes	Yes
Subst. × Year-Month FE	No	No	No	No	Yes	Yes
Original Prescribed	No	No	No	No	No	Yes
N	13,490	13,490	13,490	13,490	13,490	13,195
R^2	0.506	0.601	0.626	0.697	0.812	0.820

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

Notes: Results of a linear least square regression. An observation is a substitution decision of an individual when substitution is possible. Each patient in the sample has at least six leading decisions with opposed substitution. The outcome variable takes the value 1 if an individual opposed a substitution conditional on facing a substitution decision. Thus, we exclude cases where substitution is not an option, either because the prescription was for the cheapest available product or because the prescriber opposed substitution. $\mathbb{1}\{AfterFirstSub_{it}\}$ *is an indicator that takes the value 1 if an individual has substituted in the last period for the first time.* SD_{ist} *takes the value 1 if the prescribed product is the same as the product that the individual bought the last time the individual made a purchase in a specific substitution group. Standard errors are clustered at the individual level, adjusted for within-cluster correlation, and reported in parentheses.*

In a second step, we turn to the analysis in section 4.1. In detail, we show that independent of the requirements of leading purchase occasions with opposed substitution, the impact of state dependence does not eliminate the general average treatment effects of a first substitution. In Figure F.1 we show the results of a naive regression as described in equation 1 for different numbers of preceding purchase occasions of opposed substitution. We show results with and without the state dependence dummy. The results are in line with the average treatment effects of the models in Table F.1. The control of state dependence decreases the effect size, but independent of the number of leading purchase occasions with opposed substitution, the effect on the probability of opposed substitution is negative and significant.

Figure F.1: Naive Regression with Different Numbers of Leading Decisions of Opposed Substitution, Controlling for State Dependence



Notes: The graph shows the coefficients of β for OLS regressions in equation 1. An observation is a substitution decision of an individual when substitution is possible. We differentiate between results with and without controlling for state dependence. When controlling for state dependence we include a dummy that takes the value 1 if the prescribed product is the same as the product that the individual bought the last time the individual made a purchase in a specific substitution group. Each coefficient refers to a sample with different minimum numbers of leading purchase occasions with opposed substitution. The regression includes substitution group $\times$ month and individual fixed effects. The error bars represent 95% confidence intervals. Standard errors are clustered at the individual level, adjusted for within-cluster correlation.

We now turn to the event study with a state dependence control. The model is identical to the one in equation 2 except for the introduction of a state dependence dummy. Considering a sample with at least six leading purchase occasions with opposed substitution, we estimate:

$$Opp_{ist} = \sum_{d=2}^{d=11} \beta_d \mathbb{1}\{PurchOccToFirstSub_{it} = d\} + SD_{ist} + \psi_{st} + \alpha_i + \varepsilon_{ist}, \tag{8}$$

where Opp_{ist} is a dummy variable that takes the value 1 if a patient opposed substitution. $\mathbb{1}\{PurchOccToFirstSub_{it} = d\}$ is an indicator that takes the value of d occasions before or after the first observed substitution. SD_{ist} takes the value 1 if the prescribed product is the same

as the product that the individual bought the last time the individual made a purchase in a specific substitution group. ψ_{st} are substitution group $\times$ year-month fixed effects. Further, α_i are individual fixed effects.

The results in Figure F.2 in comparison to Figure 3 show that the impact for purchase occasions after the first substitution decreases, but remains negative and significant when including the state dependence dummy.

Finally we turn to the instrumental variable regression. Using the same instrument as described in equation 3, we add a dummy $\mathbb{1}\{PrescAfterExit_{ist}\}$ that takes the value 1 for individuals with a current or previous prescription for a product that exited the Swedish market and is not available even if a patient wishes to oppose substitution. In comparison to the model in equation 3 we additionally add the control SD_{ist}.

We present the results in Table F.2. The first stage shows that the instruments are still strong. The second stage again confirms the results of the OLS regressions. Controlling for state dependence decreases the effect size of the impact of a first substitution. However, results remain strong and significant. For example, model (4) of the instrumental variable estimation reveals that a first substitution decreases the probability of following opposed substitutions by 49.9 percentage points.

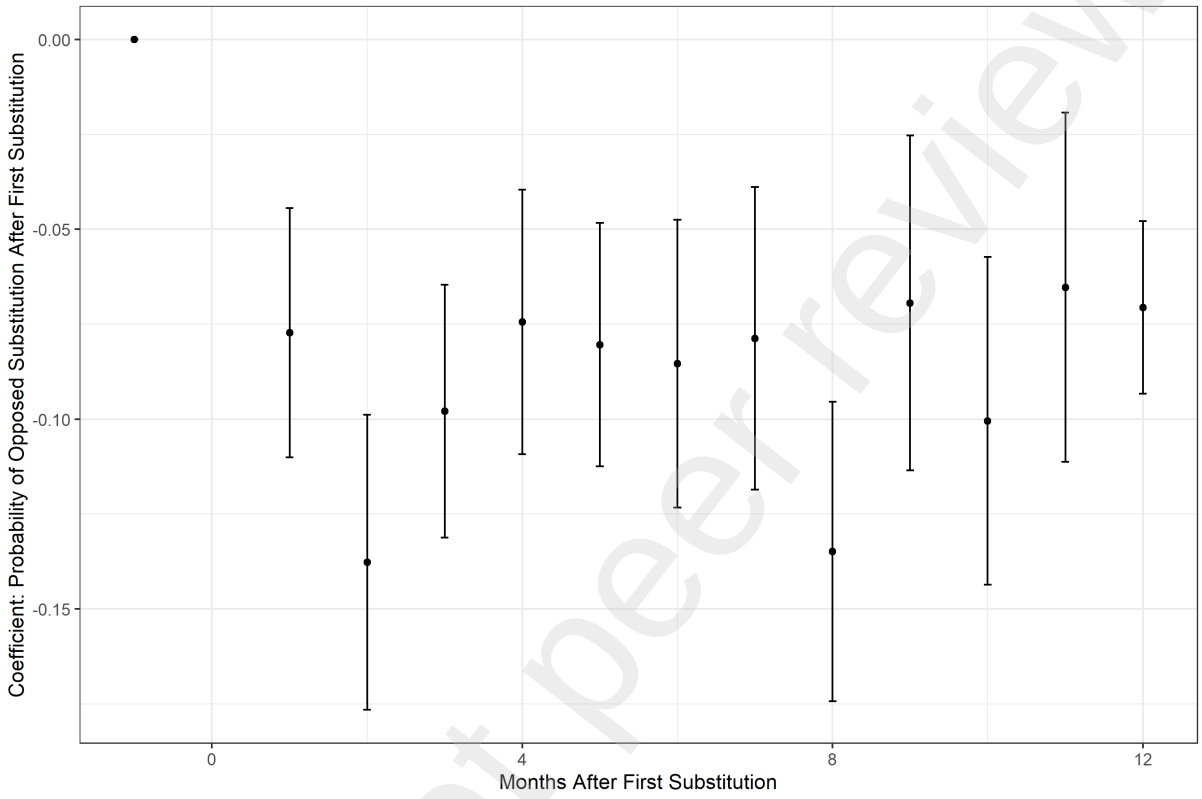
Table F.2: Instrumental Variable Regression, Controlling for State Dependence

Panel A: First Stage				
	$\mathbb{1}\{AfterFirstSub_{it}\}$			
	(1)	(2)	(3)	(4)
$\mathbb{1}\{PrescAfterExit_{ist}\}$	0.613***	0.351***	0.215***	0.218***
	(0.021)	(0.028)	(0.040)	(0.041)
SD_{ist}	−0.194***	−0.285***	−0.179***	−0.156***
	(0.018)	(0.016)	(0.010)	(0.011)
Constant	0.445***			
	(0.021)			
Substitution Group FE	No	Yes	Yes	No
Year-Month FE	No	Yes	Yes	No
Individual FE	No	No	Yes	Yes
Substitution × Year-Month FE	No	No	No	Yes
F-statistics	858	160.4	29.5	28.5
N	11,129	11,129	11,129	11,129
R^2	0.160	0.433	0.739	0.824
Panel B: Second Stage				
	Opp_{ist}			
	(1)	(2)	(3)	(4)
$\mathbb{1}\{AfterFirstSub_{it}\}$	−0.546***	−0.529***	−0.275***	−0.280***
	(0.059)	(0.088)	(0.093)	(0.080)
SD_{ist}	0.284***	0.221***	0.224***	0.229***
	(0.021)	(0.029)	(0.023)	(0.019)
Constant	0.865***			
	(0.030)			
Substitution Group FE	No	Yes	Yes	No
Year-Month FE	No	Yes	Yes	No
Individual FE	No	No	Yes	Yes
Substitution × Year-Month FE	No	No	No	Yes
N	11,129	11,129	11,129	11,129
R^2	0.360	0.486	0.717	0.832

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

Notes: Results of the instrumental variable regression. An observation is a substitution decision of an individual when substitution is possible. Each patient in the sample has at least six leading decisions with opposed substitution. Thus, we exclude cases where substitution is not an option, either because of a prescription was for the cheapest available product or because the prescriber or pharmacist opposed substitution. Further, we only consider the first observation of a forced substitution. Panel A shows the first stage of the two-stage least square estimation. Here, the instrument $\mathbb{1}\{PrescAfterExit_{ist}\}$ takes the value 1 for individuals with a current or previous prescription for a product that exited the Swedish market and is not available even if a patient wishes to oppose substitution. The F-statistics refers to the test statistics of testing significance of the instrument. SD_{ist} takes the value 1 if the prescribed product is the same as the product that the individual bought the last time the individual made a purchase in a specific substitution group. Panel B shows results of the second stage. The outcome variable takes the value 1 if an individual opposed a substitution conditional on facing a substitution decision. Standard errors are clustered at the individual level, adjusted for within-cluster correlation, and reported in parentheses.

Figure F.2: Event Study: Opposed Substitutions Before and After First Substitution, Controlling for State Dependence



Notes: The graph shows the results of the regression presented in equation 8. One observation corresponds to an individual pharmaceutical decision where an individual either substitutes or opposes substitution. Each patient in the sample has at least six leading decision of opposed substitutions. The outcome variable takes the value 1 if an individual opposed a substitution conditional on facing a substitution decision. Thus, we exclude cases where substitution is not an option, either because the prescription was for the cheapest available product or because the prescriber or pharmacist opposed substitution. In the graph, we show coefficient estimates of $\hat{\beta}_d$*. The estimates show the probability of opposing substitution for the dth purchase occasions after the first substitution, which is excluded. The regression includes substitution group* $\times$ *year-month as well as individual fixed effects as well as a dummy* SD_{ist} *that takes the value 1 if the prescribed product is the same as the product that the individual bought the last time the individual made a purchase in a specific substitution group. The default is an opposed substitution during the months before the initial substitution. The error bars represent 95% confidence intervals. Standard errors are clustered at the individual level, adjusted for within-cluster correlation.*